38

ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF STATISTICS

DETERMINANTS OF TRAFFIC FATALITIES AND INJURIES IN ADDIS ABABA: A STUDY BASED ON BINARY AND ORDINAL LOGISTIC REGRESSION MODELS

A THESIS SUBMITTED TO THE GRADUATE PROGRAM OF ADDIS ABABA UNIVERSITY, STATISTICS DEPARTMENT IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

For

MASTER OF SCIENCE IN STATISTICS

By

BISRAT MISGANAW

June, 2010

Addis Ababa

Ethiopia

ADDIS ABABA UNIVERSITY
SCHOOL OF GRADUATE STUDIES
DEPARTMENT OF STATISTICS

DETERMINANTS OF TRAFFIC FATALITIES AND INJURIES IN ADDIS ABABA: A STUDY BASED ON BINARY AND ORDINAL LOGISTIC REGRESSION MODELS

By

BISRAT MISGANAW

APPROVED BY BOARD OF EXAMINERS:

DEPARTMENT HEAD	SIGNATURE	DATE
Mekonnen Tadesse EXAMINER	SIGNATURE	20/10/10 DATE
Prof. Eshetu Wencheko EXAMINER	SIGNATURE	20/10/10 DATE

Tel. 251-1-239465 P.O.Box 1176 Addis Ababa, Ethiopia

ACKNOWLEDGEMENT

First, and foremost, I thank God for giving me the opportunity to pursue my graduate study at the Department of Statistics, Addis Ababa University.

I would like to gratefully and sincerely thank Dr.Emmanuel G/ Yohannes, my thesis advisor and instructor, for his invaluable comments and suggestions that contributed to the successful realization of the study.

My special thanks also go to the Ministry of Education for sponsoring my study.

ABSTRACT

Background: Road traffic accidents (RTAs) are a major public health concern, resulting in an estimated 1.2 million deaths and 20-50 million injuries worldwide each year. In developing countries RTAs are among the leading cause of death and injury. Ethiopia experiences the highest rate of such accidents in Sub-Saharan Africa. Out of all the accidents registered in Ethiopia, Addis Ababa accounts for 60% on average.

Objective: The objective of this study is to identify the factors that contribute to the occurrence of traffic accidents leading to human injuries and death.

Method: This study applied binary and ordinal logistic regression models to identify factors influencing traffic fatalities and injuries. Stratified sampling with proportional allocation and simple random sampling technique are used to select samples from recorded frame. The data is obtained from Addis Ababa Traffic Control and Investigation Department (AATCID). A total of 443 accidents involving human injury that occurred within 365 consecutive days from July 8, 2008 to July 7, 2009 are sampled.

Results: Descriptive analysis indicates that 149 (33.6 percent), 196 (44.2 percent) and 98 (22.2 percent) are found to be slight injury, serious injury and fatal, respectively. Binary and ordinal logistic regression analyses showed that drivers aged 18-30 years caused the larger number of accidents followed by age group 31-50. Drivers with lower educational background, absence of lighting, wet surface and asphalt pavements, morning and evening hours, places like offices, residential and commercial neighborhoods, automobiles and taxies/minibuses were found to be strong and significant predictors of severity fatalities/serious injuries for traffic accidents in Addis Ababa.

Conclusion: For effective and safe traffic management, the concerned transportation authorities can consider the above mentioned predictors as potential causes of accidents in their order of importance. The findings of this study can serve as a basis for policymakers to create preventive measures for traffic accidents.

Table of contents **page**

Acknowledgments.......................................i

Abstract.......................................ii

List of tables.......................................v

Acronyms.......................................v

CHAPTER ONE: INTRODUCTION.......................................1

1.1 Background.......................................1
1.2 Statement of the problem.......................................2
1.3 Objectives of the study.......................................3
1.4 Significance of the study.......................................3
1.5 Limitation of the study.......................................4
1.6 Definition of terms.......................................4
1.7 Organization of the study.......................................5

CHAPTER TWO: LITERATURE REVIEW.......................................6

2.1 Factors associated with traffic accident.......................................7

2.1.1 Human factors.......................................7

2.1.2 Vehicle factors.......................................8

2.1.3 Environmental factors.......................................9

2.1.4 Roadway factors.......................................10

2.1.5 Other factors.......................................12

CHAPTER THREE: MATERIALS AND METHODOLOGY.......................................13

3.1 Source of data.......................................13
3.2 Data collection.......................................13
3.3 Variables included in the study.......................................14

a. The response variable.......................................14
b. Explanatory variables.......................................14
i. Human factors.......................................14
ii. Vehicle factors.......................................14
iii. Environmental factors.......................................14
iv. Roadway factors.......................................14
v. Other factors.......................................14
c. Description of variables included in the analysis.......................................15
3.4 Methodology.......................................19

3.4.1 Introduction to logistic regression analysis....19
3.4.2 Binary logistic regression model....19
3.4.3 Polytomous logistic regression model....20
3.4.4 Ordinal logistic regression model....21
3.4.5 Cumulative logit models....22
3.4.6 Estimation techniques....23
3.4.7 Model adequacy checking....23
3.4.7.1 Overall model fit....23
3.4.7.2 Test of goodness of fit....23
3.4.7.3 R-square estimate....25
3.4.7.4 Test of parallelism....27

CHAPTER FOUR: STATISTICAL DATA ANALYSIS....28
4.1 Major human, roadway, vehicle, and environmental factors with injury severity levels....28
4.2 Binary logistic regression analysis....31
4.2.1 Goodness of fit test....31
4.2.2 Classification table....31
4.2.3 Model estimation....32
4.3 Ordinal logistic regression analysis....35
4.3.1 Overall model fit....35
4.3.2 Goodness of fit test....35
4.3.3 Model summary....36
4.3.4 Classification table....36
4.3.5 Test of parallelism....37
4.3.6 Model estimation results....37

CHAPTER FIVE: DISCUSSION, CONCLUSION AND RECOMMENDATION....41
5.1 Discussion....41
5.2 Conclusion and Recommendation....45

REFERENCES....46
APPENDICES....48

List of tables

Page

Table 4.1: Distribution of human, vehicle, roadway and environmental factors with injury level.. 29

Table 4.2: Classification of injury severity levels for binary response..............................31

Table 4.3: Parameter estimates of the binary logistic model...33

Table 4.4: Classification of injury severity levels for ordinal response..............................36

Table 4.5: Parameter estimates of the ordinal logistic model..38

Abbreviations and Acronyms

AATCID	Addis Ababa Traffic Control and Investigation Department
AIDS	Acquired Immunodeficiency Syndrome
ML	Maximum Likelihood
MLE	Maximum Likelihood Estimate
OR	Odds Ratio
RTA	Road Traffic Accidents
RTI	Road Traffic Injuries
SPSS	Statistical Package for Social Sciences
UN	United Nations
US	United States
USA	United States of America
WHO	World Health Organization

CHAPTER ONE

INTRODUCTION

1.1 Background

Road traffic accident is defined as any vehicle accident occurring on a public highway. It includes collisions between vehicles and animals, vehicles and pedestrians, vehicles and fixed objects or vehicles and vehicles. Road traffic deaths accounted for 23% of all injury deaths worldwide in 2002. It has also been estimated that nearly 1.2 million people, male-to-female ratio being 2.34 to 1, are killed and 20-50 million people are injured or disabled each year in road traffic accidents. On average 3,300 road users are killed and about 100,000 are injured and/or disabled each day in traffic (WHO, 2004).

Out of all the accidents registered in Ethiopia, Addis Ababa holds accounts for 60% on average. This is partly because the city has five outlets that connect it to different regions. In addition to this about 77% vehicles in Ethiopia are registered here. Thus Addis Ababa, having a great concentration of vehicles and traffic, takes the lion's share in car accidents. Statistical data from the Addis Ababa Traffic Control and Investigation Department (AATCID) show that Addis Ababa is experiencing around 700 accidents per month and the costs of such fatalities and injuries due to traffic accidents have a great impact on various aspects of the society (Tesema and Tibebe, 2005).

Road traffic injures (RTIs) are growing as the vehicle use of developing-countries rises. By 2020, RTIs are expected to be the third leading cause of death and disability worldwide, by some calculations matching the toll of AIDS. Residents of developing countries are at much higher risk of RTIs than are residents of high-income countries. They are also at greater risk of death when a crash occurs. Developing countries also have inadequate trauma systems and are often unable to care for crash victims. Unless action is taken to improve road safety systems, poor countries will continue to bear the heavy toll of road traffic injuries (Lauren and Hill, 2005).

The 2004 World Health Report (WHO) shows that of the 1.2 million people killed in road crashes worldwide, 85% are in developing countries. Sub-Saharan Africa alone with only 4% of the global vehicle registered accounts for 10% of the road fatalities, and the economic, social and health consequences are grave. Conversely, the high income nations, with 60% of the total global vehicle fleet contribute only 14% of the annual road deaths. Human error, road environment and vehicle factors are reported by the traffic police as the main causes of road crashes. Two countries, South Africa and Nigeria, account for most of the reported deaths in Sub-Saharan Africa. The South Africa figure of over 9,000 has been consistent over time, while Nigeria with 6,185 deaths has declined from a high of over 9,200 in the early 1990s. Ethiopia, Kenya, Uganda, Tanzania and Ghana are the other countries that experience high number of road deaths (WHO, 2004).

Studies in Addis Ababa show that among the 3100 records which are collected within 310 consecutive days (04/03/06 – 07/01/07) from Addis Ababa Traffic Control and Investigation Department (AATCID), 1141(37%) involved human injuries. The rest are considered as accidents with material damage only. The highest mean number of injuries per accident (0.9) is obtained in residential areas by drivers in the age group of 18-30 who have elementary school level of education (Tewolde, 2007).

1.2 Statement of the Problem

At all levels, whether at national or international level, road traffic accidents continue to be a growing problem. In connection with this, according to a World Health Organization /World Bank Report, deaths from non-communicable diseases are expected to grow from 28.1 million a year in 1990 to 49.7 million by 2020, which is an increase in absolute number of 77%. Traffic accidents are the main cause of this rise. Road traffic injures are expected to take higher place in the rank order of disease burden in the near future.

The tragedy is more or less similar in Addis Ababa. The rate of traffic accidents in Addis Ababa increases with the increase of motor vehicles and population size. The rise in automobile ownership together with the poor condition of the roads has resulted in the high level of traffic safety and congestion problems. In Ethiopia, more than 1,800 people died while more than 7,000 were crippled or injured in 2003. Moreover the death rate is 136 per 10,000 vehicles. The share of Addis Ababa city in the total number of accidents was 60% in 1989 with annual average traffic accident growth of 34.4%. Nowadays Addis Ababa is experiencing around 700 accidents per month resulting in various levels of injury (Tesema and Tibebe, 2005).

Accidents are not random along the road network, and that drivers are not involved in accidents at random. There are complex circumstantial relationships between several characteristics (driver, road, vehicle, etc.). As such, one cannot improve safety without successfully relating accident frequency and severity to the causative variables (Yitambe, 2004).
This study explores the major human, vehicle, road and environmental factors that contribute to human fatalities and injuries in Addis Ababa.

1.3 Objectives of the Study

The general and specific objectives of the study are the following:

General Objective: To identify the major factors that contributes to the occurrence of traffic accidents leading to human injuries and death.

Specific Objectives:

- To classify the major correlates of the traffic fatalities into human, vehicle, road and environmental categories and investigate the impact of each.
- To demonstrate the application of the ordered logit to road safety problem.
- To identify the determinants of traffic accidents by using binary logistic regression model.

1.4 Significance of the Study

This study is mainly concerned with road traffic accident in Addis Ababa. Emphasis is given to study and measure the traffic behavior of pedestrians and drivers. In addition, problems related to the road environment, condition of vehicles and other factors will be identified. Therefore, the significance of the study can be stated as follows:

- ✓ The findings obtained from the study will provide information and knowledge about the patterns of road accidents in the city, which in turn, could help to develop countermeasures that could reduce the number and severity of accidents.
- ✓ The research outcome will add expertise to the existing body of knowledge.
- ✓ The research outcome will help policy makers to design appropriate strategies to reduce traffic accidents.
- ✓ The research will open the door for other researchers who are going to undertake further study in the area.

1.5 Limitation of the Study

The data from AATCID related to vehicle service years and driving experience are collected by interviewing the drivers themselves and the accuracy of the data is questionable to some extent.

1.6 Definition of Terms

Accident Frequency is a measure of the relative magnitude of a specific type of accident in relation to all accidents for a specified time period such as one or three years.

Accident Rate is a measure of the rate of occurrence of accidents in relation to any one of a number of exposure factors, such as per million vehicles entering or per million km in a study area.

Property Damage means physical damages of vehicles or properties but not to persons.

Fatal Accident is one where at least one person is killed with death occurring within 30 days of the accident.

Road Traffic Accident is defined as rare, random and multiple factor events always preceded by a situation in which one or more road users have failed to cope with the road and its environment.

Serious Injury means injury requiring admission to hospital as an in-patient.

Slight Injury is a traffic casualty, which is minor injury not requiring hospital admission as an in-patient.

Traffic Accident pertain to any vehicle accidents occurring on public high way (i.e. originating on, terminating on, or involving a vehicle partially).

1.7 Organization of the Study

This study contains five chapters. Chapter one presents background of the study, objectives, statement of the problem, its significance and limitation of the study. Chapter two deals with review of related literature on traffic accident. Chapter three discusses the material and methodology of the study such as sources of data and variables to be included in the study with their coding and description. Methods of data analysis are also described in this chapter. Chapter four presents statistical data analysis. Finally, discussion, summary of the main findings and conclusions of the study are dealt with in chapter five.

CHAPTER TWO

LITERATURE REVIEW

Road traffic injury is a major but neglected public health challenge that requires concerted efforts for effective and sustainable prevention. Of all the systems with which people have to deal with every day, road traffic systems are the most complex and the most dangerous. Worldwide, an estimated 1.2 million people are killed in road crashes each year and as many as 50 million are injured. Projections indicate that these figures will increase by about 65% over the next 20 years unless there is new commitment to prevention. In spite of this, the tragedy behind these figures attracts less media attention than other, less frequent types of tragedy. Every day around the world more than 3000 people die from road traffic injury. Low-income and middle-income countries account for about 85% of the deaths and for 90% of the annual disability adjusted life years (DALYs) lost because of road traffic injury. Projections show that, between 2000 and 2020, road traffic deaths will decline by about 30% in high-income countries but increase substantially in low-income and middle-income countries. Without appropriate action, by 2020, road traffic injuries are predicted to be the third leading contributor to the global burden of disease and injury (WHO, 2004).

Gekonge (1990) found that 85% of causes were attributed to human; while 6% to the vehicle and 9% to the traffic environment. The environmental factors include: road defects, animals, various obstruction and the weather. He found that road defects such as potholes do not seem to be major causes of road accidents against the belief that potholes and other road defects have been thought to be the main causes of road safety problems in Kenya. Running or walking into the road was found to be the most common cause of accidents where pedestrians are involved. Losing control, excessive speed, misjudging and overtaking improperly are the next in order.

Another study in Kenya also shows that the major causes of RTAs are driver error and pedestrians by considering police records from 1977 to 1996 (Yitambe, 2004). Passengers, pedal cyclists and vehicle defects accounted for a small number of road traffic accidents. Nantulya and Musiime (2001) observed that driver error was the common cause of RTAs accounting for 41% of the accidents on record. Other causes recorded include passengers, pedestrians, motor cyclists, and pedal cyclists. The human factor was highlighted as a primary cause of RTAs.

2.1 Factors Associated with Traffic Accident

Systems theory assumes that accidents are generated by numerous factors from different levels of the hierarchy. These factors often have interactions and confounding effects with each other. Such factors may be considered as accident causes if they either increase or decrease the probability of accident occurrence. Therefore, in order to prevent accidents, one must know which of the numerous traffic risk factors have a real strong influence on the number and probability of accidents. The major causes of road traffic accidents can be grouped under the following factors.

2.1.1 Human Factors

Tewolde (2007) used a Poisson regression model to identify factors that mainly determine the crash-related injuries. According to the results of his study, driver age is found to be significantly associated with the occurrence (number) of injuries. Among the age categories, drivers in the age group 18-30 are responsible for the larger number of injuries (51%) and the larger number of accidents (46%). These indicate that young drivers are accountable for the magnitude (severity) of accidents.

A study in Turkey using descriptive statistics found that there is a significant relationship between the age of drivers and the number of accidents. The accident rates of the young are more than the rest of the examinees. Furthermore, the age groups of 20-29 and 30-39 in the traffic are more than other age groups. It was also observed that as the age of drivers increases from the group of 40-49 to the group of 50-59, the rates of accidents decrease (Sari *et al.*, 2009). Similarly, Yitambe (2004) found that 39% of vehicles involved in the accidents in Kenya were driven by young men aged 18 to 24 years, meaning young drivers are reckless drivers and behave in a manner that threatens road safety.

Contrary to the above findings, a study in Tanzania (Akarro, 2009) using chi -square test found that there is no relationship between type of injury and age. This implies that irrespective of age, fatal accidents are equally likely in every age category.

According to a more recent study in Turkey, men are seen in accidents dominantly in comparison to women. The characteristic differences between male and female drivers are valuable in accident occurrences. Female drivers are more careful and responsible in traffic so the rates of their accident are less than that of men (Sari *et al.*, 2009). Similarly, a study in the USA using the Cox proportional survival model shows that male drivers' fatality rate was 55% higher than that of female drivers (Massie *et al.*, 1994). Female drivers' injury rate was 25% lower than that of males, but the total accident rate was 16% higher than that of male drivers.

Education is a very important parameter for traffic safety as is the case in most fields. Sari *et al.*, (2009) discussed the relationship between education and the number of accidents in Turkey. The highest risk groups for accident are people who graduated from primary school. For people who graduated from high school, the risk of accident decreases a bit; university graduates' risk of accident decreases highly. These results show that as the education level increases, the accident rate decreases.

A study in Finland shows that ownership of vehicle (own vehicle/not own vehicle) did not have a significant effect on drivers' accident time distributions. Own vehicles tended to be older, with tyres in poorer condition, their drivers either in the young or in the old age groups, and accumulated fewer kilometers in a year (Roine, 1999).

According to studies in Addis Ababa (Tewolde, 2007) with regards to the number of injuries and the number of accidents, drivers with 2-5 years of experience take the major share i.e. above 29%. Drivers with 5-10 years of experience and above 10 years of experience are having nearer figures when compared to the former one. When it comes to the number of injuries per accident the highest mean (.53) is attained by those with less than one year experience and the next higher mean (.50) is attained by those with 1-2 years of experience. But driver-vehicle relationship and sex of deriver were found to be statistically insignificant.

2.1.2 Vehicle Factors

Studies in Turkey using chi square test show that type of vehicle is one of the most important parameters on the number of accidents. It was seen that bicycles have the highest risk for accidents followed by motorcycles. When van, bus, minibus and lorry are considered, the number of accident decreases. Contrary to common belief, cars have the lowest risk for traffic and accident occurrence (Sari *et al.*, 2009). Similarly, in Pakistan the result of descriptive statistical analysis shows that among victims, 37% were motorcyclist, 23% bicyclist, 22% motorized two-wheelers and 18% bus users (Khan *et al.*, 2007).

Tewolde (2007) also examined the relationship between type of vehicle and the occurrences of traffic accidents in Addis Ababa by using Poisson model. According to the results, automobiles and taxies are responsible for the largest number of injuries with values 26.8% and 26%, respectively.

A study based on chi-square in Bahir Dar shows that 47.6 percent of all accidents are inflicted by private and governmental cars (automobiles and pickups). This is followed by trucks, which contribute 17.7 percent of accidents. Ranking third are taxies and minibuses that inflict 16.1 percent of the accidents (Yayeh, 2003).

A study in Finland by using Cox proportional survival model found that vehicle age was an insignificant survival variable. This implies that vehicle age did not have a major effect on the share of accident (Roine, 1999).

2.1.3 Environmental Factors

In Turkey the result of chi-square test shows that there is a significant relationship between the weather condition and the number of accidents. Most of the accidents happen on sunny days (Sari *et al.*, 2009).

Luoma and Sivak (1992) made comparisons of associated factors on traffic accidents in Finland and USA. They found that proportionally more fatal accidents occur in daylight conditions in the U.S. than in Finland. More than 85% of U.S. and more than 75% of Finnish fatal accidents occur in good conditions. Also more fatal accidents in winter conditions occur in Finland, consistent with the fact that the frequency of winter conditions is higher in Finland.

Mustakim *et al.,* (2008) used Multiple Liner Regression model to identify the effect of Hour of the Day on traffic accident in Malaysia. The highest number of accident occurred at 1601 to 1800 namely 4.01 pm to 6.00 pm, recorded for 754 accident cases and second highest was at 1401 to 1600 namely 2.01 pm to 4.00 pm involving 665 accident cases. The lowest recorded was at 0401 to 0600 or 4.01 am to 6.00 am with 69 accident case.

According to a study in Turkey (Sari *et al*., 2009), the "time" parameter was examined as month, day and hour in detail. The chi-square test demonstrates that the month of the year has no significant effect on the number of accidents. The result shows that each month has almost the same rate of accident. But, when day is considered, the results show that it is an effective factor on the number of accident. The number of accidents increases towards the weekends, even though the increase in the number of accidents on Fridays is slightly low. Sunday has the lowest accident number while most traffic accidents take place on Saturday. In a similar vein, the time of the day is an effective parameter for accidents. Comparison of two periods of peak traffic hours: 06:00–09:59 (the beginning of the working day) and 16:00–19:59 (the end of the working day) demonstrates that time period of 16:00–19:59 is riskier than the other one for traffic accidents.

A study in Oman shows that the highest number of accidents occurred on weekends compared to weekdays and accident number during weekdays does not show any significant variation (Shahalam *et al*., 2001).

A logistic model analysis in Calgary gives that wet surface condition resulted in a lower likelihood of serious injury to pedestrians. Due to short-lived summer season in the country, the road surface remains dry, providing an optimal condition for drivers to travel at a much higher speed. The favorable driving condition results in decreased driver concentration putting pedestrians at greater risk pedestrians (Anowar *et al.,* 2010).

2.1.4 Roadway Factors

By using descriptive statistics, a study of Patna city in India found that around 70% accidents occurred on straight roads. T-Junction seems to be the least prone to accident. The other junctions, Y-Junction, and 4-Way intersection, have almost equal share of accidents of around 8-10% (Singh and Misra, 1998).

Studies in Bahir Dar (Yayeh, 2003) revealed that 77 percent accidents have occurred on good-asphalted roads. Moreover, the records of police traffic data in the town show that 302 (87 percent) of accidents occurred on dry conditions of the road. The main types of road accidents that occur in urban areas related to road layouts are intersections, stretch of roads and along the roads. In the study area 275 (79 percent) of all road accidents occurred at stretches or along roads while 65 (18.7 percent) of the accidents occurred at junctions.

Tewolde (2007) used a Poisson regression model to identify the most influential places for traffic accident in Addis Ababa. Among several categories of places, both the number of injuries and number of accidents are the highest for office areas. But for the number of injuries per accident, residential areas are the places with the highest mean.

Accidents in Turkey mostly took place on main roads, and streets have the second-highest risk of accidents as investigated using the chi-square test. It should be noted that main roads are typically busy with traffic volume so this situation gives rise to increase in the number of accidents. The accidents happening on the main roads can be related to the urbanization and city planning (Sari *et al.*, 2009). Similarly, a chi-square test shows that accidents in Bahir Dar are usually highest where mixed traffic prevails, particularly on main streets of shopping streets, office areas and on older mixed residential roads with narrow rights way, and lowest on well designed roads with light traffics of outer areas of the town (Yayeh, 2003).

A logistic model analysis in Calgary shows that, the road class was found to have quite a significant effect on pedestrian safety. Although two-way divided roads have a better geometric design, such roads appeared to increase the serious injury risk to pedestrians. The physical separation of opposing traffic by medians may result in an increase in vehicular speed. Therefore, if a crash happens, a pedestrian is more likely to suffer from more severe injury due to the higher impact speed (Anowar *et al.*, 2010).

2.1.5 Other Factors (Reasons of Accident, Accident Type)

Studies in Patna city, India, indicate that pedestrian deaths as a percentage of all road fatalities are extremely high i.e. more than 90% of all road fatalities (Singh and Misra, 1998).

A study in Tirana using Multivariate-adjusted binary logistic models showed that younger age, high speed and especially alcohol consumption were significant predictors of fatal accidents. Fatal accidents were more prevalent on intercity roads and involved especially vans and trucks (Roshi *et al.*, 2008).

Among many causes of accidents in Oman, the main causes are negligence and high speed. The highest number of fatalities and injuries occur from head-on collisions (Shahalam *et al.*, 2001).

In Bahir Dar, pedestrians are the road users that are most affected by road accidents. On average, 52 percent of traffic casualties are pedestrians. Passengers, drivers and two-wheeled vehicles riders are also frequently involved in traffic accidents (Yayeh, 2003).

CHAPTER THREE

MATERIALS AND METHODOLOGY

3.1 Source of Data

Accidents are recorded by the Addis Ababa Traffic Control and Investigation Department (AATCID) on daily basis. This study is based on secondary data obtained from this department.

The data provide information on accidents that occur within 365 consecutive days from July 8, 2008 to July 7, 2009. The total number of accidents involving human injury recorded by AATCID in this period was 1669. Among these accident records, 560, 740 and 369 were registered for slight injury, serious injury and fatal, respectively. A pilot sample was used in order to test the recorded data and to determine and estimate the required sample size. From recorded AATCID cases 64, 80 and 45 pilot samples are taken for slight injury, serious injury and fatal, respectively. We have three strata: slight injury, serious injury and fatal. Stratified sampling with proportional allocation and simple random sampling technique are used to select samples from recorded frame for each stratum. By using the above technique 149, 196 and 98 samples respectively are obtained from AATCID data for slight injury, serious injury and fatal.

3.2 Data Collection

According to the definition of the Traffic Police Department in Addis Ababa, a road traffic accident is considered to be any event of human injury and/or death as a consequence of a physical collision between a responsible party (motor vehicle including motorcycle, car, van or truck, and/or bicycle) and an injured/ damaged party (motor vehicle(s); bicycle(s)/ bike(s); pedestrian(s); or any physical object, e.g., building or tree).

In cases when the crash involves a motor vehicle and/or bicycle with pedestrians, the information about the driver/rider of the motor vehicle/bicycle (age, sex, alcohol consumption, speed, type of motor vehicle) is recorded at the Traffic Police Department in Addis Ababa, regardless of the responsible party. When the crash involves two motor vehicles, the record is made only for the responsible party.

3.3 Variables Included in the Study

The following variables are considered in this study:

a. The Response Variable

The response variable of the study is the injury severity level of accidents involving humans, which is an ordered categorical variable categorized into three: slight injury, serious injury and death.

b. Explanatory Categorical Factors

The predictor factors to be analyzed as determinants of traffic accidents involving human injuries are grouped into human, vehicle, roadway, environmental, and other factors.

i. **Human Factors:** - Human factors are believed to have strong impact on traffic accident. Factors such as sex of driver, age of driver, driving experience, driver-vehicle relationship and driver educational background are included.

ii. **Vehicle Factors:** - Vehicle factors included in this study like vehicle direction, vehicle type, and vehicle age may have an influential impact on traffic accident as well as human injuries.

iii. **Environmental Factors:** - There are certain environmental factors that may increase or decrease the risk of traffic accidents. Atmospheric conditions, light condition, day of accident, time of accident, and road condition factors are included in this study.

iv. **Roadway Factors:** - Factors like place of accident, road situation, road junction, road class, and pavement roughness are included.

v. **Other Factors:** - Factors such as reason for accident and accident type are also included in the study.

c. Description and Coding of the Study Variable

Detailed description of human, vehicle, environmental, roadway factors to traffic human injuries are presented as follows.

Response Variable

Injury Severity Level	Binary Logistic Regression	0) Slight Injury 1) Serious Injury or Fatal
	Ordinal Logistic Regression	1) Slight Injury 2) Serious Injury 3) Fatal

Human Factors

No.	Factors	Categories
1	Gender of Driver	0) Male 1) Female
2	Age of Driver	0) 18-30 1) 31-50 and 2) 51 and above
3	Driver-Vehicle Relationship	0) Owner 1) Hired 2) Other (which may be relative, friend, someone who rented the vehicle …etc)
4	Driving Experience	0) Less than or equal to 1 year 1) Greater than 1 year and less than or equal to 2 years 2) Greater than 2 years and less than or equal to 5 years 3) More than 5 years

5	Driver Educational Background	0) Elementary school 1) Junior secondary school 2) Secondary school 3) Above secondary school

Vehicle Factors

No	Factors	Categories
6	Vehicle Type	0) Automobile 1) Taxies and Mini-buses(Up to 12 Seats) 2) Cargo(Upto 10 Quintals, 11-40 Quintals, 41-100 Quintals) 3) Buses (13-45 Seats, Above 46 Seats) 4) Station Wagon 5) Others (truck, cart, train,...,etc)
7	Vehicle Age	0) Less than or equal to 2 years 1) Greater than 2 years and less than or equal to 5 years 2) Greater than 5 years and less than or equal to 10 years 3) More than 10 years
8	Driving Direction	0) Straight going 1) Turning 2) Reverse going 3) Others (entering, 'U' bend, out from garden, crossing, parking)

Environmental Factors

No.	Factors	Categories
9	Atmospheric Conditions	0) Good condition 1) Rainy

10	Light Condition	0) Daylight 1) Dark but Lighted 2) Dark
11	Day of Accident	0) Weekend 1) Week day
12	Time of Event	0) Morning 1) Day 2) Evening 3) Night
13	Road Condition	0) Dry 1) Wet

Roadway Factors

No.	Factors	Categories
14	Pavement Roughness	0) Asphalt 1) Non- Asphalt
15	Road Junction	0) Non-Junction 1) Junction
16	Road Situation	0) Straight-flat 1) Not straight
17	Road Class	0) Roundabout 1) Two-way 2) One-way
18	Place of Accident	1) Office 2) Residential 3) Commercial centers 4) Religious centers 5) Recreation 6) School

Other Factors

No.	Factors	Categories
19	Accident Type	0) Crashing with pedestrian 1) Crash with another vehicle 2) Crashing with objects 3) Others (falling from vehicle, inverted etc.)
20	Reason of Accident	0) Rash and negligent driving 1) Wrong overtaking 2) No priority to pedestrian 3) No priority to vehicle 4) Others
21	Sub-City	1) Arada 2) Addis ketema 3) Lideta 4) Kirkos 5) Yeka 6) Bole 7) Akaki kaliti 8) Nifas silk 9) Kolfe 10) Gulele

3.4 Methodology

Multiple logistic regression analysis is used to study the effect of the predictors on the response variable-injury severity level of human controlling for other factors. The estimated coefficients tell us the increased or decreased chance of traffic human injuries given a set of levels of the independent factors while controlling for the effects of the other factors in the model. Estimates of odds ratio greater than 1.0 indicate that the risk of human injuries is greater in a category than that for the reference category. Estimates less than 1.0 indicate that the risk of human injuries is less than that for the reference category of each variable.

3.4.1 Logistic Regression Analysis

Regression methods are an integral component of any data analysis concerned with describing the relationship between a categorical response variable and one or more explanatory factors. It is often the case that the outcome variable is discrete, taking on two or more values. Logistic regression model can be used mainly for two reasons. The first is from a mathematical point of view, it is an extremely flexible and easily used function, and the second is that it leads itself to meaningful interpretation.

3.4.2 Binary Logistic Regression Model

To understand polytomous and ordinal logistic regression, one needs to understand the standard binary logistic regression model. The model can have an arbitrary number of parameters and terms in the model representing qualitative variables, quantitative variables, and interaction terms in order to model dichotomous or categorical outcome variable. When explanatory variables are included to model probabilities, a problem is that probabilities are restricted to the domain between 0 and 1, whereas a linear effect for an explanatory variable could take the fitted value outside this interval. Instead of the probability of an event, one may consider the odds: the ratio of the probability of success to the probability of failure. When the probability of success is p, the odds are $p/(1-p)$. In contrast to probabilities, odds can assume any value from 0 to infinity, and odds can be considered to constitute a ratio scale. The *logarithm* transforms a multiplicative scale to an additive scale.

Indeed, one of the most widely used transformations of probabilities is the *log odds*, defined by

$$\log it(p) = \ln\left(\frac{p}{1-p}\right),$$

where $\ln(x)$ denotes the natural logarithm of a positive number x. The logit function is an increasing function assuming values between 0 and 1, and it is defined over the set of real numbers. The logistic regression model is a model where $\log it(p)$ is a linear function of the explanatory variables.

For a binary response variable, the logit transformation of success probability, p_i of the i^{th} individual can be modeled as a linear combination of *k* explanatory variables

X_{1i}, X_{2i}, ..., X_{ki}, so that

$$\log it(p_i) = \log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki}.$$

This implies

$$p_i = \frac{e^{\beta_0+\beta_1 X_{1i}+\beta_2 X_{2i}+\ldots+\beta_k X_{ki}}}{1+e^{\beta_0+\beta_1 X_{1i}+\beta_2 X_{2i}+\ldots+\beta_k X_{ki}}}.$$

where a model parameter β_i will be interpreted as the change in the log odds for a one unit increase in x_i, holding all the other predictor constant.

3.4.3 Polytomous Logistic Regression Model

Polytomous logistic regression is the extension for the (binary) logistic regression when the categorical dependent outcome has more than two levels. One possible way to handle such situations is to split the categorical response variable in several ways and apply binary logistic regression to each dichotomous variable. However, this will result in several different analyses for only one categorical response. A more structured approach is to formulate one model for the categorical response by means of so-called *generalized logits.* Suppose that Y has k categories and the probability for category i is given by

$P(Y=i)= p_i \qquad i=1,\cdots,k$.

Then the k *generalized logits* are defined by

$$\log it(Y=i)= \ln\left(\frac{P_i}{1-(p_1+\ldots+p_{k-1})}\right)=\ln\left(\frac{P_i}{P_k}\right) \qquad i=1,\cdots,k-1.$$

This means that the generalized logits relate the probabilities P_i for the categories $i=1,\cdots,k-1$ to the reference category k. for m covariates the general polytomous logistic regression model becomes

$$\log it(Y=i)=\beta_{io}+\beta_{i1}X_1+\beta_{i2}X_2+\ldots+\beta_{im}X_m \qquad i=1,\cdots,k-1.$$

Note that the polytomous logistic model is given by *k-1* equations if Y has k category and that we have one logistic coefficient β_{ij} for each category/covariate combination. Hence it is not possible to summarize the effect of a covariate on the response Y by a single measure such as one odds ratio. Although the polytomous model offers the advantage of simultaneously testing the effect of a covariate on all response categories, polytomous logistic regression generates a cumbersome amount of statistical information.

3.4.4 Ordinal Logistic Regression Model

When the response variable Y is ordinal, the categories can be ordered in a natural way. The polytomous logistic regression model can be applied but does not make use of information about the ordering. One way to take account of the ordering is the use of *cumulative probabilities, cumulative odds* and *cumulative logits*. Considering k ordered categories, these quantities are defined by:

$$P(Y\le i)=p_1+p_2+\ldots+p_i$$

$$odds(Y\le i)=\left(\frac{P(Y\le i)}{1-P(Y\le i)}\right)=\frac{P_1+\ldots+P_i}{P_{i+1}+\ldots+P_k} \qquad i=1,\cdots,k-1.$$

Currently, the most popular model for ordinal responses uses logits of cumulative probabilities, often called *cumulative logits*.

$$\log it(Y \leq i) = \ln\left(\frac{\mathrm{P}(\mathrm{Y} \leq \mathrm{i})}{1 - P(Y \leq i)}\right) \qquad i = 1, \cdots, k.$$

The cumulative logistic model for ordinal response data is discussed below.

3.4.5 Cumulative Logit Models

Cumulative logit model is one of the most commonly used models for the analysis of ordinal categorical data and belongs to the class of generalized linear models. It is a generalization of a binary logistic regression model when the response variable has more than two ordinal categories. In *cumulative logit model*, the effect of β is the same for all i, often referred to as a *proportional odds model* (McCullagh and Nelder, 1989). In other words, the model assumes that the effect of each independent variable is the same for each cumulative probability. This model provides a single odds ratio (OR) estimate for all response categories, which can be obtained by exponentiation of the β coefficient. This estimate is quite convenient in terms of the model's ease of interpretation and parsimony.

The proportional odds model is used to estimate the odds of being at or below a particular level of the response variable. If there are i levels of ordinal outcomes, the model makes k-1 predictions, each estimating the cumulative probabilities at or below the i^{th} level of the outcome variable. This model can estimate the odds of being at or beyond a particular level of the response variable as well, because below and beyond a particular category are just two complementary directions. The model's threshold varies for each of the equations and satisfies the condition $\beta_{o1} \leq \beta_{o2} \leq \ldots \leq \beta_{ok-1}$, where $\beta_o's$ are cut points, usually nuisance parameters of little interest.

A model can simultaneously describe the effect of an explanatory variable on all the cumulative probabilities for y. This model is defined by:

$$\text{logit}\left[P(Y \leq i | x_1, x_2, ..., x_p)\right] = \ln\left[\frac{P(Y \leq i | x_1, x_2, ..., x_p)}{P(Y > i | x_1, x_2, ..., x_p)}\right] = \beta_{io} + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p \quad i = 1, \ldots, k-1.$$

When $\beta > 0$, this means that the corresponding independent variable is more likely to have higher values of y. Also the parameterization of this model is in accord with the usual formulation, in the sense that a positive β corresponds to a positive association (higher x tending to occur with higher y). In other words if $\beta > 0$ (odds greater than one) in this study this means increased likelihood of accident to be fatal/serious injury.

3.4.6 Estimation Techniques

The method used to calculate the logit coefficients is the Maximum likelihood estimation (MLE). To obtain the maximum likelihood estimate of β, a Fisher Scoring iterative estimation method or Newton-Raphson iterative estimation method can be used. ML methods maximize the log likelihood which reflects how likely it is (the odds) that the observed values of the dependent may be predicted from the observed values of the independents.

3.4.7 Model Adequacy Checking

3.4.7.1 Overall Model Test

Before proceeding to examine the individual coefficients, we want to look at an overall test of the null hypothesis that all coefficients for all of the variables in the model are zero.

In other words,

Ho: $\beta_1 = \beta_2 = \cdots = \beta_k = 0$

H1: Not all $\beta_i = 0$ (at least one of the predictors is significantly related to the response variable)

This can be tested using the likelihood ratio chi-square test at *k* degrees of freedom.

$\chi_c^2 = 2(\ln L_k - \ln L_0)$

where $\ln L_k$ is the log likelihood value for the model containing all *k* factors and $\ln L_0$ is the log likelihood value for the model containing only the intercept.

The null hypothesis is rejected if $\chi_c^2 \geq \chi^2_{\alpha,k}$ or the p-value is less than 0.05.

3.4.7.2 Test of Goodness Fit

Once a model has been developed, we would like to know how effective the model is in describing the outcome variable. This is referred to as goodness of fit test. In testing the hypothesis that the model fit the data, the two common approaches are Pearson's X^2 statistic and the likelihood-ratio statistic G^2 (see details in Agresti (1996)).

The test is written as:

Ho: The model adequately fits the data

H1: Not Ho

The null hypothesis is rejected if the p-value is less than 0.05.

Each model for a contingency table has a set of cell expected frequencies, which are numbers that perfectly satisfy the model and give the best fit to the observed counts. The model goodness of fit is tested by comparing the expected frequencies, denoted by f_e; to the observed frequencies f_0. The Chi-squared test statistic summarizes the discrepancies. Larger differences between the f_0 and f_e yield larger values of the test statistic and stronger evidence that the model is inadequate. Two chi-squared statistics, having similar properties, are commonly used to do this. The Pearson statistic is:

$$X^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

Another statistic, the likelihood-ratio statistic, is

$$G^2 = 2\sum f_o \log \frac{(f_o)}{f_e}$$

It equals the difference between the $(-2\log l)$ values for the model being tested and for the most complex model possible. Software for generalized linear models calls this statistic the deviance. Both X^2 and G^2 statistics equal 0 when there is a perfect fit (i.e., all $f_0 = f_e$). Large values indicate a poor fit.

Hosmer-Lemeshow Test: The null hypothesis for this test is that the model fits the data, and the alternative hypothesis is that the model does not fit the data. The test statistic is constructed by first divide the data into 10 groups. From each group, the observed and expected number of events will be computed.

Then, the test statistic is given by

$$\hat{C} = \sum_{k=1}^{g=10} \frac{(O_k - E_k)^2}{nP_k(1-P_k)}, \quad \text{where} \quad E_\kappa = nP_k, \quad P_k = \sum_{j=1}^{c_k} \frac{m_j p_j}{n_k}$$

C_k = number of covariate patterns in the k^{th} decile.

where g=10 is the number of groups, O_κ is the observed number of events in the κ^{th} group, E_κ is the expected number of events in κ^{th} group. If the observed number of events differs from what is expected by the model, the statistic $\hat{C}$ will be large and there will be evidence against the null hypothesis that the model is adequate to fit the data. This statistic has an approximate chi-square distribution with $(g-2)$ degrees of freedom. If p-value is greater than 0.05, we fail to reject the null hypothesis that there is no difference between observed and model predicted values, implying that the model estimates fit the data.

The Wald statistic is also an alternative test which is commonly used to test the significance of individual logistic regression coefficients for each independent variable (that is, to test the null hypothesis in logistic regression that a particular logit (effect) coefficient is zero i.e. H_0: $\beta_i = 0$ against $\beta_i \neq 0$). For a dichotomous independent variable, the Wald statistic (W) is

$$W = \frac{\hat{\beta}_i^2}{\operatorname{var}(\hat{\beta}_i)}.$$

For large sample size this statistic has an approximate chi-square distribution with one degree of freedom.

3.4.7.3 R-square Estimate

We cannot use the coefficient of determination on R^2 (defined for linear regression) in ordinal regression. In ordinal regression so called pseudo R-square gives information about how much variance is explained by the independent variables. However, in ordinal regression, variance is split into categories. Hence Cox and Snell's, Nagelkerke's, and McFadden's pseudo-R^2 statistics will be used in ordinal regression to estimate the variance explained by the independent variables.

- The Cox-Snell's R^2 (Cox and Snell, 1989)

$$R^2_{CS} = 1 - \left(\frac{L\left(\hat{\beta}^0\right)}{L\left(\hat{\beta}\right)} \right)^{2/n}$$

is based on the log likelihood for the model compared to the log likelihood for a baseline model. However, with categorical outcomes, it has a theoretical maximum value of less than 1, even for a "perfect" model.

- Nagelkerke's R^2 (Nagelkerke, 1991) is an adjusted version of the Cox-Snell R-square that adjusts the scale of the statistic to cover the full range from 0 to 1. It is given as

$$R^2_N = \frac{R_{cs}^2}{1 - L\left(\hat{\beta}^0\right)^{2/n}}$$

- McFadden's R^2 (McFadden, 1974) is another version, based on the log-likelihood kernels for the intercept-only model and the full estimated model, and is given by

$$R^2_M = \left(\frac{L\left(\hat{\beta}\right)}{L\left(\hat{\beta}^0\right)} \right)$$

where $L\left(\hat{\beta}\right)$ is the log-likelihood function for the model with the estimated parameters and $L(\hat{\beta}^0)$ is the log-likelihood with just the thresholds, and n is the number of cases (sum of all weights).

3.4.7.4 Test of Parallelism

We also want to test the assumption that the regression coefficients are the same for all response categories. For models without a scale component, the test of parallel lines helps to determine whether it is reasonable to assume that the values of the location parameters are constant across categories of the response.

The test of parallelism is:

Ho: The lines are parallel (slope coefficients are the same across response categories).

H1: Not Ho

This can be tested using the likelihood ratio chi-square test at k degrees of freedom.

$$\chi_c^2 = 2(\ln L_k - \ln L_0)$$

where $\ln L_k$ is the log likelihood value for the model with separate lines or planes and $\ln L_0$ is the log likelihood value for the constrained model, the model that assumes the lines are parallel. The null hypothesis is not rejected if the p-value is greater 0.05. That means the effect of the independent variable is the same for different logit functions. In other words, the slope for each factor included in the study for slight injury compared to serious injury/fatal, is the same as the slope for slight/serious injury compared to fatal categories. If the assumption of parallelism is violated, we should consider using multinomial regression, which estimates separate coefficients for each category.

CHAPTER FOUR

STATISTICAL DATA ANALYSIS

The analysis in this study pertains to searching the significant factors and investigating their effects. Binary logistic regression and ordinal logistic regression will be used for analysis. The statistical analyses are performed using Statistical Package for Social Sciences, version 15.0 (SPSS).

4.1 Major Human, Roadway, Vehicle and Environmental Factors with Injury Severity Levels

The major human, roadway, vehicle and environmental factors of the human injuries are presented in Table 4.1 below. The total number of human injuries covered in the study is 443, out of which, 149 (33.6 percent), 196 (44.2 percent) and 98 (22.2 percent) are found to be slight injury, serious injury and fatal, respectively.

Table 4.1: Distribution of human, vehicle, roadway, and environmental factors with injury levels

Variables		Injury severity level of accident							
		Slight injury		Serious injury		Fatal		Total	
		Count	Share	Count	Share	Count	Share	Count	Share
Age of Driver	18- 30	60	30.3%	91	46.0%	47	23.7%	198	44.7%
	31-50	61	31.9%	87	45.5%	43	22.5%	191	43.1%
	>= 51	28	51.9%	18	33.3%	8	14.8%	54	12.2%
Driving Experience	1 year at maximum	25	59.5%	11	26.2%	6	14.3%	42	9.5%
	1 - 2 years	23	56.1%	11	26.8%	7	17.1%	41	9.3%
	2 -5 years	36	32.4%	57	51.4%	18	16.2%	111	25.1%
	More than 5 years	65	26.1%	117	47.0%	67	26.9%	249	56.2%
Driver Educational Background	Elementary school	12	21.1%	27	47.4%	18	31.6%	57	12.9%
	Junior secondary school	20	30.3%	22	33.3%	24	36.4%	66	14.9%
	Secondary school	79	33.5%	107	45.3%	50	21.2%	236	53.3%
	Above secondary school	38	45.2%	40	47.6%	6	7.1%	84	19.0%
Atmospheric Conditions	Good condition	122	31.4%	178	45.9%	88	22.7%	388	87.6%
	Rainy	27	49.1%	18	32.7%	10	18.2%	55	12.4%
Time of Accident	Morning	42	23.5%	97	54.2%	40	22.3%	179	40.4%
	Day	21	40.4%	20	38.5%	11	21.2%	52	11.7%
	Evening	40	27.0%	68	45.9%	40	27.0%	148	33.4%
	Night	46	71.9%	11	17.2%	7	10.9%	64	14.4%
Pavement Roughness	Asphalt	131	31.9%	187	45.5%	93	22.6%	411	92.8%
	Non- Asphalt	18	56.3%	9	28.1%	5	15.6%	32	7.2%
Place of Accident	Office	50	25.8%	90	46.4%	54	27.8%	194	43.8%
	Residential	16	25.8%	25	40.3%	21	33.9%	62	14.0%
	Commercial	25	37.9%	29	43.9%	12	18.2%	66	14.9%
	Religious centers	20	46.5%	18	41.9%	5	11.6%	43	9.7%
	Recreation	16	36.4%	24	54.5%	4	9.1%	44	9.9%
	School	22	64.7%	10	29.4%	2	5.9%	34	7.7%
Accident Type	Crashing with pedestrian	80	25.2%	151	47.6%	86	27.1%	317	71.6%
	Crash with another vehicle	31	48.4%	25	39.1%	8	12.5%	64	14.4%
	Crashing with objects	22	66.7%	9	27.3%	2	6.1%	33	7.4%
	Others	16	55.2%	11	37.9%	2	6.9%	29	6.5%
Reason of Accident	Rash and negligent driving	38	36.9%	49	47.6%	16	15.5%	103	23.3%
	Wrong overtaking	26	50.0%	15	28.8%	11	21.2%	52	11.7%
	No priority to pedestrian	52	25.7%	104	51.5%	46	22.8%	202	45.6%
	No priority to vehicle	15	30.6%	18	36.7%	16	32.7%	49	11.1%
	Others	18	48.6%	10	27.0%	9	24.3%	37	8.4%

The proportion of injury severity levels on human beings, as can be seen in Table 4.1, differ by type of factors. Regarding sex, male drivers have a higher share from injury 96.2% (Appendix A).

As Table (4.1) shows, the severity level of injuries varies by age of driver. The young drivers (18-30) category accounted for the largest, 44.7%, share of injury. The oldest age group (over 51 years) had the smallest, 12.2%, share. When driving experience is considered with regards to the number of injuries, drivers with more than 5 years of experience take more than half of the share. Drivers with 2-5 years of experience take the second share (25%).

Among the four categories of educational background, those drivers with secondary school level education are responsible for the largest share of injuries (53.3%) and the minimum share is taken by drivers with elementary school level education (12.9 percent).

Table 4.1 shows that among environmental factors, atmospheric conditions and time of accident were also found to be the major factors for human injuries. The highest proportion of accidents are observed in the morning (40.4 percent) while in the evening (33.4 percent) and the majority of injuries are observed when atmospheric condition is good (87.6 percent).

Related to the road factor, 93% of the injuries have occurred on roadways with asphalt pavement, while the corresponding percent on non-asphalt roads was about 7%. The road junction distribution of crashes shows that human injuries at junctions and without junctions are 27% and 73%, respectively (Appendix A).

Accidents mostly took place around office buildings (areas) (44 percent), followed by commercial areas (15 percent), and residential areas (14 percent). When we come to types of traffic accidents, accidents crashing with pedestrian (71.6 percent) take the major share compare to other accident types. The reasons behind different factors of human injuries are also shown in Table (4.1). Among many causes for accidents, the highest is failure to give priority to pedestrian (45.6 percent) followed by rash and negligent driving (23.3 percent). Among ten sub-cities in Addis Ababa, The highest proportion of fatalities and injuries are observed in Bole (15.1 percent), followed by Yeka (14.2 percent) and Kirkos (12.6 percent) (Appendix A).

4.2 Binary Logistic Regression Analysis

Binary logistic regression analysis has been used to examine the effect of each independent variable on severity level of human injuries, while controlling for the other independent variables.

4.2.1 Goodness of Fit Test

Goodness of fit test for the fitted binary logistic regression model is assessed using the Hosmer and Lemeshow test. Accordingly, the Hosmer and Lemeshow test for the model resulted in p-values of 0.397, which tell us that the model adequately fit the data at 0.05 level of significance. Similarly, the model diagnostic checks presented in Appendix D do not show any violation of model assumption.

4.2.2 Classification Table

Table 4.2: Classification of injury severity levels for binary response

Observed			Predicted		
			Injury Severity Level		Percentage Correct
			Slight Injury	Fatal /Serious Injury	
Step 1	Injury Severity Level	Slight Injury	96	53	64.4
		Fatal /Serious Injury	27	267	90.8
	Overall Percentage				81.9

In addition to goodness-of-fit test, we need to look at the classification table which tells us how many of the cases where the observed values of the dependent variable are fatal/serious injury or slight injury have been correctly predicted. In this study, 64.4% were correctly classified for slight injury group and 90.8% for fatal/serious injury group. Overall 81.9% were correctly classified.

4.2.3 Model Estimation

According to the binary logistic regression results, out of 21 independent variables used in this study, eleven are found to be statistically significant with respect to injury severity levels.

The most important explanatory variables of injury severity in Addis Ababa are identified using forward conditional selection method in logistic regression.

The results of the model estimation are presented in Table 4.3. The table contains the significant variables, along with the estimated coefficients, standard errors of the estimate and p-value. It also includes the odd ratios for ease of interpretation. It should be noted that some of the variables within a factor were not significant but were retained in the models as long as at least one of the variables for the factor was statistically significant.

Table 4.3: Parameter estimates of the binary logistic model

Variables		Estimate	S.E.	Wald	df	Sig.	OR
Step 1(a)	**Age of Driver**			15.291	2	.000	
	18-30	1.758	.452	15.160	1	.000	5.804
	31-50	1.380	.428	10.401	1	.001	3.975
	>= 51 (ref.)						1.000
	Driving Experience			20.667	3	.000	
	1 year at maximum	-1.898	.478	15.780	1	.000	.150
	1 - 2 years	-1.356	.456	8.836	1	.003	.258
	2 -5 years	-.753	.344	4.805	1	.028	.471
	More than 5 years (ref.)						1.000
	Driver Educational Background			13.910	3	.003	
	Elementary school	2.112	.603	12.281	1	.000	8.268
	Junior secondary school	1.202	.484	6.178	1	.013	3.327
	Secondary school	.863	.346	6.237	1	.013	2.371
	Above secondary school (ref.)						1.000
	Vehicle Type			16.486	5	.006	
	Automobile	1.420	.445	10.180	1	.001	4.138
	Taxies and Minibuses	1.840	.523	12.357	1	.000	6.298
	Cargo	1.000	.457	4.793	1	.029	2.719
	Buses	.674	.515	1.716	1	.190	1.963
	Other vehicles	.807	.561	2.067	1	.151	2.240
	Station wagon (ref.)						1.000
	Atmospheric Conditions						
	Good condition	1.817	.492	13.646	1	.000	6.150
	Rainy (ref.)						1.000
	Time of Accident			17.165	3	.001	
	Morning	1.618	.403	16.147	1	.000	5.044
	Day	1.004	.496	4.103	1	.043	2.729
	Evening	1.389	.408	11.578	1	.001	4.009
	Night (ref.)						1.000
	Road Condition						
	Dry	-1.898	.634	8.967	1	.003	.150
	Wet (ref.)						1.000
	Pavement Roughness						
	Asphalt	1.591	.511	9.692	1	.002	4.907
	Non- Asphalt (ref.)						1.000
	Road Junction						
	No Junction	.602	.294	4.205	1	.040	1.826
	Junction (ref.)						1.000
	Place of Accident			18.254	5	.003	
	Office	1.875	.472	15.750	1	.000	6.522
	Residential	2.104	.573	13.496	1	.000	8.195
	Commercial	1.386	.535	6.717	1	.010	4.000
	Religious centers	1.404	.594	5.585	1	.018	4.073
	Recreation	1.430	.583	6.017	1	.014	4.180
	School (ref.)						1.000
	Accident Type			19.935	3	.000	
	Crashing with pedestrian	1.559	.492	10.038	1	.002	4.756
	Crash with another vehicle	.401	.566	.502	1	.479	1.494
	Crashing with objects (ref.)						1.000
	Others	.358	.677	.279	1	.597	1.430
	Constant	-7.718	1.228	39.469	1	.000	.000

The model shows that age of a driver is a significant variable. The odds of traffic accidents causing fatal/serious injury are higher for drivers in the age groups below 18-30 and 31-50 years as compared to those aged above 51 years.

As compared to drivers with driving experience of more than 5 years, the odds of traffic accidents causing fatal/serious injury for drivers with driving experience 1 year at maximum, 1-2 years and 2-5 years are decreased by a factor of 0.15, 0.26, and 0.47 times, respectively.

Regarding driver educational background, drivers with elementary, junior secondary and secondary level of education are about 8.3, 3.3 and 2.4 times more likely to get involved in fatal/ serious injuries as compared to those with above secondary school level of education, respectively. The results also show that the likelihood of traffic accidents causing fatal/serious injury is higher for drivers with automobiles (4.1 times), cargo (2.7 times), and taxies or minibuses (6.3 times) as compared to those driver with station wagon.

The model shows that traffic accidents at good conditions were 6.2 times more likely to cause fatal/serious injury as compared to those accidents in rainy condition. Moreover, time of the accident is one of the most important variables for severity level of traffic accidents. The odds of traffic accidents causing fatal/ serious injury are higher in the morning, day and evening as compared to accidents that occur at night. It is also observed that the odds of traffic accidents causing fatal/serious injury on dry roads are 0.15 times lower than those at wet roads.
The risk of fatal/serious injury is 5 times higher for drivers driving on asphalt roads as compared with drivers driving on non-asphalt roads. Road junction is also another important variable explaining the severity level of traffic accidents. The odds of traffic accidents causing fatal/ serious injuries at non-junction roads are 1.8 times more likely as compared to accidents at junction roads.

It is also observed that the odds of an accident causing fatal/serious injury due to crashing with pedestrians are higher as compared to crashing with objects. Place of accident is also another important variable explaining human fatalities and injuries in crash. Office and residential areas are more likely for fatal /serious injury as compared to school areas.

4.3 Ordinal Logistic Regression Analysis

In this study, ordinal logistic regression analysis is also used to examine the effect of each predictor in the model for causing of fatalities and injuries during traffic accident, while controlling for the other independent variables.

4.3.1 Overall Model Test

Model Fitting Information

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	940.052			
Final	612.603	327.449	56	.000

Link function: Logit.

Figure 4.1: Overall Model Test

Before proceeding to examine the individual coefficients, we want to look at the overall test of the null hypothesis that all coefficients for all of the variables in the model are zero. We can base this on the change in –2 log-likelihood when the variables are added to a model that contains only the intercept. Accordingly, in our model the difference between the two log-likelihoods (i.e, a chi-square value of 327.449) is significant at the 1% level. This means that we can reject the null hypothesis that the model without predictors is as good as the model with the predictors.

4.3.2 Goodness of Fit Test

Goodness of fit test for the fitted ordinal logistic regression models is assessed using the Pearson and deviance-based chi-square test. Accordingly, the Pearson test for the traffic model resulted in p-values of 0.116, which indicates that the model adequately fits the data at 0.05 level of significance.

4.3.3 Model Summary

The *Nagelkerke* modification of R-square is a more reliable measure of the relationship. *Nagelkerke's* R^2 will normally be higher than the *Cox and Snell* measure. Among the three measures *Nagelkerke's* R^2 is the most-reported of the R-squared estimates. In our case it is 0.594, indicating a moderately strong relationship of 59.4 % between the predictors and the prediction.

4.3.4 Classification Table

Table 4.4: Classification of injury severity levels for ordinal response

Observed		Predicted Response Category			Percentage Correct
		Slight Injury	Serious Injury	Fatal	
Injury Severity Level of Accident	Slight Injury	105	44	0	70.5%
	Serious Injury	26	152	18	77.6%
	Fatal	3	40	55	56.1%
Overall Percentage					70.4%

Cross tabulating the dependent variable with predicted values gives the above classification table, showing how often predicted values matched actual values. In this study by using ordinal logistic regression model, 70.5 % were correctly classified for slight injury group, 77.6 % for serious injury group and 56.1 % for the fatal group. Overall 70.4 % were correctly classified.

4.3.5 Testing Parallelism

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	612.603			
General	578.028(a)	34.575(b)	56	.989

When we fit an ordinal regression, we assume that the coefficients between the predictors and the logits are the same. According to our results, we do not reject the null hypothesis that the lines are parallel (p>0.05). In other words, the coefficients that describe the relationship between, say, the slight injury versus fatal/serious injury of the response variable are the same as those that describe the relationship between the slight/serious injury and fatal.

4.3.6 Model Estimation Results

According to the ordinal logistic regression results in Table 4.5, out of 21 independent variables used in this study, 12 are found to be statistically significant with respect to the injury severity levels.

The most important human, vehicle, environmental and road factors for fatalities and injuries are presented in Table 4.5 below. The table contains the significant variables, along with the estimated coefficients, standard errors of the estimate and p-value. It also includes the odd ratios.

Table 4.5: Parameter estimates of the ordinal logistic model

Variables		Estimate	S.E.	Wald	df	Sig.	OR
Threshold	[SEVERITYLEVEL = 1]	7.932	1.684	22.175	1	.000	
	[SEVERITYLEVEL = 2]	11.336	1.740	42.458	1	.000	
	Age of Driver	-.723	.160	20.359	1	.000	
	18-30	2.224	.444	25.110	1	.000	9.244
	31-50	1.437	.412	12.129	1	.000	4.208
	>= 51 (ref.)	0(a)	.	.	0	.	1.000
	Driving Experience	.546	.114	22.955	1	.000	
	1 year at maximum	-1.817	.443	16.805	1	.000	0.163
	1 - 2 years	-1.480	.430	11.832	1	.001	0.228
	2 -5 years	-1.004	.293	11.736	1	.001	0.366
	More than 5 years (ref.)	0(a)	.	.	0	.	1.000
	Driver Educational Background	-.658	.121	29.492	1	.000	
	Elementary school	2.107	.465	20.569	1	.000	8.224
	Junior secondary school	1.851	.433	18.291	1	.000	6.366
	Secondary school	1.123	.327	11.792	1	.001	3.074
	Above secondary school (ref.)	0(a)	.	.	0	.	1.000
	Vehicle Type	-.236	.064	13.782	1	.000	
	Automobile	1.759	.440	15.950	1	.000	5.807
	Taxies and Minibuses	1.493	.474	9.912	1	.002	4.450
	Cargo	.984	.445	4.894	1	.027	2.675
	Buses	1.462	.509	8.265	1	.004	4.315
	Other vehicles	1.242	.544	5.219	1	.022	3.463
	Station wagon (ref.)	0(a)	.	.	0	.	1.000
	Atmospheric Conditions	-1.529	.379	16.261	1	.000	
	Good condition	1.883	.443	18.065	1	.000	6.573
	Rainy (ref.)	0(a)	.	.	0	.	1.000
	Light Condition	1.024	.192	28.529	1	.000	
	Daylight	-3.075	.527	33.995	1	.000	0.046
	Dark but lighted	-2.555	.541	22.301	1	.000	0.078
	Dark (ref.)	0(a)	.	.	0	.	1.000
	Time of Accident	-.504	.103	24.014	1	.000	
	Morning	2.405	.455	27.936	1	.000	11.078
	Day	1.813	.523	12.022	1	.001	6.129
	Evening	1.852	.434	18.189	1	.000	6.373
	Night (ref.)	0(a)	.	.	0	.	1.000
	Road Condition	1.671	.440	14.440	1	.000	
	Dry	-1.881	.509	13.677	1	.000	0.152
	Wet (ref.)	0(a)	.	.	0	.	1.000
	Pavement Roughness	-1.752	.424	17.091	1	.000	
	Asphalt	1.754	.493	12.657	1	.000	5.778
	Non- Asphalt (ref.)	0(a)	.	.	0	.	1.000
	Road Junction	-.464	.232	4.020	1	.045	
	No Junction	.777	.274	8.038	1	.005	2.175
	Junction (ref.)	0(a)	.	.	0	.	1.000
	Place of Accident	-.263	.063	17.663	1	.000	
	Office	2.323	.488	22.693	1	.000	10.206
	Residential	3.126	.549	32.388	1	.000	22.783
	Commercial	2.074	.533	15.160	1	.000	7.957
	Religious centers	2.316	.601	14.827	1	.000	10.135
	Recreation	1.993	.568	12.301	1	.000	7.338
	School (ref.)	0(a)	.	.	0	.	1.000
	Accident Type	-.804	.133	36.323	1	.000	
	Crashing with pedestrian	3.062	.559	29.999	1	.000	21.370
	Crash with another vehicle	1.598	.615	6.744	1	.009	4.943
	Others	1.346	.713	3.561	1	.059	3.842
	Crashing with objects (ref.)	0(a)	.	.	0	.	1.000

This model shows that driver age is a significant variable. The likelihood of traffic accidents causing fatal/serious injury is higher for young drivers (aged below 30 years and aged 31-50 years) compared with the elderly drivers (aged above 51 years).

As compared to drivers with driving experience of more than 5 years, the odds of traffic accidents causing fatal/serious injury for drivers with driving experience 1 year at maximum, 1-2 years and 2-5 years were reduced by a factor of 0.16 0.23, and 0.37 times, respectively.

Driver educational background is also significantly associated with severity level of traffic accidents. Drivers with elementary, junior secondary and secondary level of education are about 8.2, 6.4, and 3.1 times more likely to get involved in fatal/serious injuries as compared to those with above secondary school level of education, respectively.

When we consider the result of Table 4.5 above, it demonstrates that type of vehicle is one of the most important factors for fatal and serious injury occurrence. The odds of traffic accidents causing fatal/serious injury is higher for drivers with automobiles (5.8 times), buses(4.3 times), cargo(2.7 times), and taxies or minibuses (4.5 times) as compared to those driver with station wagon.

Our study shows that the atmospheric condition is significantly associated with traffic accidents. The likelihood of traffic accidents causing fatal/serious injury is higher during clear conditions as compared with rainy condition. The result also indicates that there is a significant relationship between light condition and severity level of accidents. The odds of traffic accidents causing fatal/serious injury are higher for drivers driving in dark condition as compared to daylight and dark but lighted condition.

Moreover, time of the accident is one of the most important variables for severity level of traffic accidents. The odds of traffic accidents causing fatal /serious injury are higher in the morning, day and evening as compared to accidents at night. Regarding road condition, the likelihood of fatal/serious injury is higher on wet road condition as compared to dry condition.

In terms of road surface condition, this study found out that drivers driving on asphalt roads are about 5.8 times more likely to get involved in fatal /serious injury accidents as compared to those driving on non-asphalt roads. When we consider the contribution of road junction to traffic accident, the odds of traffic accidents causing fatal/serious injuries at no-junction roads are 2.2 times higher as compared to accidents at junction roads.

Furthermore, the odds of accident being fatal /serious injury due to crashing with pedestrians, and crashing with another vehicle were more likely as compared to crashing with objects. Place of accident is also another important variable explaining human fatalities and injuries in crash. Offices and residential areas are more likely for fatalities/serious injuries as compared to school areas (surroundings).

CHAPTER FIVE

DISCUSSION, CONCLUSION AND RECOMMENDATION

5.1 DISCUSSION

This study is designed to identify some of the determinants of the fatalities and injuries due to traffic accidents based on AACTID 2009 data. Binary logistic regression and ordinal logistic regression have been used to analyze the data. The effect of the factors of interest on the severity level will be discussed here in detail.

A. Human Factors

People are the most active actors in the system. While their actions might be influenced by subconscious motives and subliminal cues, they are also the most adaptive element in the traffic system. They can create risk situations as well as respond to ever changing new demands of the traffic environment.

The findings of the study show that accident rates of young drivers are more than the rest of the examinees. Similar results are also reported from other studies (Sari *et al.*, 2009; Yitambe, 2004). This may be related to reckless driving, psycho-biological immaturity, an excessive belief in one's own abilities, lack of experience, driving culture, and lifestyle induced risky type of exposure such as night driving.

According to the findings, drivers who have more than 5 years of experience are significantly exposed to traffic accidents causing fatal/serious injury than those who have less driving experience. This could be due to overconfidence and carelessness.

Driver educational background is found to have quite a significant effect on human safety. Table 4.5 shows that the highest risk group for accident is drivers with primary school educational level. The risk of accident decreases a bit among junior high school graduates. It can be clearly seen from the table that for drivers with above secondary school level of education, the risk of accident decreases more. These results show that as the education level increase, the fatal/serious injury accident rate decreases. This finding is consistent with other studies (Sari *et al.*, 2009). This indicates that education is a considerable factor to prevent road traffic accidents. Thus, it can be said that understanding, interpreting and obeying the regulations are parallel to education and behaving more logical.

B. Vehicle Factors

Among vehicle factors included in this study, vehicle type was found to be an important factor which affects human safety due to traffic accidents. Even if buses and taxies/ minibuses play an essential role in public transportation, our results showed that these vehicles (in addition to automobiles, cargo vehicles and buses) pose a significantly greater fatal/serious injury risk to pedestrians. Note that cargo vehicles have larger mass, greater momentum and longer stopping distance. For any given speed, the greater the mass of the vehicle, the greater would be its force of impact at collision with the pedestrians leading to higher injury severities. Furthermore, it is possible that drivers of small vehicles but high speedy are more likely to weave around in traffic, change lanes, dart ahead of others or even take corners and curves faster. This finding seemed to be in accordance with other studies (Tewolde, 2007).

C. Environmental Factors

According to the results, atmospheric condition is found to be another significant factor in determining the injury severity. As can be seen in Table 4.5, most of the accidents happened in good conditions. It is obvious that the movements of vehicles in the traffic also depend on weather conditions among many parameters. When it is rainy, drivers are probably more careful.

Statistical tests showed that lighting condition is a risk factor affecting crash severity. Poor light conditions contributed to causing fatal/serious injuries, which indicates that poor light conditions could increase the probability of causing fatalities when a crash occurs. This is possibly due to visibility at night in the absence of street light is limited by the range of headlights and glare from oncoming vehicles headlights. In addition, drivers are also unable to distinguish pedestrians from the shaded surroundings due to their darker outfits. All these can lead to drivers braking later or taking less effective avoidance maneuvers leading to increased risk of crash and serious injury. This finding is consistent with other studies (Luma and Sivak, 1992).

In our study, the time of accident is significantly associated with fatal/serious injury risk. The odds of traffic accidents causing fatal/serious injury are higher in the morning, during the day, and in the evening as compared to accidents during the night. This finding is in accordance with other studies (Sari *et al.*, 2009). This indicates that in the morning (the beginning of the working day) and evening (the end of the working day) the roads are typically busy with traffic volume, and during daytime, there are increased activities such as commercial activities and work related (office) activities, etc. So this situation gives rise to increase in the number of accidents.

In terms of road surface condition, this study found out that wet surface condition resulted in a higher likelihood of fatal/serious injury than dry surface. This could be because on wet road surfaces it is difficult to stop vehicles easily. This implies that the possibility of death/serious injury will be high.

D. Roadway Factors

The environment with its roadway network creates the framework for the behavior of traffic and exposes those who are on the network to various accident risks. Traffic environment can support and promote safe behavior, but it can also encourage or lead to risky behavior.

The findings of the study show that the majority of the fatal/serious injuries have occurred on roadways with asphalt pavement. This finding is consistent with previous studies (Yayeh, 2003). This is because of asphalt roads may result in an increase in vehicular speed. Therefore, if a crash happens, a pedestrian is more likely to suffer from more severe injury due to the higher impact speed.

Road junction is an important roadway factor that contributes to severity level of traffic accidents. The findings of this study show that the odds of traffic accidents causing fatal/serious injuries at no-junction roads are higher as compared to accidents at junction roads. This finding is consistent with the study by Singh *et al.*, 1998.

This could be due to the reason that at junction not only drivers but also pedestrians are more careful and also there may be traffic lights at junctions.

According to the results, place of accident was significantly associated with fatal/serious injury risk. Accidents mostly took place on offices, residential and commercial neighborhoods. It should be noted that these places are typically busy with traffic volume giving rise to increase in the number of accidents. A study by Tewolde (2007) also reported the same result.

The findings of this study also show that the odds of accident being fatal/serious injury due to crashing with pedestrians, and crashing with another vehicle were higher as compared to crashing with objects. As expected, the serious injury risk significantly increased for pedestrians who tended to cross roadway without proper right of way. Since drivers do not expect pedestrians at locations not designated for pedestrians, they might fail to detect the pedestrian in time and take evasive actions accordingly. This might result in higher impact speed during collision leading to greater injury risk. Negligent crossing behavior has been reported to be a major cause of pedestrian collision and injury by Yayeh (2003).

5.2 CONCLUSION AND RECOMMENDATION

Drivers with driving experience of more than 5 years are significantly exposed to traffic accidents causing fatal/serious injury. This could be due to overconfidence and carelessness.

Young drivers aged 18-30 years and aged 31-50 year, drivers with lower educational background, at no-junction roads, absence of lighting, wet surface and asphalt pavements are more likely to result in fatal/serious injuries.

Accidents occurring in the morning and evening and accidents that occur in office areas, residential and commercial neighborhoods are more likely to result in fatal/serious injuries. Among types of vehicles, automobiles and taxies/minibuses create fatal or serious injury risk.

Gender of driver, vehicle ownership, vehicle age, driving direction, day of accident, road situation, road class, reason of accident and sub-city have no significant association with the influence of fatalities/serious injuries during traffic accident.

It can be learnt from this study that special attention should be given to reduce the severity of accidents by taking the following into consideration.

- ✓ Pedestrian crossings could be raised to lower vehicular speed and advanced stop line could be used to reduce the conflicts at crosswalks, thereby improving pedestrian safety.
- ✓ By avoiding overspeed, respecting pedestrian priority, and follow the traffic rules and regulations drivers may reduce traffic accidents.
- ✓ Since, there is a high volume of traffic in the morning (the beginning of the working day), evening (the end of the working day) and areas around offices, residential and commercial neighborhoods drivers should be more careful in such areas and times of the day.

References

[1] Agresti, A. (1996). *An Introduction to Categorical data Analysis*. John Wiley and Sons, Inc., New York.

[2] Akarro, R. R. (2009). Peoples Opinions on the Causes of Road Accidents in Selected Places of Tanzania. *European journal of Social Sciences*; 9(4): 615-626.

[3] Anowar, S., Yasmin, S. and Tay, R. (2010). A Logistic Model for Examining Pedestrian Serious Injury Risk in Traffic Crashes Department of Civil Engineering, University of Calgary.

[4] Cox, D. R. and Snell, E. J. (1989). *The Analysis of Binary Data*, Second Edition, London: Chapman and Hall.

[5] Gekonge, N. (1990). Traffic accident statistics and their economic implications in Kenya. *East African Medical Journal*; 12(1): 187-220.

[6] Khan, M.H., Ahmed, I., Zia, N., Babar,T.S. and Babar, K.S. (2007). Road traffic accidents; study of risk factors Community Medicine Department Gomal Medical College; 14(2): 323-327.

[7] Lauren, P. and Hill, S. (2005). Road traffic Injuries-Can we a global epidemic? Retrieved from the World Wide Web http://www.thedoctorwillseeeyounow.com/articles/other/road-33/.

[8] Luoma, J. and Sivak, M. (1992). Traffic accidents in Finland and the U.S.A: A cross-cultural comparison of associated factors 4: 1-25.

[9] Massie, D., Campbell, K. L. and Williams, A. (1994). Traffic accident involvement rates by driver age and gender. Accident Analysis and Prevention 27(1): 73–87.

[10] McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models*, Second Edition, New York: Chapman and Hall.

[11] McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior, in P. Zarembka (ed.), *Frontiers of Econometrics*, Acadamic Press.

[12] Mustakim, F., Yusof, I., Rahaman, I. and Samad, A.A. (2008). Black spot Study and Accident Prediction Model Using Multiple Liner Regression with Special Reference to Malaysia 7(3): 66-83.

[13] Nagelkerke, N. J .D. (1991). A note on a generalized definition of the coefficient of determination. *Biometrika*; 78: 691-692.

[14] Nantulya, V. and Muli-Musiime, F. (2001). Kenya: Uncovered the social Determinants of Road Traffic Accidents. 11:119-132.

[15] Roine, M. (1999). Accident risks of car drivers in wintertime traffic. Case of Finland. Retrieved from the World Wide Web http://www.inf.vtt.fi/pdf/.

[16] Roshi, E., Qirjako, G., Hysa, B. and Burazeri, G. (2008). Factors Associated with Fatal Traffic Accidents in Tirana, Albania: Cross-sectional Study 49: 734-740.

[17] Sari, M., Mutlu,O. and Zeytinoglu, A. (2009). Effects of Human and External Factors on Traffic Accidents in Turkey; 21(1): 9-17.

[18] Shahalam, A.M., Ali, G., Bakheit, C.S. and Fadhel, E.B. (2001). Factors Analysis: traffic Accidents and Potential Causes- a Case Study in Oman 5(8): 1-94.

[19] Singh, S.K. and Misra, A. (1998). Road accident analysis: A case Study of Patna City. *Urban Transport Journal*; 2(2): 60-75.

[20] Tesema, T. and Tibebe, B. (2005). Rule Mining and Classification of Road Accidents Using Adaptive Regression Trees. International Journal of Simulation Systems Science &Technology Special Issue on Soft Computing for Modeling and Simulation; 6: 10-11.

[21] Tewolde, M. (2007). Analysis on traffic accidents involving human injuries in the case of Addis Ababa. (Unpublished M.Sc. Thesis), Department of Statistics, Addis Ababa University.

[22] World Health Organization and the World Bank (2004). World report on road traffic injury prevention.

[23] Yayeh, A. (2003). The extent, variations and causes of road traffic accidents in Bahir Dar. (Unpublished M.A. Thesis), Department of Geography, Addis Ababa University.

[24] Yitambe, A. (2004). Disaster Avoidance Strategies in Kenya: The Case of Road Traffic Accidents; 3:1-11.

APPENDIX

Appendix [A]

Distribution of human, vehicle, roadway, and environmental factors with injury levels

		Injury severity level of accident							
		Slight injury		Serious injury		Fatal		Total	
		Count	Share	Count	Share	Count	Share	Count	Share
Sex of Driver	Male	144	33.8%	189	44.4%	93	21.8%	426	96.2%
	Female	5	29.4%	7	41.2%	5	29.4%	17	3.8%
Age of Driver	18-30	60	30.3%	91	46.0%	47	23.7%	198	44.7%
	31-50	61	31.9%	87	45.5%	43	22.5%	191	43.1%
	>= 51	28	51.9%	18	33.3%	8	14.8%	54	12.2%
Driver-Vehicle Relationship	Owner	22	30.6%	26	36.1%	24	33.3%	72	16.3%
	Hired	94	33.0%	127	44.6%	64	22.5%	285	64.3%
	Other	33	38.4%	43	50.0%	10	11.6%	86	19.4%
Driving Experience	1 years at maximum	25	59.5%	11	26.2%	6	14.3%	42	9.5%
	1 - 2 years	23	56.1%	11	26.8%	7	17.1%	41	9.3%
	2 -5 years	36	32.4%	57	51.4%	18	16.2%	111	25.1%
	More than 5 years	65	26.1%	117	47.0%	67	26.9%	249	56.2%
Driver Educational Background	Elementary school	12	21.1%	27	47.4%	18	31.6%	57	12.9%
	Junior secondary school	20	30.3%	22	33.3%	24	36.4%	66	14.9%
	Secondary school	79	33.5%	107	45.3%	50	21.2%	236	53.3%
	Above secondary school	38	45.2%	40	47.6%	6	7.1%	84	19.0%
Vehicle Type	Automobile	34	32.4%	50	47.6%	21	20.0%	105	23.7%
	Taxies and Minibuses	18	19.6%	49	53.3%	25	27.2%	92	20.8%
	Cargo	31	31.3%	49	49.5%	19	19.2%	99	22.3%
	Buses	22	37.9%	18	31.0%	18	31.0%	58	13.1%
	Station wagon	26	52.0%	20	40.0%	4	8.0%	50	11.3%
	Other vehicles	18	46.2%	10	25.6%	11	28.2%	39	8.8%
Vehicle Age	2 years at maximum	18	36.7%	23	46.9%	8	16.3%	49	11.1%
	2 -5 years	34	40.0%	41	48.2%	10	11.8%	85	19.2%
	5 -10 years	62	35.4%	66	37.7%	47	26.9%	175	39.5%
	Over 10 years	35	26.1%	66	49.3%	33	24.6%	134	30.2%
Vehicle Direction	Straight going	123	32.7%	173	46.0%	80	21.3%	376	84.9%
	Turning	12	44.4%	6	22.2%	9	33.3%	27	6.1%
	Reverse going	11	47.8%	9	39.1%	3	13.0%	23	5.2%
	Others	3	17.6%	8	47.1%	6	35.3%	17	3.8%
Atmospheric	Good condition	122	31.4%	178	45.9%	88	22.7%	388	87.6%

Conditions	Rainy	27	49.1%	18	32.7%	10	18.2%	55	12.4%
Light Condition	Daylight	97	31.4%	158	51.1%	54	17.5%	309	69.8%
	Dark but lighted	48	48.5%	34	34.3%	17	17.2%	99	22.3%
	Dark	4	11.4%	4	11.4%	27	77.1%	35	7.9%
Day of Accident	Weekend	67	43.5%	58	37.7%	29	18.8%	154	34.8%
	Week day	82	28.4%	138	47.8%	69	23.9%	289	65.2%
Time of Accident	Morning	42	23.5%	97	54.2%	40	22.3%	179	40.4%
	Day	21	40.4%	20	38.5%	11	21.2%	52	11.7%
	Evening	40	27.0%	68	45.9%	40	27.0%	148	33.4%
	Night	46	71.9%	11	17.2%	7	10.9%	64	14.4%
Road Condition	Dry	141	34.8%	178	44.0%	86	21.2%	405	91.4%
	Wet	8	21.1%	18	47.4%	12	31.6%	38	8.6%
Pavement Roughness	Asphalt	131	31.9%	187	45.5%	93	22.6%	411	92.8%
	Non- Asphalt	18	56.3%	9	28.1%	5	15.6%	32	7.2%
Road Junction	No Junction	98	30.2%	147	45.4%	79	24.4%	324	73.1%
	Junction	51	42.9%	49	41.2%	19	16.0%	119	26.9%
Road Situation	Straight-flat	133	32.8%	184	45.3%	89	21.9%	406	91.6%
	Not straight	16	43.2%	12	32.4%	9	24.3%	37	8.4%
Road Class	Roundabout	78	33.2%	105	44.7%	52	22.1%	235	53.0%
	Two -way	57	32.9%	79	45.7%	37	21.4%	173	39.1%
	One -way	14	40.0%	12	34.3%	9	25.7%	35	7.9%
Place of Accident	Office	50	25.8%	90	46.4%	54	27.8%	194	43.8%
	Residential	16	25.8%	25	40.3%	21	33.9%	62	14.0%
	Commercial	25	37.9%	29	43.9%	12	18.2%	66	14.9%
	Religious centers	20	46.5%	18	41.9%	5	11.6%	43	9.7%
	Recreation	16	36.4%	24	54.5%	4	9.1%	44	9.9%
	School	22	64.7%	10	29.4%	2	5.9%	34	7.7%
Accident Type	Crashing with pedestrian	80	25.2%	151	47.6%	86	27.1%	317	71.6%
	Crash with another vehicle	31	48.4%	25	39.1%	8	12.5%	64	14.4%
	Crashing with objects	22	66.7%	9	27.3%	2	6.1%	33	7.4%
	Others	16	55.2%	11	37.9%	2	6.9%	29	6.5%
Reason of Accident	Rash and negligent driving	38	36.9%	49	47.6%	16	15.5%	103	23.3%
	Wrong overtaking	26	50.0%	15	28.8%	11	21.2%	52	11.7%
	No priority to pedestrian	52	25.7%	104	51.5%	46	22.8%	202	45.6%
	No priority to vehicle	15	30.6%	18	36.7%	16	32.7%	49	11.1%
	Others	18	48.6%	10	27.0%	9	24.3%	37	8.4%
Sub-city of the Accident	Arada	12	35.3%	12	35.3%	10	29.4%	34	7.7%
	Addis ketema	11	31.4%	16	45.7%	8	22.9%	35	7.9%
	Ledeta	10	37.0%	10	37.0%	7	25.9%	27	6.1%
	Kirkos	15	26.8%	31	55.4%	10	17.9%	56	12.6%
	Yeka	21	33.3%	27	42.9%	15	23.8%	63	14.2%

Bole	22	32.8%	34	50.7%	11	16.4%	67	15.1%
Akaki	16	31.4%	20	39.2%	15	29.4%	51	11.5%
Nefas silk	23	45.1%	22	43.1%	6	11.8%	51	11.5%
Kolfie	9	23.7%	15	39.5%	14	36.8%	38	8.6%
Gulele	10	47.6%	9	42.9%	2	9.5%	21	4.7%

Appendix [B]

Parameter estimates of the binary logistic model

		Estimate	S.E.	Wald	df	Sig.	OR
Step 1(a)	**Driver Sex**						
	Male	.193	.839	.053	1	.818	1.213
	Female						1
	Age of Driver			16.672	2	.000	
	18-30	2.190	.540	16.470	1	.000	8.938
	31-50	1.818	.514	12.517	1	.000	6.162
	>= 51						1
	Driver-Vehicle Relationship			1.745	2	.418	
	Other	-.661	.508	1.697	1	.193	.516
	Hired	-.476	.459	1.073	1	.300	.621
	Owner						1
	Driving Experience			22.224	3	.000	
	1 years at maximum	-2.239	.543	17.039	1	.000	.107
	1 - 2 years	-1.611	.518	9.665	1	.002	.200
	2 -5 years	-.768	.375	4.191	1	.041	.464
	More than 5 years						1
	Driver Educational Background			10.106	3	.018	
	Elementary school	1.958	.651	9.044	1	.003	7.084
	Junior secondary school	1.126	.532	4.483	1	.034	3.084
	Secondary school	.803	.391	4.211	1	.040	2.232
	Above secondary school						1
	Vehicle Type			15.270	5	.009	
	Automobile	1.697	.529	10.280	1	.001	5.460
	Taxies and Minibuses	2.050	.601	11.632	1	.001	7.768
	Cargo	1.097	.540	4.128	1	.042	2.996
	Buses	1.014	.617	2.701	1	.100	2.755
	Other vehicles	1.101	.663	2.753	1	.097	3.006
	Station wagon						1
	Vehicle Age			4.343	3	.227	
	Over 10 years	.820	.514	2.545	1	.111	2.270
	2 -5 years	.094	.546	.030	1	.863	1.099
	5 -10 years	.256	.489	.274	1	.601	1.291
	2 years at maximum						1
	Vehicle Direction			3.936	3	.268	
	Straight going	-.865	.759	1.301	1	.254	.421
	Turning	-1.068	.904	1.394	1	.238	.344
	Others	.787	1.168	.454	1	.501	2.197
	Reverse going						1
	Atmospheric Conditions						
	Good condition	2.267	.568	15.939	1	.000	9.652
	Rainy						1
	Light Condition			.274	2	.872	
	Daylight	-.185	.607	.093	1	.760	.831
	Dark but lighted	.006	.657	.000	1	.993	1.006

Dark						1
Day of Accident						
Weekend	-.681	.300	5.157	1	.023	.506
Week day						1
Time of Accident			16.259	3	.001	
Morning	1.888	.502	14.132	1	.000	6.609
Day	1.132	.580	3.811	1	.051	3.101
Evening	1.686	.479	12.407	1	.000	5.397
Night						1
Road Condition						
Dry	-2.322	.749	9.604	1	.002	.098
Wet						1
Pavement Roughness						
Asphalt	1.698	.586	8.387	1	.004	5.464
Non- Asphalt						1
Road Junction						
No Junction	.737	.337	4.787	1	.029	2.089
Junction						1
Road Situation						
Straight-flat	-.087	.555	.025	1	.875	.917
Not straight						1
Road Class			1.755	2	.416	
Roundabout	.701	.557	1.585	1	.208	2.015
Two -way	.684	.542	1.595	1	.207	1.982
One -way						1
Place of Accident			17.373	5	.004	
Office	2.135	.559	14.575	1	.000	8.456
Residential	2.363	.659	12.850	1	.000	10.623
Commercial	1.644	.611	7.229	1	.007	5.176
Religious centers	2.178	.707	9.502	1	.002	8.828
Recreation	1.420	.651	4.754	1	.029	4.139
School						1
Accident Type			11.426	3	.010	
Crashing with pedestrian	1.749	.614	8.119	1	.004	5.748
Crash with another vehicle	.821	.699	1.380	1	.240	2.274
Others	.458	.792	.334	1	.563	1.581
Crashing with objects(ref.)						1
Reason of Accident			5.760	4	.218	
Rash and negligent driving	.449	.596	.567	1	.451	1.567
Wrong overtaking	-.676	.692	.955	1	.328	.509
No priority to pedestrian	-.049	.602	.007	1	.935	.952
Others	-.592	.686	.744	1	.388	.553
No priority to vehicle						1
Sub-city of the Accident			10.962	9	.278	
Arada	-.306	.869	.124	1	.725	.736
Addis ketema	-.177	.904	.038	1	.844	.837
Ledeta	-1.030	.916	1.266	1	.260	.357
Kirkos	.574	.837	.471	1	.493	1.776
Yeka	.037	.809	.002	1	.963	1.038
Bole	-.138	.808	.029	1	.865	.871
Akaki	-.528	.831	.403	1	.526	.590
Nefas silk	-.939	.838	1.256	1	.262	.391
Kolfie	.583	.877	.442	1	.506	1.791
Gulele						1
Constant	-8.321	2.136	15.177	1	.000	.000

Appendix [C]

Parameter estimates of the ordinal logistic model

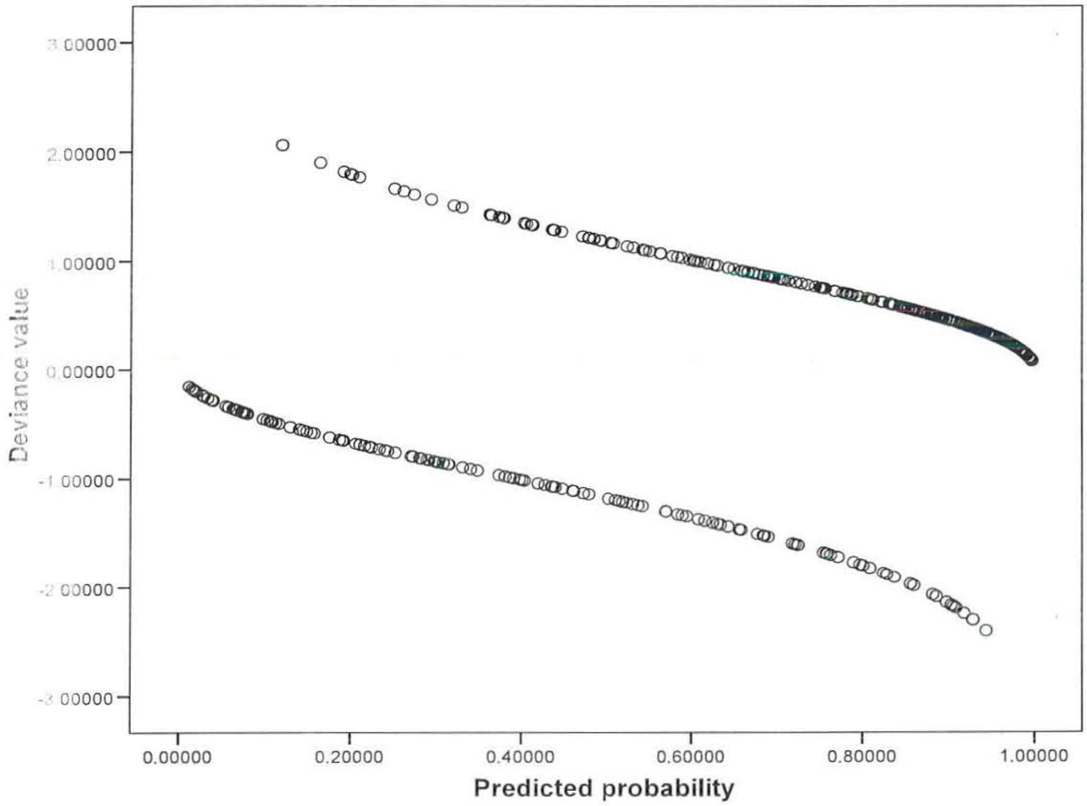
	Estimate	S.E.	Wald	df	Sig.	OR
[SEVERITYLEVEL = 1]	7.932	1.684	22.175	1	.000	
[SEVERITYLEVEL = 2]	11.336	1.740	42.458	1	.000	
Driver Sex	-.430	.521	.682	1	.409	
Male	.703	.611	1.321	1	.250	0.651
Female	0(a)	.	.	0	.	1
Age of Driver	-.723	.160	20.359	1	.000	
18-30	2.224	.444	25.110	1	.000	9.244
31-50	1.437	.412	12.129	1	.000	4.208
>= 51	0(a)	.	.	0	.	1
Driver-Vehicle Relationship	.337	.174	3.745	1	.053	
Other	-.907	.397	5.218	1	.022	0.404
Hired	-.606	.348	3.037	1	.081	0.546
Owner	0(a)	.	.	0	.	1
Driving Experience	.546	.114	22.955	1	.000	
1 years at maximum	-1.817	.443	16.805	1	.000	0.163
1 - 2 years	-1.480	.430	11.832	1	.001	0.228
2 -5 years	-1.004	.293	11.736	1	.001	0.366
More than 5 years	0(a)	.	.	0	.	1
Driver Educational Background	-.658	.121	29.492	1	.000	
Elementary school	2.107	.465	20.569	1	.000	8.224
Junior secondary school	1.851	.433	18.291	1	.000	6.366
Secondary school	1.123	.327	11.792	1	.001	3.074
Above secondary school	0(a)	.	.	0	.	1
Vehicle Type	-.236	.064	13.782	1	.000	
Automobile	1.759	.440	15.950	1	.000	5.807
Taxies and Minibuses	1.493	.474	9.912	1	.002	4.450
Cargo	.984	.445	4.894	1	.027	2.675
Buses	1.462	.509	8.265	1	.004	4.315
Other vehicles	1.242	.544	5.219	1	.022	3.463
Station wagon	0(a)	.	.	0	.	1
Vehicle Age	-.079	.098	.644	1	.422	
Over 10 years	.887	.402	4.873	1	.027	2.428
2 -5 years	-.131	.439	.089	1	.766	0.877
5 -10 years	.758	.392	3.746	1	.053	2.134
2 years at maximum	0(a)	.	.	0	.	1
Vehicle Direction	.247	.137	3.257	1	.071	
Straight going	.220	.594	.137	1	.711	1.246
Turning	1.274	.731	3.036	1	.081	3.575
Others	1.876	.814	5.312	1	.021	6.527

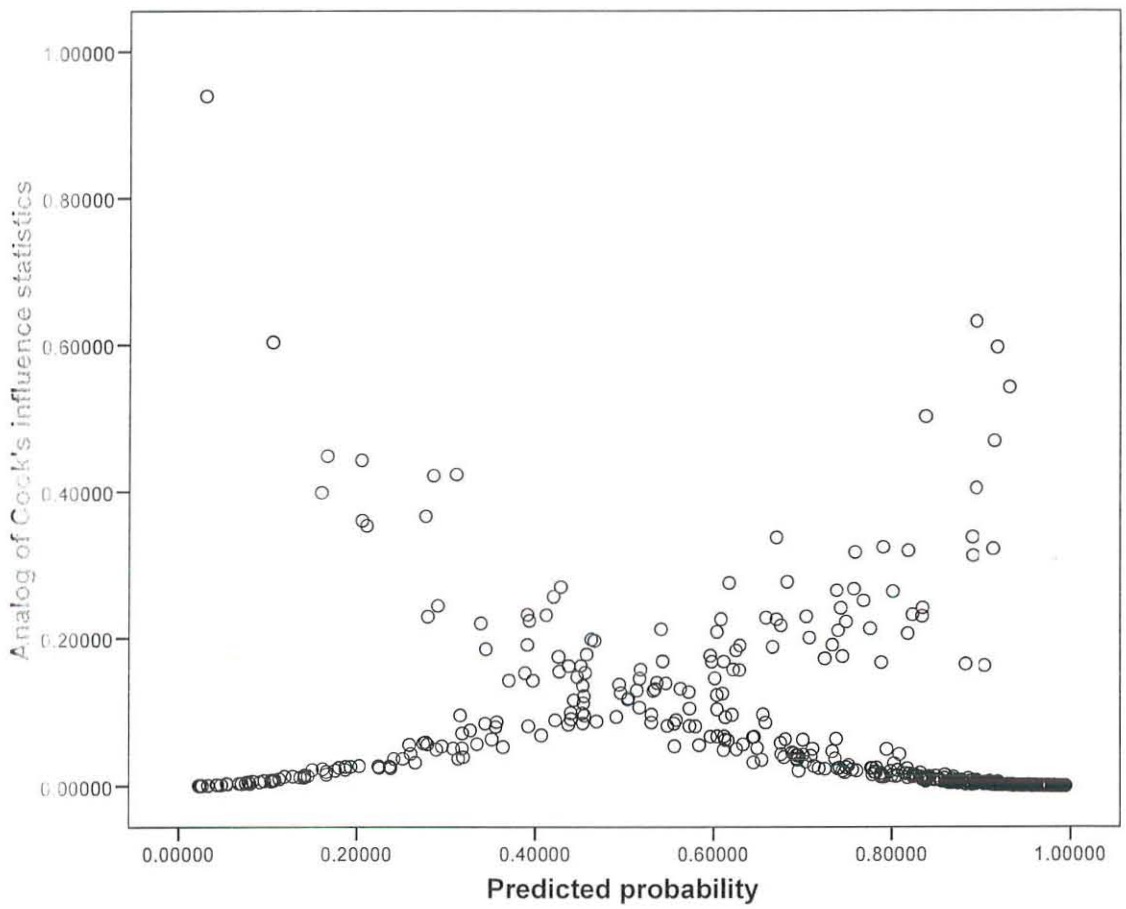
Reverse going	0(a)	.	.	0	.	1
Atmospheric Conditions	-1.529	.379	16.261	1	.000	
Good condition	1.883	.443	18.065	1	.000	6.573
Rainy	0(a)	.	.	0	.	1
Light Condition	1.024	.192	28.529	1	.000	
Daylight	-3.075	.527	33.995	1	.000	0.046
Dark but lighted	-2.555	.541	22.301	1	.000	0.078
Dark	0(a)	.	.	0	.	1
Day of Accident	.388	.214	3.277	1	.070	
Weekend	-.447	.236	3.584	1	.058	0.640
Week day	0(a)	.	.	0	.	1
Time of Accident	-.504	.103	24.014	1	.000	
Morning	2.405	.455	27.936	1	.000	11.078
Day	1.813	.523	12.022	1	.001	6.129
Evening	1.852	.434	18.189	1	.000	6.373
Night	0(a)	.	.	0	.	1
Road Condition	1.671	.440	14.440	1	.000	
Dry	-1.881	.509	13.677	1	.000	0.152
Wet	0(a)	.	.	0	.	1
Pavement Roughness	-1.752	.424	17.091	1	.000	
Asphalt	1.754	.493	12.657	1	.000	5.778
Non- Asphalt	0(a)	.	.	0	.	1
Road Junction	-.464	.232	4.020	1	.045	
No Junction	.777	.274	8.038	1	.005	2.175
Junction	0(a)	.	.	0	.	1
Road Situation	.198	.374	.279	1	.597	
Straight-flat	-.461	.439	1.104	1	.293	0.631
Not straight	0(a)	.	.	0	.	1
Road Class	.048	.164	.084	1	.772	
Roundabout	.312	.446	.488	1	.485	1.366
Two -way	.119	.436	.075	1	.784	1.126
One -way	0(a)	.	.	0	.	1
Place of Accident	-.263	.063	17.663	1	.000	
Office	2.323	.488	22.693	1	.000	10.206
Residential	3.126	.549	32.388	1	.000	22.783
Commercial	2.074	.533	15.160	1	.000	7.957
Religious centers	2.316	.601	14.827	1	.000	10.135
Recreation	1.993	.568	12.301	1	.000	7.338
School	0(a)	.	.	0	.	1
Accident Type	-.804	.133	36.323	1	.000	
Crashing with pedestrian	3.062	.559	29.999	1	.000	21.370
Crash with another vehicle	1.598	.615	6.744	1	.009	4.943
Others	1.346	.713	3.561	1	.059	3.842
Crashing with objects(ref)	0(a)	.	.	0	.	1
Reason of Accident	-.007	.087	.007	1	.935	
Rash and negligent driving	-.699	.453	2.378	1	.123	0.497
Wrong overtaking	-.590	.552	1.143	1	.285	0.554
No priority to pedestrian	-1.180	.458	6.634	1	.010	0.307

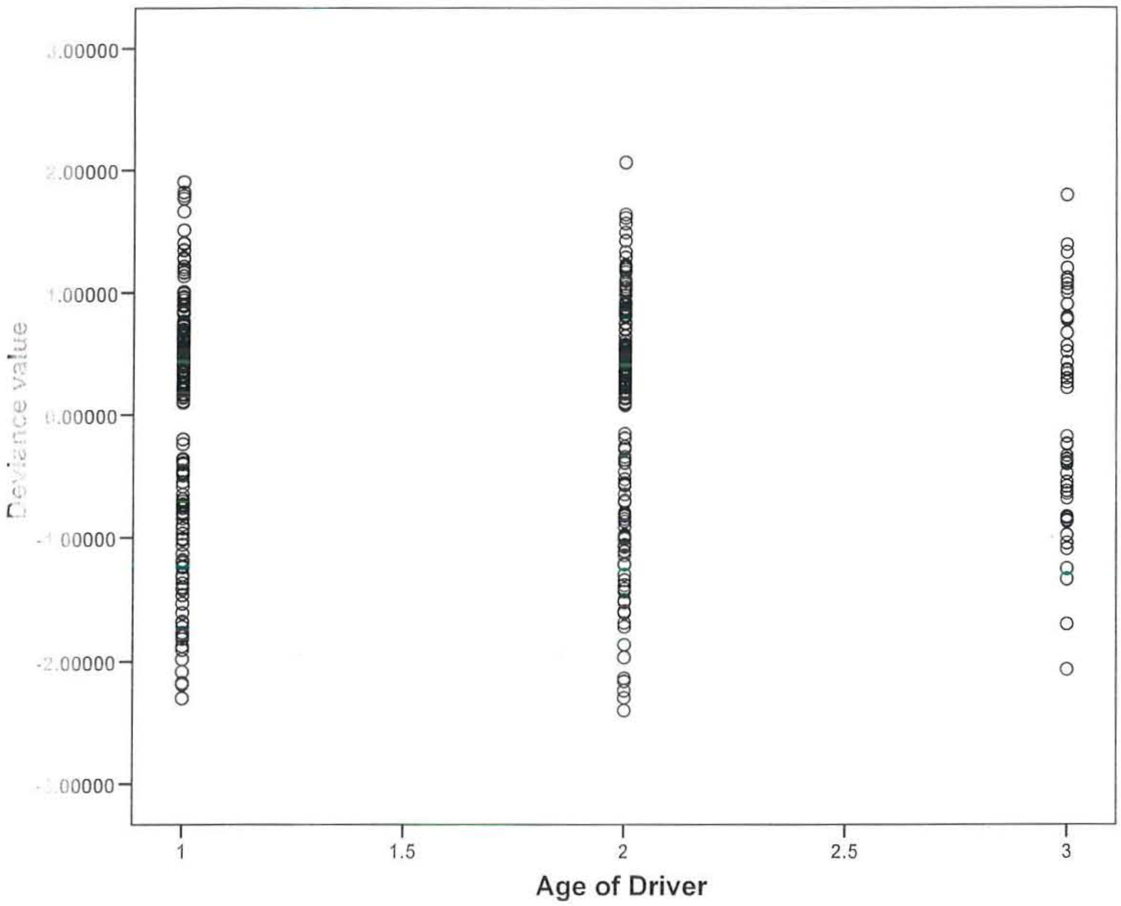
Others	-1.269	.577	4.833	1	.028	0.281
No priority to vehicle	0(a)	.	.	0	.	1
Sub-city of the Accident	-.018	.041	.190	1	.663	
Arada	.534	.679	.619	1	.432	1.706
Addis ketema	-.215	.683	.099	1	.753	0.806
Ledeta	.121	.724	.028	1	.867	1.129
Kirkos	.292	.645	.206	1	.650	1.339
Yeka	.174	.640	.074	1	.785	1.190
Bole	-.241	.635	.144	1	.705	0.786
Akaki	-.103	.650	.025	1	.874	0.902
Nefas silk	-.735	.660	1.240	1	.265	0.480
Kolfie	.562	.669	.705	1	.401	1.754
Gulele	0(a)	.	.	0	.	1

Appendix [D]

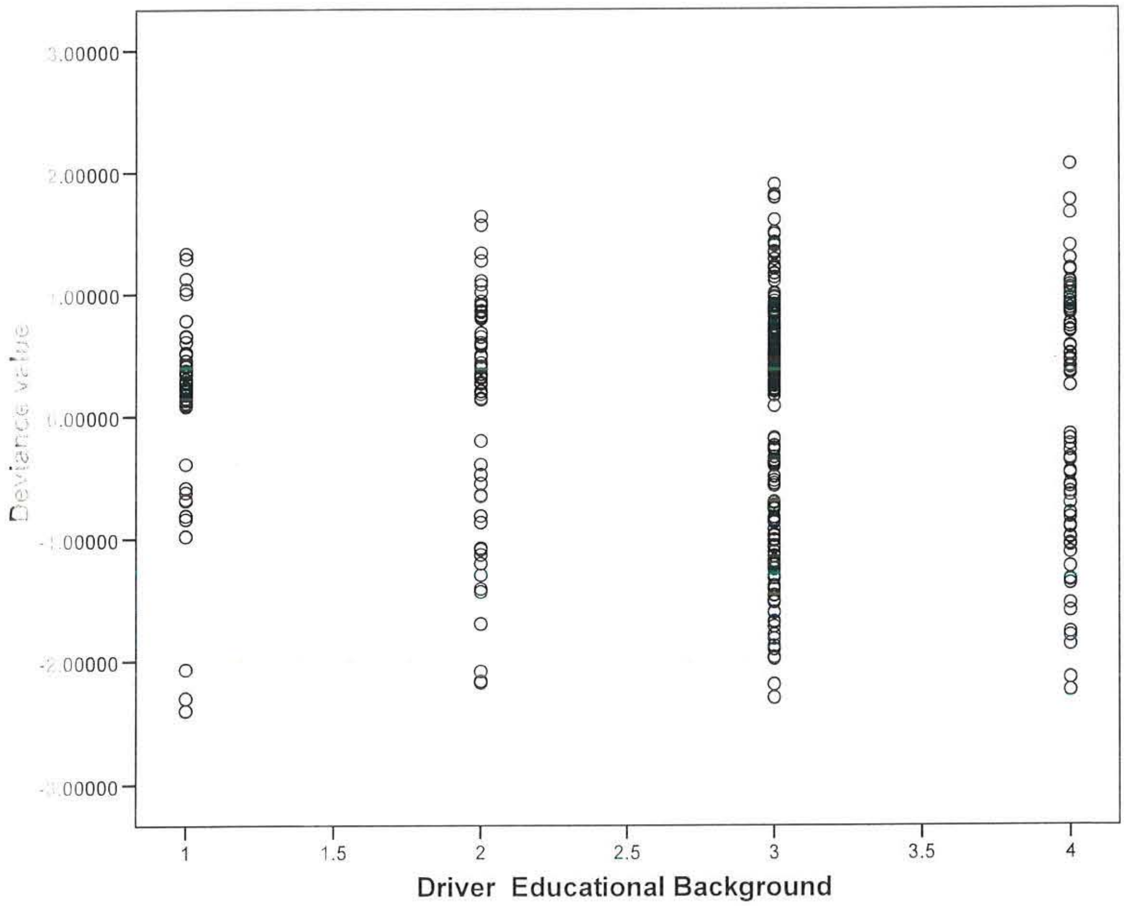
Deviance Value and Cook's Distance

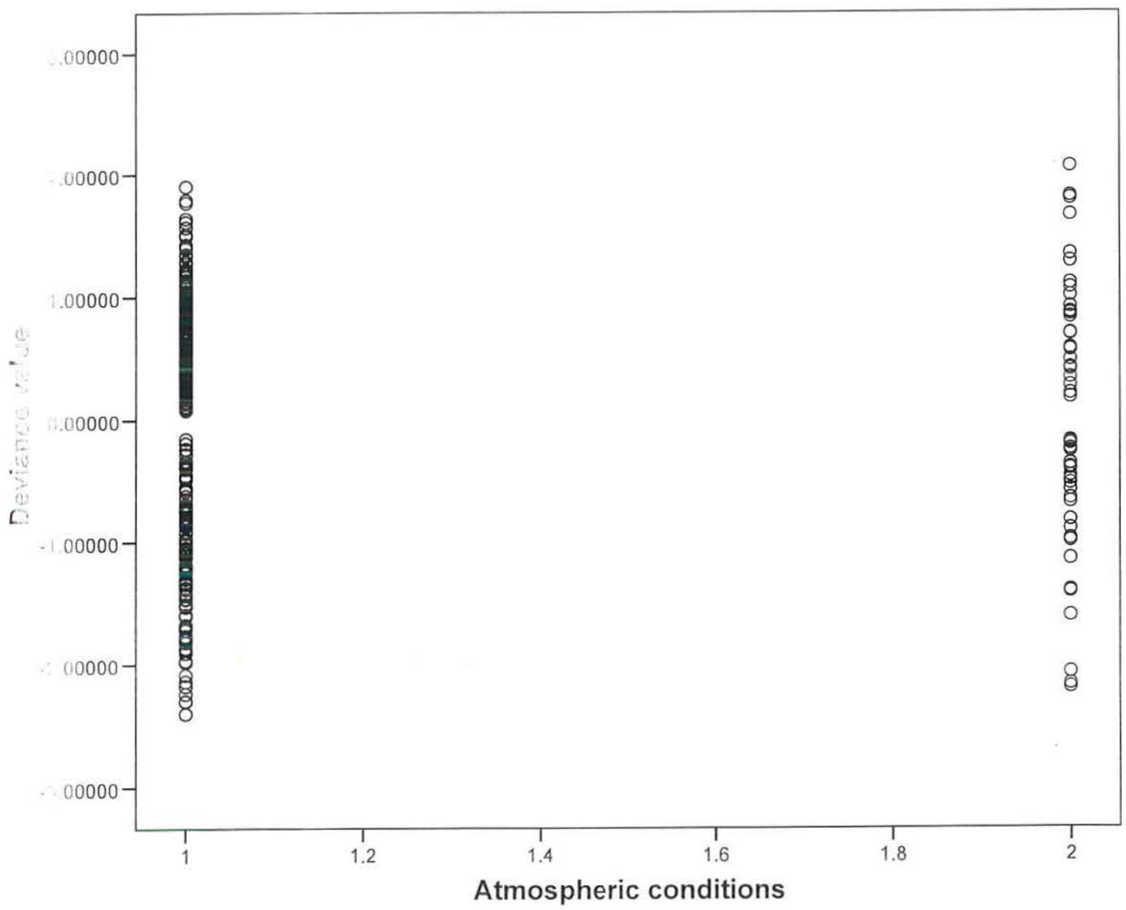


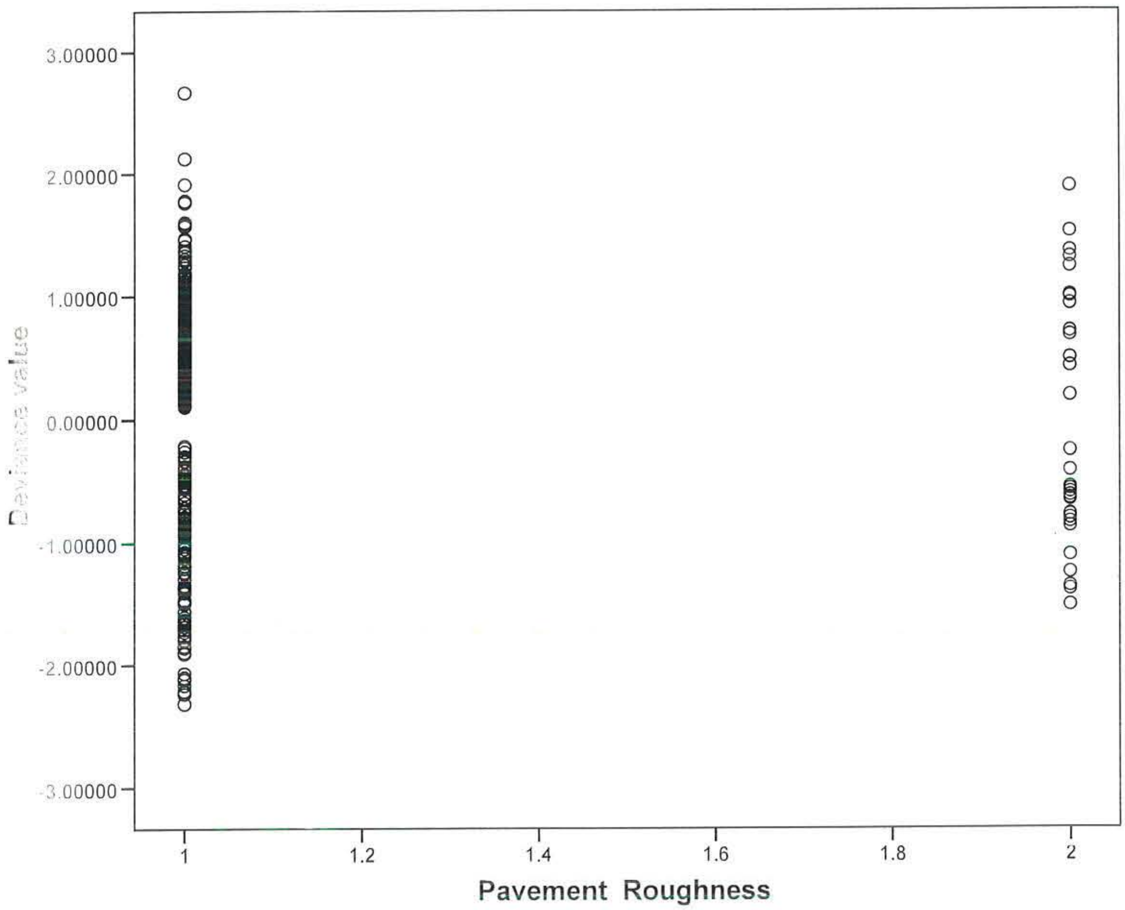

1.00000
0.80000
0.60000
0.40000
0.20000
0.00000
Analog of Cook's influence statistics
0.00000
0.20000
0.40000
0.60000
0.80000
1.00000
Predicted probability



3.00000
2.00000
1.00000
0.00000
-1.00000
-2.00000
-3.00000
Deviance value
1
1.5
2
2.5
3
Age of Driver


3.00000
2.00000
1.00000
0.00000
-1.00000
-2.00000
-3.00000
Deviance value
1
1.5
2
2.5
3
3.5
4
Driver Educational Background


1.00000
1
1.2
1.4
1.6
1.8
2
Atmospheric conditions


3.00000
2.00000
1.00000
0.00000
-1.00000
-2.00000
-3.00000
Deviance value
1
1.2
1.4
1.6
1.8
2
Pavement Roughness

Declaration

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

Declared by

Name: Bisrat Misganaw Geremew

Signature: ------------------------------

Date: June 23, 2010

Confirmed by the Advisor

Name: Dr. Emmanuel G/Yohannes

Signature: ------------------------------

Date: June 23, 2010