ADDIS ABABA UNIVERSITY

COLLEGE OF NATURAL AND COMPUTATIONAL SCIENCES

DEPARTMENT OF STATISTICS

ANALYSIS OF HUMAN DEATHS BY ROAD TRAFFIC ACCIDENT IN OROMIA REGION, ETHIOPIA

by
Merga Abdissa

Advisor: Mekonnen Tadesse (Associate Prof.)

A Thesis submitted to the Department of Statistics, Addis Ababa University, in partial fulfillment of the requirements for the degree of Master of Science in Statistics.

Addis Ababa University

Addis Ababa, Ethiopia

June, 2018

ADDIS ABABA UNIVERSITY

COLLEGE OF NATURAL AND COMPUTATIONAL SCIENCES

DEPARTMENT OF STATISTICS

Analysis of Human Deaths by Road Traffic Accident in Oromia Region, Ethiopia

Merga Abdissa

A Thesis Submitted to the Graduate programs of Addis Ababa University in Partial Fulfillment of the Requirements for the Degree of Master of Science in Statistics (Applied Statistics)

Advisor: Mekonnen Tadesse (Associate Prof.)

Addis Ababa University
Addis Ababa, Ethiopia

June, 2018

Addis Ababa University

Department of Statistics

This is to certify that the thesis prepared by Merga Abdissa , entitled: Analysis of human deaths by road traffic accident in Oromia region, Ethiopia and submitted in partial fulfillment of the requirements for the Degree Master of Science in Statistics complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the Examining Committee:

	Name	Signature	Date
Advisor:	______________________	_________	___________
Examiner:	______________________	_________	___________
Examiner:	______________________	_________	___________

DECLARATION

I, the undersigned, declare that the thesis is my original work, has not been presented for degrees in any other University and all sources of materials used for the thesis have been duly acknowledged.

Name: **Merga Abdissa**

Signature: **___________________________**

Place of submission: Department of statistics, College of Natural and Computational Sciences, Addis Ababa University

Date of submission: June, 2018

This thesis has been submitted for examination with my approval as a University advisor.

Mekonnen Tadesse (Assoc. Prof.) ______________________________

Advisor's Name **signature**

Acknowledgments

Glory is to the almighty God for making me finish one chapter of my life and for giving me endurance and health to successfully deal with my educational career.

I would, above all, like to express my heartfelt gratitude to my advisor, Mekonnen Tadesse(Associate. Prof.), for the guidance rendered to me in the work of this thesis into what it is today. I also thank for taking his time to advice through my work, as well as the criticisms and the suggestions. His encouragements and welcoming approach at any time will never be forgotten.

Secondly, my deepest thanks go to all members at Addis Ababa University especially, my instructors and all individuals in the Department of Statistics for their moral support and positive attitude when I needed them during my course work, and research period.

Once again, my heartfelt gratitude goes to all Oromia Police Commission bureau department of traffic control and investigation staff members for they were ever willing to provide me road traffic accident data.

Special thanks go to my roommates for sharing constructive ideas throughout the writing of this thesis.

I would like to extend my thanks to my best friends, Moti, Amanu, Eran, Shime and all staff members of Adea Berga Woreda Road Authority office especially Makasa, for all their encouragement and support throughout my study.

This study would have not been possible indeed without the incalculable encouragement and support offered by my wife, Bontu Tamiru, thank you Boni!.

Table of Contents

List of Tables

List of Figures

List of Acronyms

AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
CSA	Central Statistical Agency
HNB	Hurdle Negative Binomial
HP	Hurdle Poisson
GLM	Generalized linear model
GDP	Gross Domestic Product
IRR	Incidence Rate Ratio
LRT	Likelihood Ratio Tests
NB	Negative Binomial
RTA	Road Traffic Accident
UNECA	United Nation Economic Commission for Africa
V	Vuong
WHO	World Health Organization
ZI	Zero-Inflated
ZIP	Zero-Inflated Poisson
ZINB	Zero-Inflated Negative Binomial

Analysis of Human Deaths by Road Traffic Accident in Oromia Region, Ethiopia.

Merga Abdissa

ABSTRACT

Globally, road traffic accidents are leading causes of death among young people in general, and the main cause of death among those aged 15–29 years, in particular. In Ethiopia, the average number of road traffic injuries has been increasing from year to year, where Oromia regional State is the major contributor of the total fatalities occurred. The aim of this study is to identify the major factors determining the number of human death by road traffic accident in Oromia Regional State of Ethiopia using an appropriate count regression model based on data obtained from the Oromia Police Commission Bureau that have been recorded on daily basis during July 8, 2016 - July 7, 2017. The descriptive result showed that 2,712 out of 3,900 of the accidents were not fatal accidents. All formal statistical tests indicated that the data are overdispersed and that there are excess zeros. Among families of count models, the hurdle Poisson (HP) model was found to be the most appropriate to the dataset. The results of hurdle Poisson regression model showed that age of driver, experience of driver, type of vehicle, vehicle service, road condition, time of accident, location of accident, environment of accident, type of accident and accident cause were significant factors influencing the number of human death per accident in Oromia regional state of Ethiopia.

CHAPTER ONE: INTRODUCTION

1.1 Background of the study

Road traffic accidents constitute a major public health and development crisis. Over 1.2 million people die each year on the world's roads, with millions more sustaining serious injuries and living with long-term adverse health consequences. Globally, road traffic crashes are a leading cause of death among young people, and the main cause of death among those aged 15–29 years. However, low-income countries have fatality rates more than double those in high-income countries and there are a disproportionate number of deaths relative to these countries' level of motorization: 90% of road traffic deaths occur in low-and middle-income countries, yet these countries have just 54% of the world's vehicles (WHO, 2015).

Road traffic injuries place a heavy burden on national economies as well as on households. In low- and middle income countries they particularly affect the economically active age group, or those set to contribute to family and the workforce in general. Many families are driven deeper into poverty by the loss of a breadwinner, or by the expenses of prolonged medical care, or the added burden of caring for a family member who is disabled as a result of road traffic injury. The economic costs also strike hard at a national level, imposing a significant burden on health, insurance and legal systems. This is particularly true in countries struggling with other development needs, where investment in road safety is not commensurate with the scale of the problem. Data suggest that road traffic deaths and injuries cause economic losses up to 3% of gross domestic product (GDP) globally whereas in low- and middle-income countries they are estimated to 5% of GDP (WHO, 2015).

The African Region continues to have the highest road traffic death rates. The situation of road traffic accidents (RTA) is most severe in Sub-Saharan Africa where the lives of millions are lost and significant amount of property is damaged. In Ethiopia, the situation has been worsened as the number of vehicles has increased and consequently due to increased traffic flow and conflicts between vehicles and pedestrians. Despite government efforts in the road development, road crashes remain to be one of the critical problems of the road transport sector in Ethiopia (UNECA, 2009). Every year many lives are lost and much property is destroyed due to road traffic accidents in the country. The country has experienced average annual road accidents of 8115 during 2000/01-2010/11 (CSA, 2011).

Ethiopia is one among many low-income countries in Africa and, like other low-income countries it has a high rate of road traffic injury and death (Persson, 2008). Road accident in Ethiopia is one of the worst accident records in the world, as expressed per 10,000 vehicles. Moreover, road accidents are concentrated in Addis Ababa, which is the capital city of Ethiopia, and Oromia region accounting for 58 per cent of all fatalities and two-third of all injuries (Getu, 2007). During 2006-2015 the average of road traffic injuries increased in Ethiopia, where Oromia regional state is the major contributor of total fatalities occurred and Addis Ababa city administration is a major contributor of serious and slight injuries as well as property damages (Amanuel, 2017).

To expand our understanding about the most common and consistent factors on the risk of human death due to road traffic accident, we have considered possible determinants of human death due to road traffic accident using count regression model. Therefore, this study explores the human, environmental, road and vehicle characteristics associated with human death by road traffic accident in Oromia regional state of Ethiopia.

1.2 Statement of the problem

Ethiopia is one of the developing countries with low level of income coupled with high rate of population growth, Oromia is the largest region in Ethiopia with an area of 284,508 km^2 and a total population in 2018 estimated to more than 37 Million (CSA,2013). According to Amanuel (2017), Road traffic accident in Ethiopia is showing a rising trend over the last decades. Moreover, road accidents are concentrated in Oromia which is the largest region and Addis Ababa which is the capital city of Oromia region and Ethiopia. Not only the traffic accidents are concentrated in Oromia region and Addis Ababa city but also the volume of motorized traffic are very high as compared to the other regions of the country. Studies that focused on the analysis of road traffic accident in Oromia regional state using count regression models are limited.

Researchers studied the road traffic accident variables and employed statistical methods such as descriptive statistics, logistic regression, Poisson Regression models to predict accident frequency at roadway(Tewolde,2007; Anteneh, 2014; Haile and Demeke,2014;Fikadu,2015;Dawit,2016). Most of these studies focused on Addis Ababa city survey of road traffic accidents. In the study of number of human deaths, count regression models are more appropriate than other methods. Many researchers categorized the count variable as binary and did their analysis. However logistic regression undercounts the total number of deaths due to road traffic accident since multiple human deaths are collapsed into a single unit to fulfill the requirements of logistic regression.

Therefore, this study attempts to analyze number of human death due to road traffic accident in Oromia region based on per accident data using appropriate count regression model.

1.3 Objectives of the study

1.3.1 General objective

The general objective of this study is to identify major factors determining the number of human death per road traffic accident in Oromia Regional state of Ethiopia.

1.3.2 Specific objectives

The specific objectives of this study are:

1. To identify the groups that are more at risk than others among the identified factors

2. To forward valuable recommendation that might contributes the number of human deaths per road traffic accident and analysis of such data.

1.4 Significance of the study

This study focuses in identifying the predictors of the number of human death per road traffic accident in Oromia region, Ethiopia. The findings of this study

- could be helpful for policy makers, program implementers in designing intervention strategies targeted towards decreasing exposure to road traffic accidents as well as casualties
- could play an important role for further research
- might support researchers in determining and using an appropriate count regression model.

1.5 Scope of the study

The study will be limited to road traffic accident in Oromia regional state of Ethiopia.

1.6 Limitation of the study

This study is based on secondary data obtained from Oromia Police Commission Bureau. The data are records of human death due to road traffic accidents. The recorded data were available in hard and softcopies. However, the reliability of the data related to drivers age, vehicle service years and driving experience was questionable as these were collected by interviewing the drivers themselves. Among the challenges during data collection are:

- Lack of some very useful data; for example, the records in the police stations said nothing about the utilization of seatbelts and helmets, the drivers' alcohol use, as well as whether a driver was using mobile phone while driving.
- Some information like driver's age, driving experience, driver's sex were missing due to escaping of the drivers who caused the accident.
- Some valid information recorded about seriously injured peoples that might have died after some time was missing.

1.7 Outline of the study

This thesis work is presented in five chapters. Chapter one high lights major issues relating to road traffic accidents at a global level in general, and Ethiopia, in particular. The significance and objectives of the study are also described. Chapter two contains related literature reviews on factors associated with human death due to road traffic accident. Chapter three describes the methodological issues of the study, and Chapter four gives the results, and Chapter five is discussion. Finally, conclusions and recommendations of the study are presented in Chapter six.

CHAPTER TWO: REVIEW OF RELATED LITERATURE

A road traffic accident can be defined as a fatal or non-fatal injury incurred as a result of a collision on a public road involving at least one moving vehicle (WHO, 2015). According to Safe Car Guide (2004), road traffic accident can be defined as an accident that occurs on a way or street open to public traffic, results in one or more persons being killed or injured, and at least one moving vehicle is involved. Therefore, road traffic accident is a collision between vehicles, between vehicles and pedestrians, between vehicles and animals, or between vehicles and fixed obstacles.

More than 1.2 million people die each year on the world's roads, making road traffic injuries a leading cause of death globally. Most of these deaths are in low- and middle-income countries where rapid economic growth has been accompanied by increased motorization and road traffic injuries. As well as being a public health problem, road traffic injuries are a development issue: low- and middle income countries lose approximately 5% of GDP as a result of road traffic crashes. Globally an estimated 3% of GDP is lost to road traffic deaths and injuries. Ninety-percent of road traffic deaths occur in low- and middle-income countries, and while these countries also account for 82% of the world's population, they nevertheless bear a disproportionate number of deaths relative to their level of motorization, as they account for only 54% of the world's registered vehicles and for comparisons between the continents Africa has the highest contribution to road traffic accident which is 26.6 % fatality deaths per 100,000 populations (WHO, 2015).

According to the study by Tewolde (2007), the number of injuries per accident is mainly determined by variables related to drivers such as drivers' age, educational background and place of accident. He used Poisson regression model.

According to Haile and Demeke (2014), road traffic accidents have been increasing at an alarming rate, causing the loss of valuable workforce and resources. The fundamental findings of their study showed that seat belt use, vehicles type; causes of accident were the leading factors that affect the occurrences of road traffic accidents. Moreover cargo vehicles were predominantly involved in road traffic accidents, while in Bahir Dar city, the majority of the accidents were attributed to motorized three wheeler taxies.

A study by K. Mamady *et al*.,(2014) demonstrated that most of the deaths were among occupants, motorcyclists, pedestrians, and the productive workforce aged 25 - 49 years.

Regardless of which factors are more significant, the results of a study by Rahma Ahmed Al-Jabri, (2015) provided valuable information on how collision types, road specifications, time, weather, and driver characteristics are related to the occurrence of injuries in accidents.
Fikadu (2015) studied main road traffic accident problems in Addis Ababa by developing accident prediction models that link accident frequencies to their behavioral contributing factors, including narrow and heavily pedestrians movements, mismanagement of urban public transport provision, poor standard of vehicle (most of vehicles are old), poor maintenance and development of roads, the negligence of drivers due to lack of adequate training and driver testing, negligence of pedestrians due to lack of stringent laws, poor traffic control and enforcement which combine increased risk of accidents on roads of Addis Ababa. Also summarized were the main road traffic accident problems in Addis Ababa, including Poor skill and undisciplined behavior of drivers and pedestrians, over speeding, pedestrians not taking proper precautions, drivers not respecting pedestrian priority, over loaded or improperly loaded vehicles, poor inspection system of vehicles, Weak traffic law enforcement, Poor vehicle conditions, Poor infrastructure condition, lack of modern technology used to control the traffic flow.

Abdul-Rahaman (2014) used binary logistic regression to accident-related data collected from motor transport and traffic unit Northern Region traffic-police records. A total of 398 accident data from 2007-2009 was used. Among the variables obtained from police-accident reports, two independent variables were found most significantly associated with accident severity: Over loading and Obstruction.

Getu *et.al*. (2013) described the road traffic accident crashes in Ethiopia using descriptive analysis. The results of the study revealed that, the variation in road traffic crashes by time of day reflects variations in traffic volumes, and most crashes occur during daylight. Besides driver age group as a factor, driver involvement in crashes was disproportionately high for the 18-30 age group, followed by the 31-50 age group. This trend also applies for all road users killed and injured. Most crashes occurred on paved two-way two-lane roads (in cities and on interstate highways), particularly in central business districts; and residential areas. In terms of collision types, pedestrian crashes are the dominant types of collision, as motorized traffic and pedestrians share the same facilities.

By using road related data and land use data along with data on traffic accidents occurred in Tokyo in 2013, the numbers of traffic accidents in workdays and non-workdays were modeled separately and the influence of those factors upon the occurrence of traffic accidents was investigated by Paweenuch, *et.al* ,(2015). Anteneh(2015), based on road traffic accident data in Addis Ababa city, Ethiopia, found out that the majority of those affected were vulnerable road users among which pedestrians were predominant and affected while crossing the road outside the zebra cross and responsible parties were driving commercial cars and vast majority of victims died at the scene instantaneously.

Getahun (2015) captured predictor variables that had significant effects on the number of human death per road traffic accident. He used ZIP model, and found out that month in quarter, age of the driver, vehicle type, ownership of vehicle, accident time, accident type and road inclination were statistically significant factors associated with the number of human death per road traffic accident in Addis Ababa.

Muhammad Imran and Jamal Abdul Nasir (2015) in their study indicated that one cause of road traffic accidents in Pakistan is lack of traffic rules and regulation among the general public.
Noland and Quddus (2005) conducted a disaggregated spatial analysis based on enumeration district area to examine the effect of congestion on traffic casualties. In this study, congestion was spatially controlled by proxy variables. Negative binomial models were used for analysis of casualties during congested and uncongested periods. The study showed that traffic casualties are likely to happen on higher speed roads and motorways but not during traffic congestion.

Ahmed Abdella (2013) used data regarding the number of fatalities per traffic accident obtained from Addis Ababa traffic control and investigation department for a one year period from April 2012-March 2013and analyzed the data using count regression Models. Drivers in the age group

18-30 years, the drivers having less than 5 years of experience and those who are employed were highly associated with the number of fatalities per accident. Among vehicle related factors, automobile, taxi-minibuses and vehicles less than five years of service were highly associated with the number of fatalities.

According to Augustus (2012), most of the road traffic accidents in Nigeria occur in the months of June, July, September, October, November and December. June, July and September are the peak rainy months and more road accidents occur in rainy months because wet road conditions affect many drivers' ability.

Getu (2007) studied main causes of accidents and identified unavailability of proper pedestrian facilities, high volume of pedestrian traffic, drivers' fatigue, lack of awareness of traffic rules and regulations, and violation of speed limit as significant factors. Layla A. (2017) used logistic regression model in determining the effective variables leading to traffic accident. The results of his study revealed that high speed, car type and location were most significantly associated to accident.

Aguero-Valverede and Jovanis (2006) demonstrated Bayesian and negative binomial models to carry out spatial analysis of fatal and injury crashes in Pennsylvania. The study used counties as the spatial unit. The study concluded that counties with a higher percentage of the population under poverty level, higher percentage of their population in age groups 0-14, 15-24 and over 64, and increased road mileage and road density have significantly increased crash risk. The study also suggested that it was important to consider spatial correlation in road-segment and intersection-level accident.

Bedard *et.at.* (2002) used multivariate logistic regression model to determine the independent contribution of crash, driver and vehicular characteristics that lead to increasing driver's fatality risk. According to this study, over speed, not using seatbelts and severity incidences attributed to driver side impacts were associated with driver fatalities.

A study conducted to show the effect of new pavement on traffic safety in Sweden investigated whether higher pavement road standards lead to a higher or lower accident rates. The Poisson regression model fitted reasonably well for the expected number of accidents and it does not depend on the different periods considered for analysis and also on the urban road accidents (Geedipally, 2005).

Jun, et.*al*. (2015) used an approach to modeling and predicting crash frequency at rural intersections by crash type and injury severity level using a multinomial Probit kernel for the discrete choice model and introduced unobserved heterogeneity in both the crash frequency model and the discrete choice model. The results revealed that the type of traffic control and the number of entering roads are the most important determinants of crash counts and crash type/injury severity.

A study by Fekede, *et.al* (2014), demonstrated a strong positive association between fatal road traffic collisions and location of collision, time of collision (day or night), high speeds, failure to give priority and vehicular technical problems. Younger people, males and hired drivers were largely involved in road traffic crashes.

A study using binary and ordinal logistic regression models to identify factors influencing traffic fatalities and injuries in Addis Ababa, Ethiopia revealed that drivers aged 18-30 years caused the largest number of accidents. Low level educational background of drivers, absence and poor lighting along roads, wet surface and asphalt surface, morning and evening hours, places like offices, residential and commercial neighborhoods, automobiles and small taxis/minibuses were found to be associated with fatalities and serious injuries (Bisrat ,2010).

A study by Gentiana *et.al*, (2008) demonstrated a strong positive association between fatal road traffic accidents and younger age, high speed, and alcohol consumption of the responsible party in a low-income country, such as Albania, in social and economic transition using the logistic regression model.

Pan and Prakash (2013), conducted a study motorway safety by developing accident prediction models that link accident frequencies to their non-behavioral contributing factors, including traffic conditions, geometric and operational characteristics of road, and weather conditions. A number of accident prediction models were developed and assessed for their predictive ability using negative binomial regression models under three categories: first for the whole of the motorway, second for rural and urban motorway segments separately and third for motorway segments without ramp, with on-ramp and with off-ramp separately. The findings of their study uncovered the safety impacts of different non-behavioral contributing factors, in which segment length, average annual daily traffic per lane and the number of lanes always have the most profound effects on accident frequency.

Olemo (2016) used the Poisson and Negative binomial model to Exploring the major causes of road traffic accidents in Nairobi county. The study concluded that the causes of RTAs in Nairobi County are multi-factorial and can be categorized into driver factors, vehicle factors and roadway factors. Driver factors relates to all proximate factors affiliated to the diver that may result to an accident occurring. Speeding was a leading cause of accidents.

Guyu (2013) conducted a study to identify the road traffic accident black spots in Addis Ababa using descriptive statistics based on data obtained from the Addis Ababa traffic police office. The results of his study revealed that Kirkos, Bole, and Arada sub-cities were the leading in terms of the total number of road accidents. Moreover, the highest accident concentrations were located in Kirkos sub-city, and the lowest was in Gullele sub-city.

The finding of a study by Belachow Melese and Zeleke Dutamo (2015), showed that the main cause of car accident is drivers. About 80% of car accidents resulted due to the drivers fault. Among the accident that resulted from the drivers fault, not giving priority for pedestrian is the main cause while poor road condition, poor car condition, the absence of knowledge in traffic system, lack of ambulance and poor medical treatment were other causes and conditions that increase the severity of the incident. In general, the occurrence of car accident problem depends on defect on human factor, vehicle characteristics, road characteristics and environmental conditions. The age of drivers had also a significant impact on the occurrence of traffic accidents. Accidents were also highly dependent on environmental factor like weather condition, type of road surface, condition of the road and joint road shape.

Adebayo Lawrence (2015) studied that assessment of human factors as determinants of road traffic accidents among commercial vehicle drivers in Gbony in Local Government Area of Ekiti State, Nigeria. His study assesses the variables of driving under the influence of alcohol, excessive speeding, impressionistic-driving, indiscriminate parking and sleepiness in relation to road traffic accidents. Results from the study indicated that about 60% of the drivers were below 45 years of age while 50% of them were married, over 70% of them had no more than secondary education-out of which 16% had no formal education while 11% of them had primary school certificate.

A study by Mustakim and Fujita (2011) demonstrated an accident predictive model for rural roadway based on data collected at rural roadway, Malaysia. They carried out black spot study to develop accident predictive models. Multiple non-linear regression model was used to relate the discrete accident data with the road and traffic flow predictor variables. Their findings showed that the existing number of major access points, without traffic light, rise in speed, increasing number of annual average daily traffic, growing number of motorcycle and motorcar and reducing the time gap are the potential contributors of increment accident rates on multiple rural roadway.

Esmaili , *et al*. (2013) studied the trends of fatal road traffic injuries in Iran (2004–2011). Their study assessed several causes and mechanisms that led to crashes. Alcohol consumption may not be a wide problem in Iran due to religious beliefs and governmental norms, but falling asleep when driving may also be one of the causes, other causes can also be considered such as over speeding, distraction by using mobile phones and talking to or trying to impress passengers. They used Autoregressive integrated moving average (ARIMA) model.

In general, the literature seems to agree that the following factors are important determinants of road traffic accidents : Human characteristics (sex of driver, age of driver, driving experience, education level of driver, driver-vehicle relationship, ownership of vehicle) , Vehicle characteristics (vehicle type, vehicle age of service,), Environmental characteristics (accident time , road condition, environment of accident, weather condition, location of accident, day of accident),road characteristics (road pavement, road inclination), other related characteristics such as accident cause accident type (collision between vehicles, collision between vehicle and pedestrian, collision between vehicle and animal, vehicle down).

CHAPTER THREE: DATA AND METHOD

3. 1. Source of the Data

The source of the data for this study was the Oromia region police commission Bureau. This study is based on a secondary data obtained from Oromia region traffic control and crime investigation department office. Data on road traffic accidents that have occurred from July 8, 2016 to July 7, 2017 were obtained. The total number of road traffic accidents happened in this period was about 3,900,out of which 1,188 accidents involved the death of 1,541 people while 2,712 accidents did not involve fatality.

3.2 Variables of the Study

The response variable

The response variable of this study is the number of human death per road traffic accident in Oromia region. This includes any people dead due to road traffic accident such as pedestrians, road users, passengers and others.

Explanatory Variables

Potential explanatory variables considered in this study are:

- ✓ Human Characteristics: Sex of driver, age of driver, driving experience, education level of driver, driver-vehicle relationship.
- ✓ Vehicle characteristics: vehicle type, vehicle length of service.
- ✓ Environmental characteristics: accident time, road condition, environment of accident, weather condition, location of accident, day of accident.
- ✓ Road characteristics: road pavement, road inclination.
- ✓ Other related characteristics: Accident cause, accident type (collision between vehicles, collision between vehicle and pedestrian, collision between vehicle and animal, vehicle upside down, etc.).

Description and Coding of the Study Variables

Description of factors included as potential predictors of the number of human death per road traffic accident are presented in Table 3.1

Table 3.1: Working definitions of explanatory variables

No	Predictor variables	representative of variables	Coding Categories
	Human Characteristics variables		
1	Gender of the Driver	X_1	0= Male
			1= Female
2	Age of Driver	X_2	0= 18-30
			1= 31-50
			2 = 51 and above
3	Driving Experience	X_3	0= Less than 5 years
			1= 5-10 years
			2 = Above 10 years
4	Vehicle driver ownership	X_4	0= Employee
			1= Owner
			2=Other(relative's, friend)
5	Educational level of the driver	X_5	0= Elementary school and below
			1 = High School
			2 = above high School
	Vehicle characteristics variables		
6	Vehicle Type	X_6	0=Taxi-minibus(up to 12 Seats)
			1=Automobile
			2=Pick up (upto 10 quintals)
			3=Cargo(11-40 quintals)
			4= Cargo (41-100 quintals)
			5=Cargo with trailer (up to 400 quintals)
			6=Bus(13-45 seats)
			7=Bus(above 46 seats)

			8= Other (motorcycle, Bajaj, Bicycle…)
7	Vehicle length of service from production	X_7	0= Up to 5 years
			1= 6-10 years
			2= more than 10 years
	Environmental Characteristics variables		
8	Day of a week	X_8	0 = Monday
			1=Tuesday
			2=Wednesday
			3=Thursday
			4=Friday
			5=Saturday
			6=Sunday
9	Road Condition	X_9	0= Wet
			1= Dry
			2=Muddy
10	Weather Condition	X_{10}	0=Not normal(cloudy, rainy, cold, hot)
			1= Normal
11	Accident Time	X_{11}	0= Afternoon
			1= Morning
			2=Evening
			3=Night
12	Location of the Accident	X_{12}	0=central zones (special &Shoa zones)
			1=western zones(Wallaga zones)
			2= south west zones (Jima, I/A/Bor, B/Bedele)

			3= Easter Zones (Bale, Hararghe)
			4= southern zones (Arsi, Guji, Borana)
13	Environment of the Accident	X_{13}	0= Factory place
			1=Residence place
			2= Commercial place
			3=Worship places
			4=Entertainment place
			5= School
			6= Hospital
			7=In village rural area
			8=Out of village in rural area
			9=Government office place
	Road Characteristics Variables		
14	Road inclination	X_{14}	0= not straight(sloped, curved, scarp, uphill)
			1= straight
15	Road Pavement	X_{15}	0= Good Asphalt
			1=Defected Asphalt
			2 =Gravel
			3= site Clearing
	Other Characteristics Variables		
16	Accident Type	X_{16}	0=Collision between Vehicles
			1= Collision between vehicle and pedestrian
			2= Collision between vehicle and animal
			3= vehicle upside down

			4= Other (vehicle down, vehicle crash to inert)
17	Accident Cause	X_{17}	0= not given priority to pedestrian
			1= overload
			2= Overspeed
			3= Turning illegal position
			4= Steering problem
			5= Brake problem
			6= Release of Tyre
			7=following not keeping distance

3.3 Methods of Data Analysis

In this section both exploratory (descriptive) and inferential statistical data analyses methods that have been used are discussed. STATA (version 13) was used for data analyses.

3.3.1. Exploratory Data Analysis

We started with an exploratory data analysis to gain insight into the dataset. Data exploration is a very helpful tool in the selection of appropriate models and understanding of the data nature. Descriptive statistics were used to observe a possible link between explanatory variables and human death per road traffic accident.

3.3.2. Statistical Data Analysis

This study used count regression models. Count regression models are appropriate when the dependent variable is a non-negative integer valued count, y = 0, 1, 2,…, where y is measured in natural units on a fixed scale. These values represent the number of times an event occurs in a fixed domain. For count data, the standard framework for explaining the relationship between the outcome variable and a set of explanatory variables includes the Poisson and Negative Binomial regression models. Unlike linear regression, count data regression models have counts as the response variable that can take only nonnegative integer values.

Among the models that deal with count dependent variables are:

- Poisson Regression
- Negative Binomial Regression
- Zero-inflated Poisson
- Zero-inflated Negative Binomial
- Hurdle models (Poisson-hurdle and Negative-binomial hurdle).

The classical Poisson regression model for count data is often of limited use in these disciplines because empirical count data sets typically exhibit over-dispersion and/or an excess number of zeros. The former issue can be addressed by extending the plain Poisson regression model in various directions: e.g., using sandwich co variances or estimating an additional dispersion parameter(in a so-called quasi-Poisson model). Another more formal way is to use a negative binomial(NB) regression. All of these models belong to the family of generalized linear models (GLMs, see Nelder and Wedderburn 1972; McCullagh and Nelder 1989). However, although these models typically can capture over-dispersion rather well, they are in many applications not sufficient for modeling excess zeros. Since Mullahy (1986) and Lambert (1992), there is increased interest, both in the econometrics and statistics literature, in zero-augmented models that address this issue by a second model component capturing zero counts. Hurdle models (Mullahy, 1986) combine a left-truncated count component with a right-censored hurdle component. Zero-inflation models (Lambert, 1992) take a somewhat different approach: they are mixture models that combine a count component and a point mass at zero. A comprehensive and up-to-date account of count models and methods as well as interpretations of fitted count models are provided by Cameron and Trivedi (2013).

3.3.2.1 Poisson Regression Model

A Poisson regression model allows modeling the relationship between a Poisson distributed response variable and one or more explanatory variables. It is suitable for modeling the number of events that occur in a given time period or area.

The explanatory variables can be either numeric or categorical. We can also include interaction terms in the model. Poisson Regression Model provides a standard framework for the analysis of count data. Let Y_i represent counts of events occurring in a given time or exposure periods with rate μ_i. Y_i are Poisson random variables with probability mass function (pmf) for Y given by

$P(Y = y_i \mid \mu_i) = \frac{e^{-\mu_i}\mu_i^{y_i}}{y_i!}$, $\mu_i \geq 0,$ (3.1)

where y_i = 0, 1, 2 … are discrete counts (in our case y is number of human death per accident) and μ is the mean of the Poisson distribution. The mean and variance of Poisson distribution are given as $E(y_i) = Var(y_i) = \mu_i.$ (3.2)

One specification that is mostly used for the mean parameter μ_i is the exponential specification. This specification ensures that μ_i is non-negative and it is given as $E(y_i) = \mu_i = \exp(x_i'\beta)$ where $x_i' = (1, x_{i1}, \ldots, x_{ip})$, is a vector of explanatory variables and β is a $(p + 1)$ dimensional column vector of unknown parameters to be estimated.

Estimation involves estimating the regression parameters specifically using the maximum likelihood estimation. The likelihood function of the Poisson model based on a sample of n independent observations is given by

$$L(\beta) = \prod_{i=1}^{n} \frac{e^{-\mu_i}\mu_i^{y_i}}{y_i!} \quad (3.3)$$

Taking logs we find the log-likelihood

$$\ell = \log L(\beta) = \sum_{i=0}^{n}[y_i \ln\mu_i - \mu_i - \ln y_i!] \quad (3.4)$$

By taking the partial derivations of the log-likelihood function and setting them equal to zero, we obtain the likelihood equations for estimating the parameters. Thus, the first derivative of ℓ with respect to the parameters is:

$$\frac{d\ell(\beta)}{d\beta} = \sum_{i=1}^{n}(y_i - \mu_i)x_i \quad (3.5)$$

If $E(y_i) < var(y_i)$, we have over-dispersion, and when $E(y_i) > var(y_i)$, we say we have under-dispersion.

There are two basic criteria commonly used to check the presence of over-dispersion: the deviance, $D(y, \mu_i)$ or the Pearson $\chi 2$ statistic. Deviance, $D(y, \mu_i)$ or the Pearson $\chi 2$ greater than its degrees of freedom indicate over-dispersion. For the Poisson regression, $D(y, \mu_i)$ and $\chi 2$ are given by

$$D(y, \hat{\mu}_i) = 2\,\Sigma\{y_i ln\frac{y_i}{\hat{\mu}_i} - (y_i - \hat{\mu}_i)\} \quad (3.6)$$

$$\chi 2 = \sum_{i=1}^{n} \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i} \quad (3.7)$$

Where, y is the number of events, n is the number of observations and $\hat{\mu}_i$ is the fitted Poisson mean.

The null hypothesis to be tested is that the mean and variance of the response variable are equal against the alternative hypothesis that the variance exceeds the mean.

When over dispersion occurs, the Poisson maximum likelihood estimator obtained will be incorrect (Cameron and Trivedi, 2013), for which negative binomial regression can be used to correct.

3.3.2.2 Negative Binomial Regression Model

Negative Binomial Regression model is an extension of Poisson regression developed to allow for possible over-dispersion in the data. Negative binomial regression adds an ancillary parameter that allows over-dispersion (but not under dispersion). It is a standard method used to model over-dispersed Poisson data. Given that over-dispersion is the norm, the negative binomial model has more generality than the Poisson model. Over-dispersion is most often caused by highly skewed response/dependent variables (often due to variables with high numbers of zeros).The probability mass function of negative binomial distribution is given by

$$f(y_i ; \mu_i ; \delta) = \frac{\Gamma(y_i + \frac{1}{\delta})}{y_i! \Gamma(\frac{1}{\delta})} (1 + \delta\mu_i)^{-1/\delta} (1 + \frac{1}{\delta\mu_i})^{-y_i} \quad , y_i \geq 0; \delta > 0 \qquad (3.8)$$

with mean and variance given by:

$E(y_i) = \mu_i = \exp(x_i'\beta)$, and $var(y_i) = \mu_i (1 + \delta\mu_i)$

where the constant δ is the dispersion factor . We relate the parameters μ_i to the covariates x_i through the log-link function so that $\log\mu_i = x_i'\beta$

where $x_i' = (1, x_{i1}, x_{i2}, \dots, x_{ip})$, $1 \times p$ row vector of covariates and $\beta = (\beta_0, \beta_1, \beta_2, \dots, _p)'$ is the corresponding $(p + 1) \times 1$ column vector of unknown regression parameters. The maximum likelihood estimation method is used to estimate the parameter vector.

The likelihood function of the NB model based on a sample of n independent observations is given by

$$L(y_i ; \mu_i ; \delta) = \prod_{i=1}^{n} \{ \frac{\Gamma\left(y_i + \frac{1}{\delta}\right)}{y_i! \Gamma\left(\frac{1}{\delta}\right)} (1 + \delta\mu_i)^{-\frac{1}{\delta}} \left(1 + \frac{1}{\delta\mu_i}\right)^{-y_i} \}$$

The log-likelihood function is

$\ell = \log L(\mu_i , \delta; y_i)$

$$\ell = \sum_{i=1}^{n} \{ -\log(y_i!) + \sum_{k=1}^{y_i} \log(\delta y_i - \delta k + 1) - (y_i + \frac{1}{\delta}) \log(1 + \delta\mu_i) + y_i \log(\mu_i) \} \qquad (3.9)$$

The likelihood equations for estimating μ_i and δ are obtained by using the partial derivatives of the log-likelihood function and setting them equal to zero. Thus, we obtain the first derivative of $\ell = \log L(\mu, \delta; y_i)$ with respect to the underlying parameters as follows:

$$\frac{\partial \ell}{\partial \beta} = \frac{\partial \ell}{\partial \mu}\frac{\partial \mu}{\partial \beta} = \sum_{i=1}^{n}\left\{\frac{(y_i - \mu_i)}{(1+\delta\mu_i)}\right\} x_i \, , \tag{3.10}$$

$$\frac{\partial \ell}{\partial \delta} = \sum_{i=1}^{n}\left\{\left(\sum_{k=1}^{y_i}\frac{y_i - k}{(\delta y_i - k\delta + 1)}\right) + \frac{(\log(1+\delta\mu_i)}{\delta^2} - \frac{(y_i + \frac{1}{\delta})\mu_i}{(1+\delta\mu_i)}\right\} \tag{3.11}$$

3.3.2.3 Zero-inflated count regression models

These models were introduced by Lambert (1992). Real-life count data are frequently characterized by over dispersion and excess zeros. Zero inflated count models provide a parsimonious yet powerful way to model this type of situation. Such models assume that the data are a mixture of two separate data generation processes: one generates only zeros, and the other is either a Poisson or negative binomial data-generating process. The main motivation for a zero inflated count models provide a way of modeling the excess zeros in addition to allowing for over dispersion. Major source of over-dispersion is a relatively large number of zero counts and the resulting over-dispersion cannot be modeled accurately with negative binomial model. In such cases, one may use zero-inflated Poisson or zero-inflated negative binomial models to fit the data.

3.3.2.4 Zero- inflated Poisson Regression Model

The excess zeros are a form of over dispersion and fitting a zero inflated Poisson model can account for the excess zeros, but there are also other sources of overdispersion that must be considered. If there are sources of overdispersion that cannot be attributed to the excess zeros, failure to account for them constitutes a model misspecification, which results in biased standard errors. In ZIP models, the underlying Poisson distribution for the first subpopulation is assumed to have a variance that is equal to the distribution's mean. If this is an invalid assumption, the data exhibit over dispersion (or under dispersion).

The probability distribution of a zero inflated Poisson random variable is given by

$$P(Y_i = y_i) = \begin{cases} \omega_i + (1-\omega_i)e^{-\mu}, & y_i = 0 \\ (1-\omega_i)\dfrac{e^{-\mu_i}\mu_i^{y_i}}{y_i!} \quad y_i = 1, 2, \ldots & 0 \le \omega_i \le 1 \end{cases} \tag{3.12}$$

The mean and variance of Zero-inflated Poisson (ZIP) distribution is given as

$E(Y_i) = (1-\omega_i)\mu_i$, and $var(Y_i) = E(Y_i)(1+\omega_i\mu_i)$.

The parameters μ_i and ω_i depend on covariates X_i and Z_i respectively, where $\log(\mu_i) = X_i^T\beta$ and $\log\left(\frac{\omega i}{1-\omega i}\right) = Z_i^T\gamma$

The log-likelihood function, $\ell = \log L(\mu_i, \omega_i; y_i)$ for ZIP model is given below:

$$\ell = \sum_{i=1}^{n}\left\{I_{(Y_i=0)}\log[\omega_i + (1-\omega_i)\exp(-\mu_i)] + I_{(Y_i>0)}[\log(1-\omega_i) - \mu_i + y_i\log(\mu_i) - \log(y_i!)]\right\}$$

with respect to the covariate X_i and Z_i the log-likelihood function can be written as follows:

$= \log(Y_i/X_i, Z_i)$

$$\ell = \sum_{i=1}^{n}\{I_{(y_i=0)}\log[\exp(Z_i^T\gamma) + exp(-exp(x_i^T\beta))] + I_{(y_i>0)}[[y_i x_i^T\beta - exp(x_i^T\beta) - log(1 + exp(z_i^T\gamma))]]\}$$

A useful diagnostic tool that can aid in detecting over dispersion is the Pearson chi square statistic. Pearson chi-square statistic is defined as

$$X^2 = \sum_{i=0}^{n}\frac{(y_i-\mu_i)^2}{V(\mu_i)}$$

Comparing the computed Pearson chi-square statistic to an appropriate quintile of a chi squared distribution with n-p df constitutes a test of over-dispersion. If overdispersion is detected, the ZINB model often provides an adequate alternative.

3. 3.2. 5 Zero-Inflated Negative Binomial Regression Model

Zero-inflated negative binomial regression is for modeling count variables with excessive zeros and it is usually for overdispersed count outcome variables. Furthermore, theory suggests that the excess zeros are generated by a separate process from the count values and that the excess zeros can be modeled independently. The probability distribution of a zero inflated negative binomial random variable Y (the response variable) is given by

$$P(Y_i = y_i) = \begin{cases} \omega_i + (1-\omega_i)(1+\delta\mu_i)^{-1/\delta} & , y_i = 0 \\ (1-\omega_i)\frac{\Gamma(y_i+\frac{1}{\delta})}{y_i!\Gamma(\frac{1}{\delta})}(1+\delta\mu_i)^{-1/\delta}\left(1+\frac{1}{\delta\mu_i}\right)^{-y_i} & , y_i > 0 \end{cases} \quad (3.13)$$

where $\delta > 0$ is a dispersion parameter and is assumed not to depend on covariates. The mean and variance of the ZINB model are given by

$E(Y_i) = (1-\omega_i)\mu_i$ and $Var(Y_i) = (1-\omega_i)(1 + \omega_i\mu_i + \delta\mu_i)\mu_i$

The parameters μ_i and ω_i depend on vectors of covariates x_i and z_i, respectively. We can write the model as $\log(\mu_i) = X_i^T\beta$ and $\log\left(\frac{\omega_i}{1-\omega_i}\right) = Z_i^T\gamma$

The log-likelihood function, $\ell = logL(\delta, \mu_i, \omega_i, ; y_i)$ for the ZINB model is given by

$$\ell = \sum_{i=1}^{n} \left\{ I_{(y_{i=0})} \log\left(\omega_i + (1-\omega_i)(1+\delta\mu_i)^{-1/\delta}\right) \right.$$

$$\left. + I_{(y_i>0)} \log((1-\omega_i) \frac{\Gamma\left(y_i + \frac{1}{\delta}\right)}{y_i!\,\Gamma\left(\frac{1}{\delta}\right)} (1 + \delta\mu_i)^{-\frac{1}{\delta}} \left(1 + \frac{1}{\delta\mu_i}\right)^{-y_i}) \right\}$$

The likelihood equations for estimating parameters are obtained by using the partial derivatives of the log-likelihood function and setting them equal to zero.

The method of Fisher scoring is more appropriate for ZINB regression because the second derivative $\ell = \ell(\delta,\ \mu_i\ ,\ \omega_i\ ,\ ;\ y_i\ )$, is simplified by taking expectations to obtain the parameter estimates of ZINB regression models since the zero-inflated negative binomial (ZINB) distribution is not a standard generalized linear model (GLM) type.

3.3.2.6 Hurdle Models

A hurdle model is a modified count model in which there are two processes, one generating the zeros and one generating the positive values. The two models are not constrained to be the same. The concept underlying the hurdle model is that a binomial probability model governs the binary outcome of whether a count variable has a zero or a positive value. If the value is positive, the "hurdle is crossed," and the conditional distribution of the positive values is governed by a zero-truncated count model. Originally developed by Mullahy (1986), Hurdle count models are two-component models with a truncated count component for positive counts and a hurdle component that models the zero counts. The count model is typically a truncated Poisson or negative binomial regression (with log link). The pm f of y is

$$p(y=y_i/x_i\ ,z_i,\ \beta,\gamma) = \begin{cases} f_{zero}(0; z_i;\ \gamma) & if\quad y_i = 0 \\ (1 - f_{zero}(0; z_i;\ \gamma))\ (\frac{f_{count}}{(1-f_{count}(0;x_i;B)}(y_i;x_i;\beta\ )\ ) & \text{if} y_i > 0 \end{cases} \qquad (3.14)$$

where Y_i is the value of the dependent variable for the i^{th} person i = 1, …, n), z_i is a vector of length J denoting the number of predictor variables in the zero part, χ_i represents a vector of length K denoting the number of predictor variables in the hurdle part, γ is a vector of coefficients belonging to z, and β denotes a vector of coefficients related to x. f *zero* is a probability density function at least on {0, 1} (binary) or {0, 1, 2, …} (count), and f *count* is a probability density function on {0, 1, 2, …}. Regression coefficients are estimated with maximum likelihood.

The f_{zero} part (where y_i = 0) is typically modeled with a binary logit (logistic regression) model, where all counts greater than 0 are given a value of 1. (Min & Agresti, 2005)

Using a binary logistic regression model for this part, the probability of y_i= 0 is denoted as

$f_{zero}(0; z_i;\ \gamma) = \psi_i = \frac{1}{1+e^{z_i\gamma}}$

where z_i represents the observed data and γ the vector of coefficients belonging to z_i. Obviously, the probability of a nonzero count is given by $1 - \psi_i$. The non-zero count part(*f count*) is modeled with a truncated(y_i>0) count model. This is typically a truncated Poisson model or a negative binomial model in case of overdispersion.

The parameters μ_i and ω_i depend on vectors of covariates x_i and z_i, respectively. We can write the model as

$\log(\mu_i) = X_i^T\beta$ and $\log\left(\frac{\omega_i}{1-\omega_i}\right) = Z_i^T\gamma$

3.3.3 Goodness of fit tests

3.3.3.1 Likelihood Ratio test(LRT)

The likelihood-ratio test is given by: LRT= -2(L_0 -L)where $L_0$is the log-likelihood of models under the null hypothesis and L is the log-likelihood of models under the alternative hypothesis. LR test compares the value of the statistic with a chi-square distribution with p-1 degree of freedom where p is the number of parameter and if the value of the test statistics exceeds the critical value, the null hypothesis is rejected and the overall model is significant. The null hypothesis may be stated as “the overdispersion parameter is equal to zero (i.e. the Poisson model is preferred)” while the alternative hypothesis can be stated as “the over-dispersion parameter is different from zero(i.e. the data would be better fitted by the negative binomial regression)”.

3.3.3.2 Vuong Test

The Vuong test is a non-nested test that is based on a comparison of the predicted probabilities of two models that do not nest (Vuong, 1989). That means vuong test statistics are needed to provide the appropriateness of zero-inflated models against the standard count models. Thus, comparisons between Zero-inflated count models with ordinary Poisson, or Zero inflated negative binomial against ordinary negative binomial model can be done using Vuong test. Define $m = \log\left(\frac{P_1\left(\frac{y_i}{x_i}\right)}{P_2\left(\frac{y_i}{x_i}\right)}\right)$

where $P_1(y_i/x_i)$ is the probability mass functions of ZI and $P_2(y_i/x_i)$ refers to the Poisson or NB models.

The Vuong test statistic, V is simply the average log likelihood ratio suitably normalized

$$V = \frac{\sqrt{n}\,(\frac{1}{n}\sum_{i=1}^{n} m_i)}{\sqrt{(\frac{1}{n}\sum_{i=1}^{n}(m_i - \overline{m})^2)}} = \sqrt{n}\,\frac{\overline{m}}{S_m}$$

Where, $\overline{m}$ is mean, S_m is standard deviation and n sample size. This test statistic is follows the standard normal distribution.

If V> the predetermined critical value ($Z_{\alpha/2}$), the first model is preferred (that means zero inflated Poisson or ZINB is preferred).

If V< the predetermined critical value ($-Z_{\alpha/2}$), the second model is preferred (that means Poisson or NB is preferred).

If $|V| < Z_{\alpha/2}$ the two models are the same, none of the models are preferred.

3.3.3.3 Information Criteria (AIC and BIC)

Information Criteria are the way to evaluate the model that attempt to quantify how well our model would have predicted the data. The model selection criteria that are most often used for GLMs are AIC and BIC

$AIC = -2\ell + 2p$

$BIC = -2\ell + p\log n$

where ℓ is the log-likelihood of a model that will compare with other models, *n* is the sample size of the data and p is the number of parameters in the model including the intercept. The model with minimum value of AIC or of BIC is preferable.

3.3.4 Statistical tests of individual parameters

3.3.4.1Wald Test

A Wald test is used to test the statistical significance of each parameter (the individual coefficients) in the model. That means Wald test is used to determine the significance of individual explanatory variables.

To test: $H_0: \beta_j = 0$. Versus $H_1: \beta_j \neq 0$ where $_j$ is parameter and j= 1, 2,….,p

The test statistic, $Z = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)}$ where $se(\hat{\beta}_j)$ is the standard error of $\hat{\beta}_j$ and has an approximate standard normal distribution, when H_0 is true. Equivalently, the Wald statistic, Z^2 has a chi-squared distribution with d.f. = 1, for large sample size.

CHAPTER FOUR: RESULTS

4.1Descriptive Statistics

This chapter provides the results of the analyses using the methods outlined in Chapter three. As shown in Table 4.1, the variance of the dependent variable is greater than its mean, suggesting a possibility of overdispersion. A further screening on the data also showed that more than 69% of the accidents, 2,712 out of 3,900, had no human deaths per accident, indicating excess zeroes in the dataset**.**

Table 4.1: Summary statistics for the number of human death per accident.

Mean	Variance	Skewness	Kurtosis
0.3951	0.7586	4.211	37.398

Table 4.2 shows the frequency distribution of the number of deaths per accident. It indicates that more than 69% of the accidents were nonfatal suggesting excess zeros and 30.5% of the accidents were fatal accidents.

Table 4.2: Frequency distribution of Number of human deaths per accident.

Count (No. of deaths)	Frequency	Percent
0	2712	69.5
1	961	24.6
2	169	4.3
3	29	.7
4	14	.4
5	8	.2
>5	7	.18
Total	3900	100.0

The histogram in Figure 1 is highly skewed to the right showing massive counts of zero outcomes. However, large number of human deaths per accident was less frequently observed. Dual regime event count models such as zero-inflated models and hurdle models will often tend to indicate over-dispersion in the data as a result of the large number of zero counts.

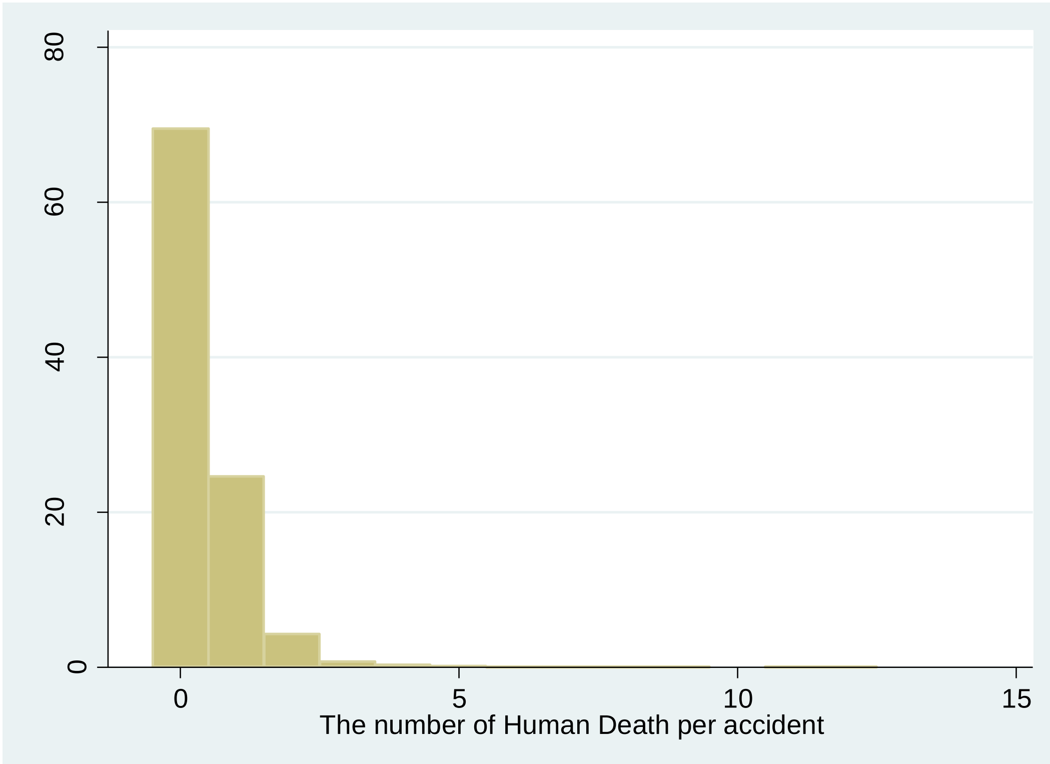

Figure 1: Distribution of response variable, count of human death per accident

Table 4.3: Descriptive statistics of fatal accidents for predictor variables of human death due to road traffic accident.

No	predictor variables	Categories	N	mean	variance
1	Gender of the Driver	Male	3605	.3953	.5066
		Female	295	.3932	1.4231
2	Age of Driver	18-30	1986	.4874	.9109
		31-50	1121	.3568	.2297
		51 and above	793	.2181	.1707
3	Driving Experience	Less than 5 years	2696	.4362	.7325
		5-10 years	698	.3510	.2281
		Above 10 years	504	.2361	.1807
4	Vehicle driver	Employee	3355	.4399	.6364

	relationship	Owner	379	.1160	.1028
		Other(relative's, friend's)	166	.3072	.2141
5	Vehicle Type	Taxi-minibuses(up to 12 Seats)	1457	.1983	.1591
		Automobile	90	.5555	.2497
		Pick up (up to 10 quintals)	375	.2960	.2089
		Cargo(11-40 quintals)	416	.2764	.2005
		Cargo (41-100 quintals)	284	.3838	.2373
		Cargo with trailer (up to 400 quintals)	503	.3837	.2369
		Buses (13-45 seats)	290	1.2517	1.1094
		Buses(above 46 seats)	114	.9824	3.9288
		Other(motorcycle, Bajaj, Bicycle,…)	371	.5364	1.4872
6	Vehicle service	Up to 5 years	2296	.3649	.2318
		6-10 years	927	.3204	.4058
		Above 10 years	677	.5997	1.928
7	Road Condition	wet	1817	.4259	.6863
		Dry	414	.4082	.2421
		Muddy	1669	.3583	.5358
8	Weather Condition	Not normal(cloudy, rainy, cold, hot)	3432	.4146	.6252
		Normal	468	.2521	.1889
9	Accident Time	Afternoon	1309	.4553	.9928
		Morning	705	.3049	.2122
		Evening	1303	.4029	.3728
		Night	583	.3516	.5136
10	Location of the Accident	central zones (special &shoa zones)	2371	.4496	.7960
		western zones(wallaga's zones)	407	.2113	.1670
		south west zones (Jima, I/A/Bor, B/Bedele)	311	.3633	.2320

		Easter Zones (Bale, Hararghe Zones)	359	.4011	.2409
		south east zones (Arsi, Guji,Borana Zones)	452	.2920	.2338
11	Environment of the Accident	Factory place	303	.1320	.1149
		Residence place	447	.8009	2.6037
		Commercial place	491	.4562	.2486
		Religious place	340	.3471	.2273
		Entertainment place	355	.3662	.2327
		School place	361	.2770	.2008
		Hospital place	325	.3323	.2225
		In village rural area	454	.2378	.1817
		Out of village in rural area	442	.2262	.1754
		Government office place	382	.6675	.8052
12	Accident Type	Collision between Vehicles	2004	.4361	.9011
		Collision between vehicle and pedestrian	341	.5366	.2494
		Collision between vehicle and animal	201	.4826	.2509
		vehicle upside down	769	.2145	.1687
		Other (vehicle down, vehicle crash to inert)	585	.3795	.2358
13	Educational level of the driver	Elementary school and below	2778	.4449	.6885
		High School	679	.4123	.3695
		Above high school	443	.0564	.0533
14	Day of the weeks	Monday	659	.3050	.2123
		Tuesday	557	.3052	.2124
		Wednesday	529	.3043	.2121

		Thursday	574	.3048	.2123
		Friday	495	.3050	.2124
		Saturday	640	.6547	1.7788
		Sunday	446	.5919	0.9612
15	Road pavement	Good Asphalt	1405	.3388	.2242
		Defected Asphalt	1220	.3877	.7691
		Gravel	1152	.4392	.7174
		site Clearing	123	.6992	1.2120
16	Road inclination	not straight(sloped, curved, scarp, uphill)	3416	.4124	.6265
		straight	484	.2727	.1987
17	accident cause	not given priority to pedestrian	525	.7085	.2068
		Over load	222	.0990	.0896
		Over speed	246	.0203	.0199
		Turning illegal position	480	.2229	.1736
		Steering problem	580	.8034	2.2272
		Brake problem	546	.3095	.2141
		Release of tyre	614	.3306	.4076
		following not keeping distance between	687	.2867	.2048

Table 4.3 shows that the mean number of human deaths per accident for male and female drivers were 0.3953 and 0.3932, respectively, indicating that the mean number of human death per accident for male drivers are almost the same as that for female drivers.

Age of driver: The average number of human death per accident is inversely related with increasing age. It was highest (.4874) among driver in the youngest age group of 18-30 years.

Driving experience: Considering the experience of drivers, the highest mean number of human death occurred among drivers with less than 5 years driving experience (0.4362), while the smallest mean number of human death (0.2361) was reported for drivers with more than 10 years driving experience.

Driver-Vehicle relationship: Employed drivers had high contribution to the mean number of human death (0.4399); on other hand, less contribution was made by owner drivers (0.1161).

Vehicle Type: Considering vehicle type categories, the highest mean number of human death of 1.2517 was observed in accidents by buses (13-45 seats) whilst the smallest mean number of human death per accident were recorded in accidents by minibuses and automobiles (0.1983).

Vehicle length of service: With regard to the categories of vehicle service age (in years), the highest mean number of human death was observed among older vehicles having above 10 years (0.5997) service.

Road condition: The highest mean number of human death per accident occurred in wet roads. On the other hand the smallest mean number of human death per accident occurred on muddy roads.

Weather condition: Table 4.3 shows that normal weather condition exhibited lower average number of human death per accident (.2521) than unusual weather conditions (cloud, rain, cold, hot) 0.4146.

Accident time: Considering the accident time variable, the highest mean number of human death per road traffic accident occurred in the afternoons and evenings time, whilst the smallest mean number of human death was reported in mornings.

Location of the Accident: The highest mean number of human death per accident occurred in central zones of Oromia region such as Oromia special zone surrounding Finfinne (Addis Ababa), Shoa zones (0.4496), whereas the lowest mean number of human death per accident occurred in western zones of the region such as Wallaga zones (0.2113).

Environment of the Accident: With regard to area of the accident, the highest mean number of human death was observed around residential areas; and the second highest mean number of human deaths was observed around government offices. The lowest mean number of human deaths per accident was observed around factories.

Accident Type: The accident type has five categories. Among these categories, accident collision between vehicle and pedestrian had the highest mean number human death (0.5366), and the second highest mean number of human death per accident occurred due to collision between vehicles and animals (.4826), and also the third highest mean number of human death occurred due to collision between vehicles (.4361). Accidents that turned vehicles upside down had the smallest mean number of human death per accident (.2145).

Educational level of driver: The highest mean number of human death (0.4449) was reported among drivers with at most elementary education; while the smallest mean number of human death was (0.0564) observed among drivers having more than high school educational level.

Day of the weeks: Among the day of the week included in the study, Saturday had the highest mean number of human death per accident as compared to other days.

Road pavement: The highest mean number of human death per accident happened on site clearing roads and gravel roads as compared to the other road pavements such as asphalt road.

Road inclination: Higher mean number of human deaths was observed in roads that were not straight (slopy, curved, sharp, uphill) than straight road inclination.

Accident cause: The highest mean number of human death occurred due to steering problem (0.8034) followed by drivers not giving priority to pedestrian (0.7085). Overspeed and overload caused the lowest mean number of human death. Turning illegal position, brake problem, release of tyre and following not keeping the minimum required distance between had intermediate mean human death per accident.

4.2. Test of overdispersion and Goodness-of-fit test

As shown by the overdispersion test results in Table 4.4, the ratios of the Deviance and Pearson Chi-square statistics to their corresponding degrees of freedom are greater than one, indicating over-dispersion in the data and the Negative Binomial (NB) regression model is preferred over the Poisson model.

Table 4.4.: Results of over-dispersion test after fitting a Poisson regression

Statistics	Value	Degrees of freedom	P-value
Deviance test statistics	4060.979	3899	0.0346
Pearson Chi-square statistic	5679.44	3899	0.0000

The likelihood ratio tests of the null hypothesis H_0; δ=0 that the errors do not exhibit over-dispersion in the dataset is rejected since the likelihood ratio test statistic value is equal to 140.30 with p_value = 0.000 and thus we conclude that there is overdispersion in the dataset and NB regression model may be the appropriate model.

4.2.2: Comparison of Models

A critical question in data analysis is how to choose the most suitable model for a specific study. Several criteria can be used to compare and select the most appropriate model among considered models. In this study, six different count regression models, namely; Poisson, negative binomial, zero-inflated Poisson and zero-inflated negative binomial, hurdle Poisson and hurdle negative binomial models were considered. Different model selection criteria: the Vuong test, Log pseudo likelihood, Akaike information criterion (AIC) and Bayesian information criterion (BIC) were used in order to identify the most appropriate fitted model. In cases of over-dispersion, the ZIP and hurdle Poisson model typically fits better than a standard Poisson model. But there are other models that allow for overdispersion: the standard negative binomial regression model, ZINB and hurdle NB model.

Table 4.5: Model selection results based on criteria for Poisson, NB, ZIP, ZINB, Hurdle Poisson and NB hurdle models.

Selection criteria	**Models**					
	Poisson	**NB**	**ZIP**	**ZINB**	**HP**	**HNB**
Logpseudo likelihood	-1616.283	-1571.679	-1307.637	-1395.554	**-914.64**	-1219.76
AIC	3400.567	3331.359	2759.279	2989.108	**2037.28**	2649.52
BIC	3927.14	3920.62	3210.623	3609.713	**2627.41**	3144.73
Vuong test (p-value)			27.74(0.000)	28.45(0.000)		

Table 4.5 shows the model selection criteria used to identify the most preferred model among the candidate models. First, the calculated value of the Vuong test statistic (27.74) for comparing ZIP versus Poisson model is greater than 1.96 implying that the ZIP model is preferred to the Poisson model for predicting the human death per accident. Similarly, the calculated value of the Vuong test statistic for comparing ZINB versus NB models is 28.45, indicating that the ZINB model is preferred to NB regression model. The Vuong test results indicate over-dispersion due to many zero observations and unobserved heterogeneity suggesting the use of ZIP, ZINB, HP or HNB models.

Finally to compare the ZIP, ZINB, Hurdle Poisson and Hurdle NB models, AIC, BIC and log pseudo likelihood were used as shown in Table 4.5. The model with the smallest AIC, smallest

BIC and largest log pseudo likelihood is preferred. Since Hurdle Poisson model has the smallest AIC, smallest BIC and maximum log pseudo likelihood, hurdle Poisson model is the most appropriate and preferred model among the six models. Thus, the hurdle Poisson regression model with the lowest value of AIC, lowest value of BIC and the highest value of Log pseudo likelihood is found to be the most appropriate and preferred model for describing the number of human death per accident due to road traffic accident.

Table 4.6: The estimated Count Model coefficients (truncated Poisson with log link) for the number of human death per accident of selected independent variables

Predictors	Coef.	IRR	Robust std. Error	Z-value	P>z	[95% CI)	
Gender of driver							
male(ref.)							
female	0.1071632	1.113116	0.0706291	1.5200	0.1290	-0.03127	0.245594
Age of driver							
18-30(ref.)							
31-50	-1.240293	.2892994	0.2578944	-4.8100	0.0000	-1.74576	-0.73483
51 and above	-1.168266	.3109057	0.4496858	-2.6000	0.0090	-2.04963	-0.2869
experience of driver							
0-5 years(ref.)							
5-10 years	-4.306517	.0134804	0.3474679	-12.3900	0.0000	-4.98754	-3.62549
above 10	-2.297937	.1004659	0.2861389	-8.0300	0.0000	-2.85876	-1.73712
Vehicle-driver ownership							
Employee(ref.)							
Owner	-0.195747	.822220	0.5317931	0.3700	0.7130	-1.238042	-0.84655
others	0.1549816	1.167636	0.3338171	0.4600	0.6420	-0.49929	0.809251
Education level of driver							
Elementary school and below(ref.)							
high school	0.0292152	1.029646	0.0278621	1.0500	0.2940	-0.02539	0.083824
above high school	0.0027787	1.002783	0.285112	0.0100	0.9920	-0.55603	0.561588
Type of vehicle							
Taxi-minibuses(up to 12 seats) (ref.)							

Automobile	2.156696	8.642535	0.5816313	3.7100	0.0000	3.296672	1.016719
pick up(up to 10 quintals)	-0.9710704	.3786775	0.759982	-1.2800	0.2010	-2.46061	0.518467
cargo(11-40 quintals)	-0.7913882	.4532152	0.8017604	-0.9900	0.3240	-2.36281	0.780033
cargo (41-100 quintals)	0.0578815	1.059589	0.8577503	0.0700	0.9460	-1.62328	1.739041
cargo with trailer (up to 400 quintals)	1.215571	3.37222	0.9678426	1.2600	0.2090	-0.68137	3.112508
buses (13-45 seats)	1.529158	4.614288	1.018837	1.5000	0.1330	-0.46773	3.526041
buses (above 46 seats)	1.871637	6.498926	1.030076	1.8200	0.0690	-0.14727	3.890549
others (motorcycle, bicycle, bajaj,....)	1.465154	4.328209	1.021895	1.4300	0.1520	-0.53772	3.468
Vehicle length of service							
0-5 years(ref.)							
5-10 years	0.9099758	2.484262	0.3738715	2.4300	0.0150	0.177201	1.642751
above 10 years	0.9707648	2.639963	0.374237	2.5900	0.0090	0.237274	1.704256
Road condition							
wet(ref.)							
dry	1.555803	4.738892	0.3470648	4.4800	0.0000	0.875569	2.236038
muddy	-0.0652523	.936831	0.0620276	-1.0500	0.2930	-0.18682	0.05632
Time of accident							
Afternoon							
Morning	-0.3435897	.7092199	0.7016853	-0.4900	0.6240	-1.71887	1.031688
Evening	-0.7503553	.4721987	0.1245018	-6.0300	0.0000	-0.99437	-0.50634
Night	-0.7333746	.4802855	0.1325627	-5.5300	0.0000	-0.99319	-0.47356
Location of accident							
Central zones(special and shoa zones)							
Western zones (Walaga's zones)	-1.290377	.2751671	0.2560859	-5.0400	0.0000	-1.7923	-0.78846
South west zones (Jima,ilu aba bor,BunoBedele)	-0.009135	.9909066	0.1221261	-0.0700	0.9400	-0.2485	0.230227
Eastern zones (Hararge's and Bale zone)	-0.210303	.8103387	0.1469683	-1.4300	0.1520	-0.49836	0.07775
South east zones (Arsi's, Guji's and Borana zone)	-0.4714403	.6241027	0.2142694	-2.2000	0.0280	-0.8914	-0.051
Environment of accident							

factory place(ref.)							
Residence place	4.686809	108.5064	1.054896	4.4400	0.0000	2.61925	6.754368
Commercial place	-0.7448221	.4748188	0.7489019	-0.9900	0.3200	-2.21264	0.722999
Religious place	-0.971702	.3784384	0.6035554	-1.6100	0.1070	-2.15465	0.211245
Entertainment place	-0.8226925	.4392474	0.9558907	-0.8600	0.3890	-2.6962	1.050819
School place	-0.3468148	.7069363	0.315963	-1.1000	0.2720	-0.96609	0.272461
Hospital place	3.169237	23.78933	1.050076	3.0200	0.0030	1.111126	5.227348
Rural village area	-1.884967	.1518341	0.8544336	-2.2100	0.0270	-3.55963	-0.21031
Out of village in rural area	-1.31526	.2684045	0.732435	-1.8000	0.0730	-2.75081	0.120286
Government office area	4.864501	129.6062	1.053231	4.6200	0.0000	2.800206	6.928795
Road Pavement							
Good asphalt(ref.)							
Defected asphalt	-0.0540752	.9473608	0.3902166	-0.1400	0.8900	-0.81889	0.710735
Gravel	-0.6978099	.4976741	0.4159444	-1.6800	0.0930	-1.51305	0.117426
Site clearing	-0.5682137	.5665365	0.4108004	-1.3800	0.1670	-1.37337	0.23694
Type of accident							
Collision between vehicles(ref.)							
Collision between vehicle and pedestrian	0.3421681	1.407997	0.251152	1.3600	0.1730	-0.15008	0.834417
Collision between vehicle and animal	0.2604785	1.297551	0.2569847	1.0100	0.3110	-0.2432	0.764159
Vehicle upside down	1.715531	5.559628	0.4066338	4.2200	0.0000	0.918544	2.512519
others (vehicle down,vehicle crash to inert)	0.3737013	1.453103	0.1858116	2.0100	0.0440	0.009517	0.737885
accident cause							
not given priority to pedestrian(ref.)							
Over load	-0.2896361	.7485359	0.3745402	-0.7700	0.4390	-1.02372	0.444449
Over speed	-0.9142836	.4008036	0.8440052	-1.0800	0.2790	-2.5685	0.739936
Turning illegal position	-0.7885209	.4545166	0.3561384	-2.2100	0.0270	-1.48654	-0.0905
Steering problem	-0.3668899	.692886	0.5212428	-0.7000	0.4820	-1.38851	0.654727
Brake problem	-0.3198858	.726232	0.6122253	-0.5200	0.6010	-1.51983	0.880054

Release of tyre	-0.3446338	.7084797	0.5199014	-0.6600	0.5070	-1.36362	0.674354
following not keeping distance between	-0.3571482	.6996688	0.2665624	-1.3400	0.1800	-0.8796	0.165305
_cons	-5.023439	.0065819	0.4870688	-10.3100	0.0000	-5.97808	-4.0688

Notes: Robust Standard Error is in Robust Std. Err.

4.2.3. Interpretation of truncated Poisson model fit results

The estimated count regression model fit results are presented in Table 4.6. The estimated coefficients can be interpreted as follows: for a one unit change in the predictor variable, the log of the response variable is expected to change by the value of the regression coefficient (coef.). In Poisson hurdle model, for every one unit increase (in units of the significant predictors, the log number of human death per accident is expected to increase or decrease by approximately the corresponding coefficient in the column of coefficient (coef.). In this model the variables whose p-value<0.05, were considered statistically significant. Interpretations are made in terms of the incidence rate ratios ($IRR=exp^{(coef.)}$) which is important to explain the change in percentage (|IRR-1|)100% of significant predictors.

In order to study the covariates related with the number of human death per accident and to predict the count of human death per accident, we fitted the Poisson hurdle regression model. The statistically significant predictors related to human death among those in zero truncated Poisson part of the model are age of driver, driver experience, type of vehicle, vehicle service, road condition, time of accident, location of accident, environment of accident, type of accident and accident cause.

Age of driver: The expected number of human death per accident due to road traffic accident for drivers in the age group 31-50 years had decreased by 71% and for drivers aged 51 years and above. It decreased by 68.9% as compared to the expected number of human death per accident for drivers in the age group 18-30 holding all other variables in the model constant.

Experience of driver: The estimated coefficients for driver's experience are negative and statistically significant. The results in Table 4.6 show that driver's experience has a significant impact on the number of human death per road traffic accident in the zero truncated group. The expected number of human death due to road traffic accident by drivers with 5-10 years driving experience was 98% lower, accidents by drivers with more than 10years experience was 89.9%

lower than the expected number of human death due to road traffic accidents by drivers with less than 5 years driving experience.

Type of vehicle: The coefficient for automobile is positive and statistically significant. The expected number of human death per road traffic accident by drivers who drive their own automobile increased by a factor of 8.64 as compared to the expected number of human death per road traffic accident for taxi-minibuses drivers holding all other variables in the model constant.

Vehicle service: The finding of this study also revealed that vehicle service (age) had a significant effect on the number of human death per road traffic accident. The expected number of human death due to road traffic accidents involving vehicles having 5-10 years of service increased by a factor of 2.48 as compared to those having 0-5 years of service (reference group) controlling for other variables in the model. In addition, the expected number of human deaths by vehicles with over 10 years of service increased by a factor of 2.65 as compared to those with 0-5 years of service.

Road condition: The expected number of human death per road traffic accident increased by a factor of 4.738 on dry roads as compared to wet roads.

Time of accident: The coefficients for evening and night time categories are negative and statistically significant. The IRRs imply that road traffic accidents in the evening are expected to have 53 percent less number of human death than road traffic accidents in the afternoon time. Likewise, accidents in the night time are expected to have 52 percent less number of human death than accidents in the afternoon time.

Location of accident: The expected number of human death per road traffic accident in the western part of the region (Wallaga's zones)was 72.3% less than that for the central zone (Oromia special zone surrounding Finfinne and shoa zone) holding all other variables in the model constant. The expected number of human death per road traffic accident in the south eastern part of the region (Arsi's,Guji's and Borana zones)was 37.58% lower than that for the central zone (Oromia special zone surrounding Finfinne and shoa's zone) holding all other variables in the model constant.

Environment of accident: The expected number of human death per accident around residence areas was 108.5 times the expected number of human death per accident around factory areas. The expected number of human death per accident around hospitals was 23.78 times the expected number of human death per accident around factory areas. The expected number of human deaths per road traffic accident around rural villages decreased by 84.8% as compared to the expected number of human deaths per road traffic accident of around factory area. Also the expected number of human deaths per accident around government offices was 129.54 times the expected number of human death per road traffic accident around factory areas.

Type of accident: The expected number of human death per road traffic accident for accidents with vehicles upside down was 5.56 times the expected number of death due to road traffic accidents as a result of collision between vehicles while the expected number of death per traffic road accident when vehicle is upside down or vehicle crashes to inert is increased by 45.3% as compared to expected number of deaths due to collision between vehicles.

Accident cause: The number of human death per road traffic accident due to illegal turning is 54.5% lower as compared to the number of human death by accidents due to drivers not giving priority to pedestrians.

Table 4.7. Parameter estimates of Zero hurdle model (binomial with logit link)

Predictors	Coef.	IRR	Robust std Error	z	P>z	95% C.I.	
Gender of driver							
male(ref.)							
Female	-2.023747	.1321593	2.583897	-0.7800	0.4340	-7.08809	3.040598
Age of driver							
18-30(ref.)							
31-50	9.363823	11658.88	12.19134	0.7700	0.4420	-14.5308	33.25842
51 and above	2.990593	19.89748	0.9363441	3.1900	0.0010	1.155393	4.825794
experience of diver							
0-5 years(ref.)							
5-10 years	-7.545835	.0005283	15.61546	-0.4800	0.6290	-38.1516	23.05991
above 10	-3.358065	.0348025	10.61037	-0.3200	0.7520	-24.154	17.43789
Vehicle-driver relationship							
Employee(ref.)							
Owner	-2.342244	.0961118	0.6382553	-3.6700	0.0000	-3.5932	-1.09129

Others	4.004295	54.83314	4.544281	0.8800	0.3780	-4.90233	12.91092
Education level of driver							
Elementary school and below(ref.)							
high school	-0.019889	.9803079	0.7527574	-0.0300	0.9790	-1.49527	1.455489
above high school	-2.347016	.0956542	6.344977	-0.3700	0.7110	-14.7829	10.08891
Type of vehicle							
Taxi-minibuses(up to 12 seats) (ref.)							
Automobile	4.881567	131.837	10.42474	0.4700	0.6400	-15.5506	25.31369
pick up(up to 10 quintals)	6.617289	747.9149	10.18349	0.6500	0.5160	-13.342	26.57656
cargo(11-40 quintals)	16.3656	$1.28e^{+7}$	25.18401	0.6500	0.5160	-32.9942	65.72536
cargo (41-100 quintals)	26.15795	$2.29e^{+11}$	22.22668	1.1800	0.2390	-17.4056	69.72145
cargo with trailer (up to 400 quintals)	38.49083	$5.20e^{+16}$	26.1681	1.4700	0.1410	-12.7977	89.77937
buses (13-45 seats)	46.94439	$2.44e^{+20}$	71.86159	0.6500	0.5140	-93.9017	187.7905
buses (above 46 seats)	47.02697	$2.65e^{+20}$	71.91633	0.6500	0.5130	-93.9265	187.9804
others (motorcycle, bicycle, bajaj,....)	45.42922	$5.37e^{+19}$	72.05245	0.6300	0.5280	-95.791	186.6494
Vehicle service							
0-5 years(ref.)							
5-10 years	4.200073	66.69119	43.65489	0.1000	0.9230	-81.362	89.76209
above 10 years	-0.022379	.9778696	0.7155403	-0.0300	0.9750	-1.42481	1.380054
Road condition							
wet(ref.)							
Dry	4.515944	91.46389	0.3491412	12.9300	0.0000	3.83164	5.200248
Muddy	-1.842185	.1584707	1.196012	-1.5400	0.1230	-4.18633	0.501955
Time of accident							
Afternoon (ref.)							
Morning	-0.306898	.7357253	1.84104	-0.1700	0.8680	-3.91527	3.301473
Evening	-2.00779	.1342851	0.7710442	-2.6000	0.0090	-3.51901	-0.49657
Night	-4.731411	.008814	1.960663	-2.4100	0.0160	-8.57424	-0.88858

Location of accident							
Central zones(special and shoa zones)							
Western zones (Walaga's zones)	-4.357243	.0128137	1.459128	-2.9900	0.0030	-7.21708	-1.4974
South west zones (Jima,ilu aba bor,BunoBedele)	-3.127891	.0438101	0.7527256	-4.1600	0.0000	-4.60321	-1.65258
Eastern zones (Hararge's and Bale zone)	-2.625282	.0724193	0.6393023	-4.1100	0.0000	-3.87829	-1.37227
South east zones (Arsi's,Guji's and Borana zone)	-2.60517	.0738906	0.8620075	-3.0200	0.0030	-4.29467	-0.91567
Environment of accident							
factory place(ref.)							
Residence place	4.07487	58.84284	27.02836	0.1500	0.8800	-48.8998	57.04949
Commercial place	5.164555	174.9596	0.9822274	5.2600	0.0000	3.239425	7.089685
Religious place	3.138904	23.07856	1.43108	2.1900	0.0280	0.334039	5.943769
Entertainment place	4.145541	63.15179	23.60026	0.1800	0.8610	-42.1101	50.4012
School place	-7.511227	.0005469	15.87554	-0.4700	0.6360	-38.6267	23.60425
Hospital place	3.87201	48.03886	28.34404	0.1400	0.8910	-51.6813	59.4253
Rural village area	-20.47262	$1.28e^{-09}$	14.00196	-1.4600	0.1440	-47.916	6.970708
Out of village in rural area	-15.77798	$1.41e^{-07}$	14.02418	-1.1300	0.2610	-43.2649	11.7089
Government office area	4.162857	64.25484	27.56451	0.1500	0.8800	-49.8626	58.1883
Road Pavement							
Good asphalt(ref.)							
Defected asphalt	-11.41474	.000011	4.172302	-2.7400	0.0060	-19.5923	-3.23718
Gravel	-12.66744	3.15e-06	14.78715	-0.8600	0.3920	-41.6497	16.31484
Site clearing	-12.00381	6.12e-06	15.19819	-0.7900	0.4300	-41.7917	17.78409
Type of accident							
Collision between vehicles(ref.)							
Collision between vehicle and pedestrian	1.927412	6.871705	10.03442	0.1900	0.8480	-17.7397	21.5945
Collision between vehicle and animal	-0.206437	.8134779	1.027506	-0.2000	0.8410	-2.22031	1.807437
Vehicle upside down	2.691796	14.75815	10.22351	0.2600	0.7920	-17.3459	22.72951
others (vehicle down, vehicle crash to inert)	-0.724617	.4845102	0.7809273	-0.9300	0.3530	-2.25521	0.805973

accident cause							
not giving priority to pedestrian(ref.)							
Over load	-18.67411	7.76e-09	35.86136	-0.5200	0.6030	-88.9611	51.61288
Over speed	-29.71508	1.24e-13	3.738553	-7.9500	0.0000	-37.0425	-22.3877
Turning illegal position	-30.25686	7.24e-14	10.80742	-2.8000	0.0050	-51.439	-9.0747
Steering problem	-47.72962	1.87e-21	73.93878	-0.6500	0.5190	-192.647	97.18773
Brake problem	-39.63408	6.13e-18	26.5886	-1.4900	0.1360	-91.7468	12.47863
Release of tyre	-48.7763	6.56e-22	74.36127	-0.6600	0.5120	-194.522	96.96911
following not keeping distance between	-15.84481	1.31e-07	13.33265	-1.1900	0.2350	-41.9763	10.2867
_cons	17.35198	3.43e+07	12.69121	1.3700	0.1720	-7.52233	42.22629

Notes: Robust Standard Error is in Robust Std. Err.

4.2.4. Interpretation of zero hurdle coefficients (binomial with logit link) model fit results

Table 4.7 provides estimated coefficients for the factor change in the odds of being in the zero counts group (no human death per road traffic accident faced) compared to the non-zero counts group (at least one human death per road traffic accident). In the zero counts group, there is no human death per road traffic accident but there are human injuries and property damages. Age of driver, vehicle-driver ownership, road condition, time of accident, location of accident, environment of accident, road pavement and accident cause had significant impacts on the probability of being in the zero counts group. The odds of being in the zero counts group increased by a factor of 19.89 for drivers in the above 50 age group as compared to those aged18-30 controlling other variables in the model. With regard to vehicle-driver ownership, the odds of being in the zero counts group decreased by about 90.4% for owner drivers as compared to employee drivers holding all other variables in the model constant. The odds of being in the zero counts group increased by a factor of 91.459 for road traffic accidents on dry roads as compared to accidents on wet roads holding all other variables in the model constant.

Time of accident had significant impact on the odds of being in the zero counts group. For instance, as compared to the afternoon, the odds of being in the zero counts group decreased by 86.56%, and 99% for evening and night time accidents, respectively. Similarly, the location of accidents had significant impact on the odds of being in the zero counts, so as compared to the central zones, the odds of being in the zero counts group decreased by 98.7%, 95.6% ,92.75 and

92.57% for western zones, south western zones, eastern zones and south eastern zones respectively. The odds of being in the zero counts group increased by a factor of 174.96 and 23.078 for road traffic accidents in commercial areas and worship areas respectively, as compared to road traffic accidents around factory areas. The odds of being in the zero counts group decreased by 99 % for road traffic accidents at road pavements on defected asphalt roads as compared to those on road pavements on good asphalt roads holding all other variables in the model constant. The odds of being in the zero counts group decreased by a factor of 99.9% for road traffic accident due to overspeeding as compared to those due to denying priority to pedestrian holding all other variables in the model constant.

In summary, the zero hurdle model tells us what affects the likelihood of clearing the zero "hurdle". Owner drivers , evening and night time, among the accident location western, eastern ,south east and south west parts of the region, poor asphalt of road pavement, overspeed and turning illegal direction, are less likely to clear the hurdle as compared to their reference categories. Driver of age above 50 years age of drivers, dry road condition, commercial and religious area of accident, are more likely to clear the hurdle as compared to their reference categories.

CHAPTER FIVE: DISCUSSION

This study was carried out to identify the major factors associated with the number of human death per road traffic accident based on Oromia region road traffic data. The total number of accidents in the one year period from July 8, 2016 to July 7, 2017 included in the present study was 3,900 among which 30.5% had at least one death per accident due to different factors. The hurdle Poisson regression model was selected as the most appropriate model from six possible candidate count models.

Drivers aged (18-30) had higher contribution to the number of human death per road traffic accident as compared to drivers above 30 years of age. This result was similar to the study by (Getahun, 2015; Tewolde, 2007; Haile and Demeke, 2014).

Drivers having less than 5 years of driving experience were involved in road traffic accidents with more human death as compared to those who have more than 5 years driving experience. Among vehicle types, number of deaths due to road traffic accidents by automobiles was significantly greater than that by other vehicle types. This result is consistent with the findings of Towelde (2007) and Ahmed (2013).

Consistent to the findings by Fikadu (2015), older vehicles were associated with accidents having higher number of human death per road traffic accident as compared to accidents by vehicles having less than 5 years of service. The reason for this may be due to poor maintenance of vehicles. Dry road conditions were responsible for large number of human death per road traffic accident as compared to wet roads.

Accident time is another significant factor associated with the number of human death per road traffic accident. Less number of human deaths per road traffic accident took place in evenings and night times as compared to afternoons. This variation in road traffic crashes by times of the day reflects variations in traffic volumes (Getu, *et.al.* 2013). This finding is similar with the findings of the studies by Fekede, *et.al* (2014) and Bisrat (2010).

Higher number of human death per accident occurred in residential areas, hospital areas and government office areas. Conversely, fewer number of human deaths per accident happened in rural villages than near factory areas**.** This finding has some similarity with the results of the study by Bisrat (2010) and Ahmed (2013)**.** Consistent with the findings of Getahun (2015), vehicle upside down and others (vehicle down, vehicle crash to inert(motionless)) accident types were associated with increased number of human death per road traffic accident as compared to the number of human death due to collision between Vehicles.

The number of human deaths per road traffic accident that took place due to turning in a forbidden direction was lower as compared to the number of deaths due to drivers denying priority to pedestrians. It was observed that the number of human deaths per road traffic accident due to denying priority to pedestrians was more than the number of deaths by accidents due to other causes. This result is similar with the finding by Haile and Demeke (2014) and Fekede,*et.al*, (2014).Similar to Belachew Melese and Zeleke Dutamo (2015), the cause of about 80% of car accidents was attributed to drivers faults among which denying priority for pedestrian is the leading one.

CHAPTER SIX: CONCLUSION AND RECOMMENDATIONS

6.1 Conclusion

This paper attempted to identify the major factors associated with the number of human death per road traffic accident using an appropriate count regression model. In a total of 3,900 road traffic accidents that occurred in Oromia region during the one year period from July 8, 2016 to July 7, 2017, a total of 1,541 people died. The **Poisson hurdle regression model (HP)** was found to be the most appropriate and preferred well-fitting model among the six count models considered for the data. The descriptive results suggested that there is high variability in the data set. The variance of the number of human death per road traffic accident was larger than its mean, suggesting the possibility of over-dispersion. In addition, all formal tests indicated that the data are overdispersed and that there are excess zeros. The Poisson hurdle model provided the most appropriate fit to the dataset.

This study captured predictor variables that had significant effects on the number of human death per road traffic accident. The selected Poisson hurdle model fit results indicated that age of driver, experience of driver, type of vehicle, vehicle service, road condition, time of accident, location of accident, environment of accident, type of accident and accident cause were statistically significant factors associated with casualty of human death per road traffic accident in Oromia region of Ethiopia.

6.2 Recommendations

Based on the findings of this study, the following recommendations are forwarded:

- ✓ There is a need for the introduction of education programs that will create awareness about road traffic accidents and the associated human deaths especially targeting road users, young drivers, passengers and pedestrians.
- ✓ The government, concerned institutions and other stakeholders could consider the major factors associated with human deaths due to road traffic accidents that are identified by this study while designing policy.
- ✓ All concerned bodies should pay special attention to vulnerable road users. There is a need to focus on road infrastructure investments that allow the separation of vulnerable from other road users on improved road rule enforcement.

- ✓ Further researchers may adopt our method and broaden the scope of this research by including other regions of the country in order to establish a better research base.

References

Abdul-Rahaman H. (2014). Identification of Factors that Cause Severity of Road Accidents in Ghana: a Case Study of the Northern Region, International Journal of Applied Science and Technology, 4:3.

Adebayo Lawrence (2015). Assessment of Human Factors as Determinants of Road Traffic Accidents among Commercial Vehicle Drivers in Gbony in Local Government Area of Ekiti State, Nigeria.

AgueroValverde, J. and Jovanis, P. P. (2006). Spatial analysis of fatal and injury crashes in Pennsylvania. Accident Analysis and Prevention,Vol. 38, 618-625.

Ahmed Abdella (2013). Statistical analysis of correlates of number of fatalities per traffic accident in Addis Ababa using count data models. Addis Ababa University Department of Statistics, Ethiopia.

Amanuel Kassa (2017). Trends, Causes, and Costs of Road Traffic Accidents in Ethiopia, College of Civil Service ,Ethiopia.

Anteneh K. (2015). Road Traffic Accident related Fatalities in Addis Ababa City, Ethiopia. Addis Ababa University School of medicine Department of Emergency Medicine.

Augustus O. A. (2012). A monthly Analysis of Road Traffic Accident in Selected Local Government Areas of Lagos State, Nigeria. Mediterranean Journal of Social Sciences, 3:11.

Bedard, M., Guyatt, G. H., Stones, M. J., & Hirdes, J. P. (2002). The independent contribution of driver, crash, and vehicle characteristics to driver fatalities. Accident Analysis & Prevention, 717-727.

Belachew Melese H. and Zeleke Dutamo A.(2015), Statistical Analysis of Road Traffic Car Accident in Dire Dawa Administrative City, Eastern Ethiopia. Science Journal of Applied Mathematics and Statistics. Vol. 3, pp. 250-256

Bisrat, M.(2010). Determinants of Traffic Fatalities and Injuries in Addis Ababa. Unpublished Msc thesis, Department of Statistics, Addis Ababa University.

Cameron A.C. and Trivedi P.K. (2013). Regression Analysis of Count Data. Second edition Cambridge University Press.

Central Statistics Agency (2011), Number of Traffic Accidents Recorded and Victims by Type of Accident in Ethiopia.

Central Statistics Agency(2013), Population Projection for Ethiopia 2007-2037 trends of population growth rate.

Dawit O. (2016). Road Traffic Accident and Safety Evaluation Case of Addis Ababa Bole Sub City, Ethiopia. Addis Ababa University School of Civil Engineering.

Esmaili A, et al. (2013), Trends of Fatal Road Traffic Injuries in Iran (2004–2011). The Australian National University, Australia.

Fekede A., Demeke A. and Gezahegn T. (2014). Magnitude of, trends in, and associated factors of road traffic collision in central Ethiopia. Department of Public Health, College of Health and Medical Sciences, Haramaya University, Ethiopia

Fikadu M.(2015). Road Traffic Accident Causes and Control Mechanisms In Addis Ababa City. Addis Ababa University Department of Public Administration and Development Management.

Geedipally, S.R, (2005). Analysis Of Traffic Accidents Before And After Resurfacing- A Statistical Approach. Department of Science and Technology. Linkopings Universitet, Sweden

Gentiana Q., Genc B., Bajram H.&Enver R. (2008), Factors Associated with Fatal Traffic Accidents in Tirana, Albania. Department of Public Health, Faculty of Medicine ,Tirana, Albania.

Getahun W.(2015). Count Regression Models of Human Death by Road Traffic Accidents in Addis Ababa, Ethiopia .Msc. Thesis presented Addis Ababa University.

Getu S. Tulu, Simon W. and Mark J. King(2013). Characteristics of Police-reported Road Traffic Crashes in Ethiopia over a Six Year Period. Queensland University of Technology .

Getu S.(2007). Causes of Road Traffic Accidents and possible counter measures on Addis Ababa-shashemene Roads ,Ethiopia. MSc Thesis Addis Ababa University department of Civil Engineering.

Guyu F. (2013). Identifying Major Urban Road Traffic Accident Black-Spots: A Sub-city Based Analysis of Evidences from the City of Addis Ababa, Ethiopia. Addis Ababa University, Ethiopia.

Haile M. and Demeke L. (2014). Analysis of Factors that Affect Road Traffic Accidents in Bahir Dar City, North Western Ethiopia. Science Journal of Applied Mathematics and Statistics, 2:5: 91-96.

Hedwig H., Elise D. ,Achim Z. Annemarie A. S.(2016). Modeling Caries Experience: Advantages of the Use of the Hurdle Model. Department of Statistics, University of Innsbruck, Innsbruck , Austria.

*Jun D. ,Marison C. and Chandra R. (2015)***,***a novel approach to modeling and predicting crash frequency at rural intersections by crash type and injury severity level. The University of Texas at Austin Department of Civil, Architectural and Environmental Engineering.*

Lambert, D.,(1992). "Zero-inflated Poisson regression, with an application to defects in Manufacturing", Technometrics, 1992; 34:1–14.

*Layla A. (2017)***.** *Using Logistic Regression in Determining the Effective Variables in Traffic Accidents in Iraq. Department of Mathematics, College of Education University of Garmian, Iraq.*

Mamady, K., Zou, B., Mafoule, S., Qin, J.B., Hawa, K., Lamine, K.F. and Hu, G.Q..(2014).Fatality from Road Traffic Accident in Guinea. Department of Epidemiology and Health Statistics, School of Public Health, Central South University, Changsha, China.

Mccullagh P. and Nelder J.A. (1989). Generalized Linear Models. Chapman and Hall, London.

Min and A. Agresti (2005). Random effect models for repeated measures of zero-inflated count data. Statistical Modeling 5: 1–19.

Muhammad I. and JamalA.(2015). Road traffic accidents; prediction in pakistan. 22(6):705-709.

Mullahy, J. (1986). Specification and Testing of Some Modified Count Data Models. Journal of Econometrics, Vol. 33, 341-365.

Mustakim, F., and Fujita, M. (2011). Development of accident predictive model for rural roadway. World Academy of Science, Engineering and Technology, 58: 126-131.

Nelder, J .A. and Wedderburn R.W.M., (1972), Generalized linear models, Journal of the statistical society Cameron, A.C., Trivedi, P.K., (1998), Regression Analysis of Count Data, Cambridge University Press, Cambridge, U.K.

Noland, R. B., and Quddus, M. A. (2005). Congestion and safety: a spatial analysis of London. Transportation Research Part A, Vol. 39, 737-754.

Olemo Clifford D.(2016),Exploring the major causes of road traffic accidents in Nairobi county. Department of social statistics, Nairobi University , Kenya.

Pan C. and Prakash R. (2013).Modeling Motorway Accidents Using Negative Binomial Regression. Proceedings of the Eastern Asia Society for Transportation Studies, Vol.9.

Paweenuch S., Harutoshi Y., Teerayut H. and Ryosuke S.(2015).Traffic accidents risk analysis based on road and land use factors using GLMs and zero-inflated models in Tokyo.

Persson A. (2008). *Road traffic accidents in Ethiopia: Magnitude, causes and possible interventions. Advances in Transportation Studies an International Journal, 15 (2008), ` pp. 5-18*

Rahma A. A.(2015). Regression Analysis for Estimation of the Influencing Factors on Road Accident Injuries in Oman, United Kingdom . Department of Mathematical Sciences.

Safe Car Guide (2004). International Injury & Fatality Statistics, pp 4

Tewolde M. (2007). Empirical Analysis on Traffic Accidents Involving Human Injuries. Msc. Thesis Presented Addis Ababa University, Department of Statistics.

UNECA United Nation Economic Commission for Africa(2009). Case Study: Road Safety in Ethiopia.

Vuong Q.H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica, 57:2:307-333

WHO Road Traffic Injuries Home Page: (2015) WWW.Who.Int/Violence-Injury-Prevention.

APPENDICES-1.

TABLE .A. 1.1 zero-inflated Poisson model

Predictors	Coef.	Robust Std. Err.	z	P>z	[95% Conf. interval]	
Gender of driver						
Male (ref.)						
female	0.1160	0.0661	1.7500	0.0790	-0.0136	0.2456
Age of driver						
18-30 (ref.)						
31-50	-0.2320	0.0316	-7.3500	0.0000	-0.2939	-0.1702
51 and above	-0.1547	0.0199	-7.7600	0.0000	-0.1937	-0.1156
Driving experience						
0-5 years(ref.)						
5-10 years	-0.2862	0.0337	-8.4800	0.0000	-0.3523	-0.2201
above 10	-0.1664	0.0272	-6.1200	0.0000	-0.2196	-0.1131
Vehicle-driver ownership						
Employee (ref.)						
Owner	0.0293	0.0220	1.3300	0.1840	-0.0139	0.0724
others	-0.0470	0.0237	-1.9800	0.0470	-0.0934	-0.0006
Education level of driver						
elementary school and below (ref.)						
high school	-0.0650	0.0236	-2.7500	0.0060	-0.1113	-0.0187
above hogh school	-0.0013	0.0081	-0.1500	0.8770	-0.0172	0.0147
Vehicle Type						
Tax-minibus (upto 12 seats)						
Automobile	-0.0135	0.0415	-0.3300	0.7450	-0.0948	0.0678
pick up(up to 10 quintals)	-0.1285	0.0731	-1.7600	0.0790	-0.2718	0.0149
cargo(11-40 quintals)	-0.1080	0.0745	-1.4500	0.1470	-0.2540	0.0379
cargo (41-100 quintals)	0.1740	0.0785	2.2200	0.0270	0.0202	0.3277
cargo with trailer (up to 400 quintals)	0.4627	0.0951	4.8700	0.0000	0.2764	0.6491
buses (13-45 seats)	0.5127	0.1000	5.1200	0.0000	0.3166	0.7087
buses (above 46 seats)	0.6482	0.2074	3.1300	0.0020	0.2417	1.0546
others (motorcycle,bicycle,bajaj,....)	0.4344	0.1477	2.9400	0.0030	0.1449	0.7239
Vehicle length of service						
0-5 years (ref.)						
5-10 years	0.2759	0.0662	4.1700	0.0000	0.1461	0.4057
above 10 years	0.3070	0.0647	4.7400	0.0000	0.1801	0.4338
Road condition						
Wet (ref.)						
dry	0.1223	0.0287	4.2700	0.0000	0.0661	0.1786
muddy	-0.0360	0.0364	-0.9900	0.3240	-0.1074	0.0355

Accident time						
Afternoon (ref.)						
Morning	0.0021	0.0254	0.0800	0.9340	-0.0477	0.0519
Evening	-0.3336	0.0546	-6.1100	0.0000	-0.4405	-0.2266
Night	-0.2988	0.0582	-5.1400	0.0000	-0.4127	-0.1848
Location of accident						
Central part (ref.)						
Western zones (Walaga's zones)	-0.0806	0.0185	-4.3600	0.0000	-0.1168	-0.0444
South west zones (Jima,ilu aba bor,BunoBedele)	0.0399	0.0150	2.6500	0.0080	0.0104	0.0693
Eastern zones (Hararge's and Bale zone)	-0.0169	0.0155	-1.0900	0.2760	-0.0473	0.0135
Southern zones (Arsi's,Guji's and Borana zone)	-0.0462	0.0336	-1.3700	0.1690	-0.1120	0.0197
Environment of accident						
Factory (ref.)						
Residence place	0.6095	0.1273	4.7900	0.0000	0.3601	0.8590
Commercial place	-0.0037	0.0791	-0.0500	0.9630	-0.1587	0.1513
Religious place	-0.0241	0.0757	-0.3200	0.7500	-0.1725	0.1243
Entertainment place	0.2425	0.1261	1.9200	0.0550	-0.0048	0.4897
School place	-0.0122	0.0092	-1.3300	0.1840	-0.0302	0.0058
Hospital place	0.4804	0.1277	3.7600	0.0000	0.2301	0.7307
Rural village area	-0.1192	0.0410	-2.9100	0.0040	-0.1996	-0.0388
Out of village in rural area	-0.0526	0.0304	-1.7300	0.0840	-0.1122	0.0071
Government office area	0.5873	0.1287	4.5600	0.0000	0.3351	0.8395
Road pavement						
Good asphalt (ref.)						
Defected asphalt	0.0357	0.0426	0.8400	0.4020	-0.0477	0.1191
Gravel	-0.5696	0.1201	-4.7400	0.0000	-0.8051	-0.3342
Site clearing	-0.4175	0.1066	-3.9200	0.0000	-0.6264	-0.2087
Type of accident						
Collision between vehicles (ref.)						
Collision between vehicle and pedestrian	0.0861	0.0322	2.6700	0.0070	0.0230	0.1492
Collision between vehicle and animal	0.0932	0.0328	2.8400	0.0040	0.0289	0.1575
Vehicle upside down	0.0970	0.0277	3.5000	0.0000	0.0428	0.1513
others (vehicle down,vehicle crash to inert)	0.0987	0.0316	3.1200	0.0020	0.0367	0.1607
Accident cause						
not given priority to pedestrian (ref.)						
Over load	-0.0237	0.0766	-0.3100	0.7570	-0.1738	0.1265
Over speed	-0.0657	0.0949	-0.6900	0.4890	-0.2516	0.1202
Turning illegal position	-0.2331	0.0713	-3.2700	0.0010	-0.3729	-0.0933
Streeng problem	-0.0122	0.0872	-0.1400	0.8890	-0.1831	0.1587
Brake problem	-0.4067	0.1186	-3.4300	0.0010	-0.6391	-0.1743
Release of tyre	0.0244	0.0828	0.2900	0.7680	-0.1378	0.1866

following not keeping distance between	-0.1631	0.0452	-3.6100	0.0000	-0.2516	-0.0746
_cons	0.2288	0.0296	7.7400	0.0000	0.1708	0.2867
inflate						
Gender of driver						
Male (ref.)						
female	40.0677	0.9004	44.5000	0.0000	38.3030	41.8323
Age of driver						
18-30 (ref.)						
31-50	-168.8032	1.3569	-124.400	0.0000	-171.462	-166.1437
51 and above	-52.3367	1.1178	-46.8200	0.0000	-54.5275	-50.1460
Driving experience						
0-5 years(ref.)						
5-10 years	78.8719	1.1696	67.4400	0.0000	76.5795	81.1642
above 10	-3.0348	0.3554	-8.5400	0.0000	-3.7313	-2.3383
Vehicle-driver ownership						
Employee (ref.)						
Owner	47.1705	2.8587	16.5000	0.0000	41.5676	52.7735
others	-48.8336	0.7309	-66.8100	0.0000	-50.2662	-47.4010
Education level of driver						
elementary school and below (ref.)						
high school	4.7320	0.9940	4.7600	0.0000	2.7838	6.6802
above hogh school	61.8415	1.6932	36.5200	0.0000	58.5230	65.1600
Vehicle Type						
Tax-minibus (upto 12 seats)						
Automobile	-87.3146	2.9411	-29.6900	0.0000	-93.0791	-81.5500
pick up(up to 10 quintals)	-86.3196	2.9325	-29.4400	0.0000	-92.0673	-80.5720
cargo(11-40 quintals)	-167.4544	3.0553	-54.8100	0.0000	-173.442	-161.4662
cargo (41-100 quintals)	-378.7300	2.3645	-160.180	0.0000	-383.364	-374.0958
cargo with trailer (up to 400 quintals)	-577.3197	3.4552	-167.090	0.0000	-584.091	-570.5476
buses (13-45 seats)	-748.5456	3.9647	-188.800	0.0000	-756.316	-740.7748
buses (above 46 seats)	-749.1947	3.9627	-189.060	0.0000	-756.961	-741.4279
others (motorcycle,bicycle,bajaj,....)	-747.7944	4.0340	-185.370	0.0000	-755.700	-739.8879
Vehicle length of service						
0-5 years (ref.)						
5-10 years	-139.3827	1.7463	-79.8200	0.0000	-142.805	-135.9600
above 10 years	-9.6943	1.1671	-8.3100	0.0000	-11.9818	-7.4067
Road condition						
Wet (ref.)						
dry	-29.8261	1.6074	-18.5600	0.0000	-32.9765	-26.6758
muddy	34.8051	1.1470	30.3400	0.0000	32.5570	37.0532
Accident time						

Afternoon (ref.)						
Morning	-2.0385	0.2212	-9.2200	0.0000	-2.4720	-1.6050
Evening	105.5431	1.2099	87.2300	0.0000	103.1717	107.9145
Night	106.3499	1.2018	88.4900	0.0000	103.9945	108.7053
Location of accident						
Central part (ref.)						
Western zones (Walaga's zones)	45.4728	0.3177	143.1200	0.0000	44.8501	46.0955
South west zones (Jima,ilu aba bor,BunoBedele)	44.7387	0.3017	148.3000	0.0000	44.1475	45.3300
Eastern zones (Hararge's and Bale zone)	30.9473	1.3458	23.0000	0.0000	28.3095	33.5850
Southern zones (Arsi's,Guji's and Borana zone)	36.3017	1.0476	34.6500	0.0000	34.2484	38.3550
Environment of accident						
Factory (ref.)						
Residence place	-74.8710	2.1270	-35.2000	0.0000	-79.0399	-70.7020
Commercial place	-112.1116	1.2518	-89.5600	0.0000	-114.565	-109.6582
Religious place	-1.9829	0.1847	-10.7400	0.0000	-2.3448	-1.6209
Entertainment place	-68.3535	1.7902	-38.1800	0.0000	-71.8623	-64.8447
School place	279.4931	6.7141	41.6300	0.0000	266.3338	292.6525
Hospital place	-72.5614	2.3767	-30.5300	0.0000	-77.2196	-67.9031
Rural village area	417.3573	6.5399	63.8200	0.0000	404.5392	430.1753
Out of village in rural area	419.2461	6.4627	64.8700	0.0000	406.5796	431.9127
Government office area	-76.0999	2.1157	-35.9700	0.0000	-80.2466	-71.9532
Road pavement						
Good asphalt (ref.)						
Defected asphalt	293.0863	2.8167	104.0500	0.0000	287.5657	298.6069
Gravel	232.3429	2.7911	83.2500	0.0000	226.8725	237.8132
Site clearing	230.4149	2.8120	81.9400	0.0000	224.9034	235.9264
Type of accident						
Collision between vehicles (ref.)						
Collision between vehicle and pedestrian	47.6345	1.2475	38.1900	0.0000	45.1895	50.0794
Collision between vehicle and animal	2.9598	0.3664	8.0800	0.0000	2.2418	3.6778
Vehicle upside down	-28.3289	1.3510	-20.9700	0.0000	-30.9767	-25.6810
others (vehicle down,vehicle crash to inert)	2.5879	0.3722	6.9500	0.0000	1.8583	3.3174
Accident cause						
not given priority to pedestrian (ref.)						
Over load	251.0423	2.1548	116.5000	0.0000	246.8190	255.2657
Over speed	826.6779	4.5854	180.2800	0.0000	817.6907	835.6652
Turning illegal position	385.2674	2.6047	147.9100	0.0000	380.1623	390.3725
Streeng problem	834.8062	4.3570	191.6000	0.0000	826.2666	843.3458
Brake problem	630.3731	3.3246	189.6100	0.0000	623.8569	636.8892
Release of tyre	833.1741	4.3356	192.1700	0.0000	824.6764	841.6718
following not keeping distance between	172.5501	1.2263	140.7100	0.0000	170.1466	174.9536

_cons	-396.8334	6.3950	-62.0500	0.0000	-409.3674	-384.2993

TABLE.A.1.2. zero-inflated Negative Binomial model

Predictors	**Coef.**	**Robust Std. Err.**	**z**	**P>z**	**[95% Conf. interval]**	
Gender of driver						
Male (ref.)						
female	0.1827	0.0685	2.6700	0.0080	0.0485	0.3169
Age of driver						
18-30 (ref.)						
31-50	-0.1729	0.0431	-4.0200	0.0000	-0.2573	-0.0885
51 and above	-0.0363	0.1004	-0.3600	0.7170	-0.2332	0.1605
Driving experience						
0-5 years(ref.)						
5-10 years	-0.1032	0.0487	-2.1200	0.0340	-0.1986	-0.0077
above 10	-0.0782	0.0344	-2.2700	0.0230	-0.1456	-0.0107
Vehicle-driver ownership						
Employee (ref.)						
Owner	0.0749	0.0994	0.7500	0.4510	-0.1200	0.2698
others	-0.0306	0.0528	-0.5800	0.5620	-0.1341	0.0729
Education level of driver						
elementary school and below (ref.)						
high school	-0.1940	0.0480	-4.0400	0.0000	-0.2880	-0.1000
above high school	-0.0486	0.0167	-2.9100	0.0040	-0.0814	-0.0159
Vehicle Type						
Tax-minibus (upto 12 seats)						
Automobile	-0.1536	0.0702	-2.1900	0.0290	-0.2911	-0.0161
pick up(up to 10 quintals)	-0.2195	0.1268	-1.7300	0.0830	-0.4680	0.0290
cargo(11-40 quintals)	-0.1869	0.1343	-1.3900	0.1640	-0.4502	0.0764
cargo (41-100 quintals)	0.0658	0.1538	0.4300	0.6690	-0.2358	0.3673
cargo with trailer (up to 400 quintals)	0.1993	0.1685	1.1800	0.2370	-0.1310	0.5296
buses (13-45 seats)	0.2090	0.1642	1.2700	0.2030	-0.1127	0.5307
buses (above 46 seats)	0.7369	0.2262	3.2600	0.0010	0.2935	1.1802
others (motorcycle,bicycle,bajaj,....)	0.3392	0.1982	1.7100	0.0870	-0.0492	0.7277
Vehicle length of service						
0-5 years (ref.)						
5-10 years	1.2718	0.1553	8.1900	0.0000	0.9673	1.5762
above 10 years	1.3600	0.1563	8.7000	0.0000	1.0537	1.6664
Road condition						
Wet (ref.)						

dry	0.2336	0.0539	4.3300	0.0000	0.1279	0.3393
muddy	-0.1137	0.0402	-2.8200	0.0050	-0.1926	-0.0348
Accident time						
Afternoon (ref.)						
Morning	-0.0591	0.0976	-0.6100	0.5440	-0.2504	0.1321
Evening	-0.4084	0.0635	-6.4300	0.0000	-0.5329	-0.2839
Night	-0.5510	0.0684	-8.0600	0.0000	-0.6851	-0.4170
Location of accident						
Central part (ref.)						
Western zones (Walaga's zones)	-0.4846	0.0686	-7.0700	0.0000	-0.6190	-0.3502
South west zones (Jima,ilu aba bor,BunoBedele)	-0.1242	0.0339	-3.6600	0.0000	-0.1907	-0.0577
Eastern zones (Hararge's and Bale zone)	-0.1637	0.0276	-5.9400	0.0000	-0.2177	-0.1096
Southern zones (Arsi's,Guji's and Borana zone)	-0.1631	0.0365	-4.4700	0.0000	-0.2346	-0.0916
Environment of accident						
Factory (ref.)						
Residence place	-0.1413	0.2584	-0.5500	0.5840	-0.6477	0.3651
Commercial place	0.1653	0.1528	1.0800	0.2800	-0.1343	0.4648
Religious place	0.0953	0.1219	0.7800	0.4340	-0.1437	0.3343
Entertainment place	0.2454	0.2065	1.1900	0.2350	-0.1594	0.6501
School place	-0.1375	0.0379	-3.6300	0.0000	-0.2118	-0.0632
Hospital place	-0.1234	0.2580	-0.4800	0.6320	-0.6290	0.3822
Rural village area	-0.2013	0.0667	-3.0200	0.0030	-0.3320	-0.0705
Out of village in rural area	-0.1901	0.0555	-3.4200	0.0010	-0.2989	-0.0813
Government office area	0.2463	0.2493	0.9900	0.3230	-0.2423	0.7350
Road pavement						
Good asphalt (ref.)						
Defected asphalt	0.0726	0.0952	0.7600	0.4460	-0.1141	0.2592
Gravel	-0.5823	0.1390	-4.1900	0.0000	-0.8547	-0.3098
Site clearing	-0.4931	0.1369	-3.6000	0.0000	-0.7614	-0.2248
Type of accident						
Collision between vehicles (ref.)						
Collision between vehicle and pedestrian	0.1230	0.0644	1.9100	0.0560	-0.0032	0.2492
Collision between vehicle and animal	0.0531	0.0895	0.5900	0.5530	-0.1223	0.2285
Vehicle upside down	0.0893	0.0439	2.0300	0.0420	0.0033	0.1753
others (vehicle down,vehicle crash to inert)	0.0603	0.0575	1.0500	0.2940	-0.0524	0.1731
Accident cause						
not given priority to pedestrian (ref.)						
Over load	-0.0537	0.1166	-0.4600	0.6450	-0.2822	0.1748
Over speed	-0.0280	0.1442	-0.1900	0.8460	-0.3106	0.2546
Turning illegal position	-0.2968	0.1114	-2.6700	0.0080	-0.5151	-0.0786

Streeng problem	-0.3024	0.1536	-1.9700	0.0490	-0.6036	-0.0013
Brake problem	-0.0828	0.1737	-0.4800	0.6330	-0.4232	0.2575
Release of tyre	-0.0455	0.1424	-0.3200	0.7490	-0.3245	0.2336
following not keeping distance between	-0.1969	0.0786	-2.5000	0.0120	-0.3509	-0.0428
_cons	0.2857	0.0593	4.8200	0.0000	0.1695	0.4020
inflate						
Gender of driver						
Male (ref.)						
female	3.1396	1.2125	2.5900	0.0100	0.7632	5.5159
Age of driver						
18-30 (ref.)						
31-50	-11.4946	0.7422	-15.4900	0.0000	-12.949	-10.040
51 and above	-4.2680	0.4030	-10.5900	0.0000	-5.0579	-3.4780
Driving experience						
0-5 years(ref.)						
5-10 years	5.7501	0.7697	7.4700	0.0000	4.2415	7.2587
above 10	1.0657	0.3667	2.9100	0.0040	0.3469	1.7844
Vehicle-driver ownership						
Employee (ref.)						
Owner	1.6687	0.3295	5.0600	0.0000	1.0230	2.3145
others	-3.1799	0.5039	-6.3100	0.0000	-4.1676	-2.1922
Education level of driver						
elementary school and below (ref.)						
high school	-0.3079	0.3041	-1.0100	0.3110	-0.9040	0.2881
above high school	4.1937	0.6449	6.5000	0.0000	2.9297	5.4577
Vehicle Type						
Tax-minibus (upto 12 seats)						
Automobile	-5.7537	1.1306	-5.0900	0.0000	-7.9697	-3.5377
pick up(up to 10 quintals)	-4.6131	0.9663	-4.7700	0.0000	-6.5070	-2.7192
cargo(11-40 quintals)	-11.4733	1.0866	-10.5600	0.0000	-13.6030	-9.3435
cargo (41-100 quintals)	-17.0252	1.2883	-13.2100	0.0000	-19.5502	-14.5001
cargo with trailer (up to 400 quintals)	-23.8146	1.6276	-14.6300	0.0000	-27.0047	-20.6246
buses (13-45 seats)	-64.4934	2.350	-29.0210	0.0000	-67.0100	-60.2001
buses (above 46 seats)	-15.2323	1.6749	-9.0900	0.0000	-18.5150	-11.9495
others (motor cycle, bicycle, bajaj,....)	-13.3406	1.7017	-7.8400	0.0000	-16.675	-10.005
Vehicle length of service						
0-5 years (ref.)						
5-10 years	4.0192	0.7272	5.5300	0.0000	2.5938	5.4445
above 10 years	21.6155	1.8830	11.4800	0.0000	17.9249	25.3062
Road condition						
Wet (ref.)						

dry	-4.3969	0.2843	-15.4600	0.0000	-4.9541	-3.8396
muddy	1.9261	0.4043	4.7600	0.0000	1.1336	2.7185
Accident time						
Afternoon (ref.)						
Morning	-1.2342	0.2843	-4.3400	0.0000	-1.7914	-0.6769
Evening	6.7708	1.0802	6.2700	0.0000	4.6536	8.8880
Night	7.3881	1.0329	7.1500	0.0000	5.3636	9.4125
Location of accident						
Central part (ref.)						
Western zones (Walaga's zones)	2.8875	0.4967	5.8100	0.0000	1.9140	3.8610
South west zones (Jima,ilu aba bor,BunoBedele)	3.5848	0.3368	10.6400	0.0000	2.9247	4.2449
Eastern zones (Hararge's and Bale zone)	1.3412	0.3249	4.1300	0.0000	0.7045	1.9779
Southern zones (Arsi's,Guji's and Borana zone)	1.5368	0.3756	4.0900	0.0000	0.8007	2.2730
Environment of accident						
Factory (ref.)						
Residence place	-16.9679	2.6918	-6.3000	0.0000	-22.243	-11.692
Commercial place	-20.9205	2.1932	-9.5400	0.0000	-25.219	-16.621
Religious place	-12.0067	1.4203	-8.4500	0.0000	-14.790	-9.2230
Entertainment place	-22.6853	2.2679	-10.0000	0.0000	-27.130	-18.240
School place	-2.4735	1.2234	-2.0200	0.0430	-4.8713	-0.0757
Hospital place	-24.9231	2.5288	-9.8600	0.0000	-29.879	-19.967
Rural village area	13.1763	1.5182	8.6800	0.0000	10.2006	16.1520
Out of village in rural area	9.2956	1.4323	6.4900	0.0000	6.4884	12.1029
Government office area	2.5401	2.2017	1.1500	0.2490	-1.7752	6.8553
Road pavement						
Good asphalt (ref.)						
Defected asphalt	11.3755	0.8304	13.7000	0.0000	9.7479	13.0031
Gravel	9.6625	1.0116	9.5500	0.0000	7.6798	11.6452
Site clearing	9.0400	1.1171	8.0900	0.0000	6.8505	11.2296
Type of accident						
Collision between vehicles (ref.)						
Collision between vehicle and pedestrian	-7.1073	0.7624	-9.3200	0.0000	-8.6016	-5.6131
Collision between vehicle and animal	-7.4429	0.9634	-7.7300	0.0000	-9.3311	-5.5547
Vehicle upside down	-1.4371	0.5493	-2.6200	0.0090	-2.5137	-0.3605
others (vehicle down,vehicle crash to inert)	-0.1187	0.3674	-0.3200	0.7470	-0.8388	0.6013
Accident cause						
not given priority to pedestrian (ref.)						
Over load	17.0455	0.9878	17.2600	0.0000	15.1094	18.9815
Over speed	31.8953	1.2502	25.5100	0.0000	29.4450	34.3457
Turning illegal position	28.0742	1.2057	23.2800	0.0000	25.7110	30.4374

Streeng problem	-3.4919	2.5759	-1.3600	0.1750	-8.5406	1.5569
Brake problem	32.0821	1.2313	26.0600	0.0000	29.6688	34.4954
Release of tyre	36.5687	1.3509	27.0700	0.0000	33.9209	39.2164
following not keeping distance between	14.2939	0.7332	19.5000	0.0000	12.8569	15.7310
_cons	-6.9896	1.4613	-4.7800	0.0000	-9.8536	-4.1256
/lnalpha	-22.3723	0.0243	-919.510	0.0000	-22.420	-22.324
alpha	0.0000	0.0000		0.0000	0.0000	