Factors Affecting Accident Severity Inside and Outside Urban Areas in Greece

ATHANASIOS THEOFILATOS[a]∗, DANIEL GRAHAM[b], GEORGE YANNIS[c]

[a] Researcher, National Technical University of Athens, Department of Transportation and Engineering, Iroon Polytechneiou 5, 15773, Athens, Greece

[b] Reader, Centre for Transport Studies, Department of Civil and Environmental Engineering, Imperial College London, SW7 2AZ London, UK Room 614, Skempton Bldg. Imperial College London SW7 2AZ

[c] Associate Professor, National Technical University of Athens, Department of Transportation and Engineering, Iroon Polytechneiou 5, 15773, Athens, Greece

Email addresses: atheofil@central.ntua.gr (A. Theofilatos), d.j.graham@imperial.ac.uk (D. Graham), geyannis@central.ntua.gr (G.Yannis)

∗ Corresponding author. Tel.: +302107721515; fax: +302107721454.

E-mail address: atheofil@central.ntua.gr

Abstract

Objectives: This research aims to identify and analyze the factors affecting accident severity through a macroscopic analysis, with focus on the comparison between inside and outside urban areas. Disaggregate road accident data for year 2008 in Greece were used.

Methods: Two models were developed, one for inside and one for outside urban areas. Since the dependent variable had two categories, Killed/Severely Injured (KSI) and Slightly Injured (SI), the binary logistic regression analysis was selected. Furthermore, this research aims to estimate the probability of fatality/severe injury versus slight injury as well as to calculate the odds ratios (relative probabilities) for various road accident configurations. The Hosmer and Lemeshow statistic and other diagnostic tests were conducted in order to assess the goodness-of-fit of the model.

Results: From the application of the models, it appears that inside urban areas three types of collision (sideswipe, rear-end, with fixed object/parked car), as well as involvement of motorcycles, bicycles, buses, two person's age groups (18-30 and older than 60 years old), time of the accident and location of the accident seem to affect accident severity. Outside urban areas, four types of collision (head-on, rear-end, side, sideswipe), weather conditions, time of the accident, one person's age group (older than 60 years old) and involvement of motorcycles and buses were found to be significant.

Conclusions: Factors affecting road accident severity only inside urban areas are young drivers, bicycles, intersections and collision with fixed objects, whereas factors affecting severity only outside urban areas are weather conditions, head-on and side collisions, demonstrating the particular road users and traffic situations which should be focused for road safety interventions for they two different types of network (inside and outside urban areas). The methodology and the results of this research may be proved a promising tool to prioritize programs and measures to improve road safety in Greece and worldwide.

Key words: Accident severity; Logistic regression; Odds and probability ratios; Diagnostic tests

INTRODUCTION

Road safety science deals with a major problem of modern society and has a complex nature, as it involves several scientific disciplines, such as highway and traffic engineers, police officers, town planners, psychologists, emergency and health staff, decision makers etc. According to the European Road Safety Observatory (CARE 2011) road traffic accidents in the member states of the European Union claim about 31.000 lives every year. More than 1.8 million people are injured. The estimated costs are of about 160 billion Euros. During the last decades, a number of measures to reduce road accidents have been taken at a national, regional and local level. The situation in Greece is still very overwhelming. According to the official statistics (Hellenic Statistical Authority, 2010) from 1985 to 2009, 44.327 people were killed, 79.893 people were severely injured and 594.114 people were slightly injured. The percentage of people killed in accidents inside urban areas in Greece (48%) is considerably higher than the respective percentage in the EU (38%) (CARE, 2011).

Various researches which examine road accident severity have been found in the literature review but only one was found to have been carried out in Greece. (Yannis et al., 2005). Many differences were identified mainly in terms of method of analysis, unit of analysis and data sources.

In those researches that attempted to conduct statistical modeling, accident severity was the dependent variable and it usually consisted of two (Al-Ghamdi, 2002; Yamamoto and Shankar, 2004; Yau, 2004; Sze and Wong, 2007; Chimba and Sando, 2009) or more classes (Quddus et al., 2002; Chang and Wang, 2006; Savolainen and Mannering, 2007; Milton et al., 2008) according to the outcome of the accident (for example, only property damage, slight injury, severe injury, death).

The proposed method was mainly defined by the nature of the dependent variable. For example, although, the simple linear regression is a very common method to analyze data and is often applied in the field of transport one very significant assumption of the linear regression is that the response variable is continuous. When the response variable (or the outcome) is discrete, it is inappropriate to use that method. Milton and Mannering (1997) state that: '*The use of linear regression models is inappropriate for making probabilistic statements about the occurrences of vehicle accidents on the road*'. in cases where the discrete response variable can take only two values, the binary logistic regression is appropriate (Hosmer and Lemeshow, 1989). Usually, logistic regression was used when severity has two categories. In those cases where more categories exist other models were used such as the multinomial (Ulfarsson and Mannering, 2004; Kim et al., 2006; Mannering and Islam, 2006; Savolainen and Mannering, 2007), the ordered probit (O' Donnell and Connor, 1996; Quddus et al., 2002; Kockelman and Kweon, 2002; Zajac and Ivan, 2002; Lee and Abdel-Aty, 2005; Christoforou et al., 2010) and the nested logit (Shankar et al., 1996; Chang and Mannering, 1999; Savolainen and Mannering, 2007), etc.

The unit of analysis in accident severity studies was the driver (Kockelman and Kweon, 2002; Abdel-Aty, 2003; Mannering et al., 2005), pedestrians or bicyclists (Sze and Wong, 2007; Lee and Abdel-Aty, 2005), motorcyclists (Pai and Saleh, 2008; Pai, 2009), crashes in general (Thomas and Bradford,1995;

Chang and Wang, 2006; Malyshkina and Mannering, 2009) and only finally fatal crashes (Yannis et al., 2010; Dupont et al., 2010).

It is important that there is a great variation in data sources among previous studies. For example, Williamson et al. (2008) used linked police crash records and hospitalization data from New South Wales, Australia, whilst Mannering et al. (2008) gathered historical accident data for the 1990–1994 timeframe in order to analyse severity in Washington State and Yau (2004) considered 1999–2000 data that were obtained from traffic accident data system (TRADS), which was developed jointly by the Transport Department, Police Force and Information Technology Services Department in the city of Hong Kong.

There are several parameters that were identified to influence severity and they are associated with driver characteristics, type of vehicle, speeding, collision types, road characteristics/geometry and environmental conditions.

Ulfarsson and Mannering (2004) who explored the differences in male and female injury severities in SUV, minivan, pickup and passenger car accidents, suggested that there are essential differences between male and female drivers in terms of behavior and psychology. As a consequence, this fact has to be further explored as well as put into practice when designing vehicles and roads. Valent et al. (2002) applied logistic regression to evaluate the association of driver characteristics and accident severity in the wider area of Udine, Northeast Italy. The results indicated that males are more likely to be engaged in a fatal accident and also that car drivers have lower probability to get killed than motorcyclists and bicyclists.

Mannering et al. (2005) examined the differences in rural and urban driver-injury severities in accidents involving large-trucks, using four years of California accident data and considering four severity categories: no injury, complaint of pain, visible injury and severe/fatal injury. Significant differences between urban and rural models were found to exist with respect to factors such as driver, vehicle and environmental characteristics. Mannering and Lee (2002) analyzed run-off-roadway accidents on a 96.6 km section of a highway in Washington State. They argued that temporal characteristics (e.g. day time, peak hour), environmental characteristics (e.g. type of weather, condition of the road surface), driver characteristics (e.g. consumption of alcohol), roadway and roadside characteristics (e.g. narrow shoulder, miscellaneous fixed objects) had a significant effect on severity.

Kim et al (2006) developed an econometric model in order to estimate the probability of accidents (with a censored regression model) and survival (with a logistic regression model) in the United States. The probability of survival is influenced by the physical characteristics of the vehicles involved in the accident and by the characteristics of the driver and the occupants. Those who used to drive a heavy vehicle had increased survival rate. Chang and Wang (2006) carried out a study that concerned the analysis of traffic injury severity. By analyzing the 2001 accident data for Tapei, they concluded that the most important factor that affected crash severity was the type of the vehicle.

Al-Ghamdi (2002), attempted to estimate the influence of various parameters on accident severity in Riyadh, the capital city of Saudi Arabia. The most significant variables were found to be location (intersection, non-intersection) and cause of accident (speed, run red light, wrong way etc.). In addition, logistic regression was found to be a promising method that can be used in future in order to improve road safety in Riyadh. It is important to highlight that in this study, the age of the driver was not found to be a significant factor. It seems that young as well as old drivers are exposed to the same hazards.

It also noted that the number of fatalities and serious injuries is not only determined by accident severity but also by accident frequency, and consequently factors affecting either/or severity and frequency of accidents should be co-examined. For example, a common factor influencing both severity and frequency of accidents is speeding (Aarts and Van Schagen, 2006), which can explain the accident dynamics both before and during the accident.

This literature reviews demonstrated that research on accident severity is mainly focused on specific areas and specific road users and their characteristics, with not that much effort on a more macroscopic analysis. Furthermore, severity multinomial and ordered regression methodologies are widely used for the analysis of accident severity, without equivalent effort on logistic regression analyses. Consequently, there is need for further investigation of the combined effect of the road type and the road user characteristics on accident severity on a more macroscopic level. Furthermore, there is also need to examine whether logistic regression analysis could allow the identification of factors affecting road accident severity and the quantification of their effects. If the factors affecting accident severity (together with those affecting accident frequency) are well identified, appropriate road safety measures could be designed and implemented in order to decrease the severity of the accidents.

OBJECTIVES AND METHODS

Objectives

The objective of this research is the analysis of accident severity, the identification of the related determinants and the quantification of their effects through a macroscopic analysis, with particular focus on the comparison between inside and outside urban areas, by the use of logistic regression analysis. This macroscopic analysis involved disaggregate data for all road accidents recorded in Greece in 2008. Occupants from all vehicle types were examined (cars, trucks, buses motorcycles, etc.) including both front and back seat occupants. Furthermore, this research aims to estimate the probability of fatality/severe injury versus slight injury as well as to calculate the odds ratios (relative probabilities) for various road accident configurations.

It is suspected that some parameters affect accident severity like the age and category of the occupant, the accident location and area type, the age and type of the vehicle, the weather, the time and the type of the accident. Appropriate logistic regression models are developed and their significance is tested by applying a hypothesis testing technique. The following typical test was used:

Null Hypothesis=H_o: the examined factors are not statistically significant at 0.05 level

Alternative Hypothesis=H_1: the examined factors are statistically significant at 0.05 level

The odds and probability ratios are very important because they allow the identification of critical parameters and enable the authorities to develop priority programs in order to reduce road accidents and casualties involved in them.

Data Collection, Description and Coding

A database with road accident disaggregate data of the Hellenic Statistical Authority (EL.STAT.) was used, which is based on the Police accident recording forms. The dataset consisted of all the accidents that occurred in 2008 in Greece. The year 2008 was chosen basically for two reasons. Firstly, the accident recording methods are getting better every year and thus the probability of errors and omitting information is minimized. Secondly, the great number of people involved in accidents every year in

Greece is sufficient for a thorough statistical analysis and extraction of conclusions. Unfortunately, the reports in accident sites did not describe the injuries in high detail. Consequently, it was not possible for this study to acquire more details on severity (e.g. no visible injury, only property damage etc.). As a result only two outcomes of severity were taken into consideration i) killed or severely injured and ii) slightly injured. The database consisted of 16.426 injured people (3.327 KSI and 13.099 SI).

In order to conduct the analysis 10 variables were selected to be examined and coded. Subsequently, the data were divided in two categories: accidents occurred inside and outside urban areas. The response variable is Variable 1, namely, SEVERITY, which is a binary (or dichotomous) variable in nature. Two severity levels exist: 0: slight injuries and 1: severe injuries and fatalities. In the Greek database, fatalities are defined as all persons killed within 30 days after the accident, serious injuries are all persons hospitalized for at least 24 hours and slight injuries are all other persons reported by the Police as injured without hospitalization.

Table 1 presents a brief description of the variables that were included in the study.

Table 1 to be inserted here

All independent variables are categorical. Since some of the categorical variables have more than two levels, the creation of dummy variables was necessary in order to import the data in the statistical software. One way to code the categorical variables is to have k-1 dummy variables for k levels of that variable.

An example of this method is being given in the next table.

Table 2 to be inserted here

If the weather when an accident happens is good (e.g. clear sky) the design variable D_1 is set to 1 and the other two are set to 0. When the weather conditions are ‘other’ such as snowing, the design variable D_2 takes the value 1 and the other two take the value 0. Finally, when it is rainy, all the design variables are set equal to zero. This coding system was used for the rest of the categorical variables as well in order to apply the logistic regression in the statistical software. It is noted that although Table 1 and 2 show that weather has 3 categories (good, rainy, other), ultimately it was decided that weather should have two categories (good and bad) since the category other represented only a very small percentage and it was unified with rainy category.

After having coded and transformed the data, their import in the statistical software SPSS followed. The original database was split in two different datasets: one for inside urban areas and one for outside urban areas. The name, the values and the type had to be defined for every variable.

RESULTS

Analysis & Results Outside Urban Areas

It is mentioned that the effect of motorways on severity was not examined due to the fact that only a small proportion of accidents occurred on motorways. All variables were checked for potential correlations. It is desirable to have minimum correlation between the dependent variables (predictors). The Spearman's and the Pearson's correlation coefficient were used. The Pearson's correlation coefficient is more appropriate for continuous variables and the Spearman's for discrete variables. The correlation coefficient *r*, shows in which extend and how (positively or negatively) two variables are correlated. Its range is between -1 and 1. Values that are close to 1 express strong positive correlation, whilst values close to -1 express strong negative correlation. If *r* is close to 0, it means that these two variables are independent. In general, values lower than 0.8 or higher than -0.8 are acceptable.

The correlation tests in each database indicated that there were not significant correlations between variables and the statistical analysis could be performed. The method that was used in order to run the logistic regression was the 'Enter' method. Each variable was manually entered in the model and if it was significant it remained. If not it was removed. After all variables were examined, each variable was removed from the final model in order to check the significance in the change.

Firstly, severity outside urban areas was analyzed. All the variables were examined for their significance against zero. The criteria that were followed in order to decide if a variable was to be accepted in the model were the following:

1. If the Wald is higher than approximately 1.7 then the variable is significant (the Sig. value which is the significance is less than 0.05).
2. If a variable is added and the change in -2Loglikelihood is higher than 3.84 then the variable is significant at 95% level of confidence (1 degree of freedom for each variable).
3. If a variable is removed and the change in -2Loglikelihood is higher than 3.84 then the variable is significant at 95% level of confidence.

It is important to note that the likelihood ratio test is generally preferred over its alternative, the Wald statistic test. In general, the likelihood statistic is superior to the Wald statistic (in the sense that it gives more reliable results). The parameters' estimates of the model for outside urban areas are presented in Table 3:

Table 3 to be inserted here

The next step was the goodness-of-fit test of the model. That is how well the model fits the data. The change in deviance will be used again for the assessment. Firstly, the model runs only with a constant. Then the final model runs again with all the significant variables included in that. For 9 degrees of freedom (9 variables in the final model) at 0.05 level, the critical value in the chi-square table is 16.92. The change in -2Loglikelihood is 238.144. That means that the model fits the data well.

The Hosmer & Lemeshow test was also used for the same purpose. The chi-square is 12.255 for 8 degrees of freedom. The Sig. value is higher than 0.140 so the chi-square is not significant. A non significant chi-square indicates that the data fit the model well.

Then, the significant variables could be further tested for potential interactions. The process was to add each interaction term to the full model. If the added term is significant there is interaction between these two variables is significant. The tests had shown that no interactions existed. All interactions fell short of significance (Sig. > 0.05). The best model in its analytical form is the following:

U = **-** 0.592 + 0.210*WEATHER – 0.203*TIME + 0.914*LESS_50CC + 0.243*MORE_50CC – 1.162*SIDESWIPE – 0.958*REAR-END – 0.244*HEAD-ON – 1.166*BUS – 0.747*SIDE + 0.387*OLDER_THAN_SIXTY_YEARS_OLD Eq. (1)

The impact of several predictors on accident severity is further explained: The estimated coefficient of the Weather variable is positive. That means that the probability of a severe/deadly accident rises when the weather is good (as the value 1 is for good weather and the value 0 for bad weather). This was expected because when the weather is good, people usually drive faster. On the other hand, bad weather especially in Greece is something rare, so drivers drive more carefully and slower. The odds ratios that will be calculated are described next:

Logit (fatal accident / good weather) = $\beta_o + \beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 + \beta_6 + \beta_7 + \beta_8 + \beta_9 + \beta_{10}$ Eq. (2)

Logit (fatal accident/ bad weather) = $\beta_o + \beta_2 + \beta_3 + \beta_4 + \beta_5 + \beta_6 + \beta_7 + \beta_8 + \beta_9 + \beta_{10}$ Eq. (3)

Where, β_i = 0, 1…10 are the coefficients of the predictors. Hence the logit difference is the estimated coefficient $\beta_1 = 0.210$ and the odds ratio ψ is $e^{0.210} = 1.234$. This value indicates that the odds of being in a severe or fatal accident in good weather conditions are 1.234 higher than those in bad weather. It is stressed that the odds ratio ψ is different than the probability ratio between KSI and SI.

The probability of being KSI (good weather) is $(e^{-0.592+0.210}/(e^{-0.592+0.210}+1) = 0.406$

The probability of being SI (good weather) is 1 – 0.406 = 0.594

So, the probability ratio is: KSI/SI = 0.406/0.594 = 0.684

The probability of being KSI (bad weather) is $(e^{-0.592}/(e^{-0.592}+1) = 0.356$

The probability of being SI (bad weather) is 1 – 0.356 = 0.644

So, the probability ratio is: KSI/SI = 0.356/0.644 = 0.553

When the variable consists of two levels, the odds that the accident will be severe or fatal accident is the logit difference and it is equal to the exponential of the coefficient of the variable.

The negative coefficient of the Time variable means that those accidents that happen at night are probably more severe than those happen at daylight. It was expected that at night fewer accidents happen because the flows are fewer as well, but they are more severe. The odds of being in a severe/fatal accident during the day are 0.816 times lower than during the night.

As expected drivers and passengers of small motorbikes (< 50cc) are more likely to get injured severely or get killed. Since the vehicle type has more than two levels, Car was used as a reference category. So, the odds ratio is $e^{0.914} = 2.495$. This means that the odds ratio of the accident being severe or fatal for the occupants of small motorbikes is 2.495 times higher than for the occupants of cars.

It was also expected that the occupants of large motorcycles (>50cc) are high likely to be severely injured or get killed. Also, the odds ratio of an accident being severe or fatal for the occupants of large motorbikes is 1.275 times higher than for the occupants of cars.

The collision type consists of several levels. Run off Road was used as a reference category. The negative value of the coefficient of Rear-end predictor shows, that the rear-end collisions decrease the probability of severely/fatal accidents. Also, the odds ratio of an accident being severe/fatal when the collision is rear-ended is 0.384 times lower than an accident where the vehicle runs off-road.

Sideswipe collisions have negative coefficient as expected because sideswipe accidents are not usually severe. The odds ratio of an accident being severe/fatal when the collision is sideswipe is 0.313 times lower than an accident where the vehicle runs off-road.

The Head-on and Side collisions have negative coefficients. The fact that a head on collision decreases the probability of a severe accident may seem counterintuitive. Nevertheless, the technological improvements that have been achieved through all these years in order to improve occupant safety should be taken into account. Other studies also indicate that head-on collisions have a negative coefficient (Al. Ghamdi, 2002). Also, side collisions decrease the probability of severe or fatal accidents, maybe because they tend to happen at lower speeds (e.g. at intersections).

Buses also seem to decrease accident severity and their driver and passengers are less likely to be KSI than the car occupants.

The Age variable consists of 5 groups. The third group (31-46 years old) was used as a reference category. Only the last group (>60 years old) was found to have an influence on severity. The odds ratio is 1.473.

The next Figure (Figure 1) shows the previous odds ratios (ψ).

Figure 1 to be inserted here

In order to find the ratio between KSI/SI, the probability of being KSI is first obtained and then it is possible to obtain the probability of being SI and hence the probability ratio. There are many different values of probability ratios between KSI and SI because there are many variables in the model. The next table shows some interesting probability ratios.

Table 4 to be inserted here

Analysis & Results Inside Urban Areas

The method of analysis was the same as in the previous database. The same variables were examined with the same reference categories and the criteria for the assessment of the model were the same as well. The next table (Table 5) summarizes the parameters' estimates of the final model which describes severity in urban areas.

Table 5 to be inserted here

The difference in the deviance between the initial model (only with the constant) and the final model is 447.452. The change in deviance is significant for 8 degrees of freedom at 0.05 level. As a consequence, the model fits the data well. Although no correlations seemed to exist those variables were further investigated for potential interactions. The same procedure was followed and once more, no interaction terms were found to be significant. The model which describes severity inside urban areas in its analytical form is the following:

U = - 1.786 + 0.539*LESS_50CC + 0.560*MORE_50CC + 0.585*BICYCLE + 0.244*EIGHTEEN_TO_THIRTY_YEARS_OLD + 0.541*OLDER_THAN_SIXTY_YEARS_OLD - 0.323*TIME – 0.574*LOCATION – 0.812*SIDESWIPE + 0.788*FIXED_OBJ. – 0.629*REAR-END - 1.005*BUS Eq. (4)

The coefficient of the Less_50cc predictor is positive. Since the vehicle type variable consists of more than two levels, the predictor Car was used as a reference category. So, the odds ratio is $e^{0.539}$= 1.714. This means that the odds ratio of the accident being severe or fatal for the occupants of light motorbikes is 1.714 times higher than for the occupants of cars.

The coefficient of the predictor More_50cc is also positive. It was also expected that the occupants of large motorcycles are high likely to be severely injured or get killed. Also, the odds ratio of an accident being severe or fatal for the occupants of large motorbikes is 1.751 times higher than for the occupants of cars. These first results are expected since here is strong evidence (Yannis et al., 2009) that helmet use in Greece is low.

The coefficient of the predictor Bicycle is positive as well. It seems that riding a bicycle without taking the appropriate measures (Yannis et al., 2009) and also in a non-friendly environment infrastructure poses hazards to bicycle riders. Bicycle riders have 1.794 more odds to get killed or severely injured than car occupants.

Bus has a negative coefficient in the model (same as outside urban areas) and as a result it decreases accident severity.

The model indicates that accidents at locations outside intersections are 1.776 times more likely to be severe or fatal than those at intersection locations.

The Time predictor has a negative coefficient (same as outside urban areas). The reasons for the increased severity at night that were identified outside urban areas also apply inside urban areas. The odds of being in a severe/fatal accident during the day are 0.724 times lower than during the night.

The Sideswipe predictor has negative coefficient as expected because sideswipe accidents are not usually severe. The odds ratio of an accident being severe or fatal when the collision is sideswipe is 0.444 times lower than the run off road collision (reference category).

The positive coefficient of the Fixed Object/Parked car means that there is an increase in the chance of a KSI outcome. It seems that inside urban areas the presence of roadside features such as utility poles, guardrail systems, fences, trees and parked cars have an important influence on severity. The corresponding odds ratio is 2.199.

The Rear-end predictor was found to be statistically significant. It has a negative coefficient. Thus, it decreases the probability of someone being Killed or Severely Injured. The odds ratio in comparison with a run off road collision is 0.533.

Finally, two Age categories were found to be significant, namely, the 18-30 years old and above 60 years old. Both increase the severity of an accident in relation to the reference group which is the third age group (31-46 years old). More specifically, the odds ratios are 1.277 and 1.717 respectively.

The next figure (Figure 2) shows the calculated odds ratios.

Figure 2 to be inserted here

The next table (Table 6) illustrates the values of various probability ratios between KSI and SI inside urban areas. The method to calculate those probability ratios is the same that was followed in the previous model.

Table 6 to be inserted here

CONCLUSION

The aim of this research is to investigate road accident severity through a macroscopic analysis, with particular focus on the comparison between inside and outside urban areas. Two models were developed by using disaggregate road accident data for 2008 in Greece; the former describes severity outside urban areas and the latter severity inside urban areas. Since the response variable has two categories (Killed/Severely Injured or Slightly Injured) the logistic regression method was selected. The results of the diagnostic tests indicate that the models provide a reasonable statistical fit. The methodology (odds and probability ratios) calculated both types of areas and the results of this research could be used by decision makers in order to develop priority programs to reduce the number of serious accidents.

Ten variables were found to have a significant effect on severity outside urban areas: weather, time, light (<50cc) and heavy motorcycles (>50cc), buses, older road users (>60 years) and four collision types (side, sideswipe, rear-end and head-on). Severity inside urban areas was found to be influenced by the following variables: light and heavy motorcycles, bicycles, buses, time of the accident, location, young (18-30 years) and older (>60 years) road users and three collision types (sideswipe, fixed object, rear-end).

Young road users (18-30 years) were found to be involved in more serious accidents only inside urban areas, possibly because they tend to take more risks during driving (OECD, 2006) and at the same time they lack experience of managing complex situations of urban traffic. Elderly road users (>60years old) were found to be involved in more severe accidents both inside urban areas and outside urban areas, confirming results from other studies (Viano et al., 1990; Robertson and Vanlaar, 2008, Sjogren et al., 1996). An additional explanation could be the fact that elderly people are more likely to die as a result of less severe injuries than younger ones (Cheung and McCartt, 2011). It is also interesting that elderly occupants are associated with more severe accidents inside urban areas having higher odds ratio (1.717 in relation to 1.473 outside urban areas).

It was also found that more severe accidents tend to happen at night. The odds ratio is lower in urban areas (0.724) than outside urban areas (0.816). As a consequence, accidents happening at night inside urban areas are 1.38 times more likely to be severe than those that happen during the day. The respective ratio outside urban areas is 1.22. Aside from alcohol usage, another possible explanation is that people drive faster or carelessly as roads are less congested during the night. In addition, insufficient illumination of roads and junctions - especially outside urban areas - in Greece could be a reason for this increased accident severity. Authorities have to improve the road lighting conditions but also to apply more systematic police enforcement during the night including drinking and driving offenders (Voas et al., 2003).

Intersection has a significant influence in severity only inside urban areas and not outside urban areas. Accidents which occur outside intersections are more severe possibly due to higher speeds. Furthermore, the type of weather seems to affect severity only outside urban areas possibly because more accidents occur on road sections (instead of intersections) outside urban areas, but also because lower speeds during adverse weather conditions inside urban areas may not affect accident severity. The combined effect of traffic volume, speed, weather conditions and road type on road accidents requires further research.

Motorcycles are more likely to be associated with severe/fatal accidents in both areas. Light motorcycles have higher odds ratio outside than inside urban areas (2.495 and 1.714 respectively), whilst heavy motorcycles have higher odds ratio inside urban areas (1.751). It is interesting that both types of motorcycles have the same odds ratio inside urban areas, but outside urban areas accidents with light motorcycles seem to be more severe. The power two wheeler (PTW) safety plans (SPRSMM, 2010) which started to be developed in the recent years, should take into account this finding on PTW accident severity and incorporate measures both for the infrastructure and the riders and drivers behavior.

Bicycles were found to be statistically significant only inside urban areas and are more likely to be involved in more severe accidents. As more people tend to use this mode in the cities, it is suggested that authorities should design and provide a bicyclist-friendly traffic environment and strive for more effective occupant protection systems (Viano, 1988) reducing thus both accident frequency and severity. Buses were found to be statistically significant both inside and outside urban areas. They are involved in less severe accidents than cars which are the reference category, being 3.2 times and 2.7 times less likely to be involved in severe accidents outside and inside urban areas respectively. More specifically the odds ratios are 0.312 outside and 0.366 inside urban areas.

Four types of collision were found statistically significant outside and three inside urban areas. Amongst them, only sideswipe and rear-end collisions are commonly present in both models. Their odds ratios are higher inside urban areas. As a consequence, rear-end and sideswipe collisions seem to be more severe outside urban areas. Another important finding of the research was that collisions with fixed objects/parked cars lead to more severe accidents inside urban areas, perhaps explained by the high frequency of roadside obstacles (utility poles, guardrails, trees, etc.). A systematic implementation programme is needed to identify and treat dangerous roadside obstacles in each case separately.

In conclusion, factors affecting road accident severity only inside urban areas are young drivers, bicycles, intersections and collision with fixed objects, demonstrating that these particular road users and traffic situations should be focused for road safety interventions in urban areas. In parallel, factors affecting severity only outside urban areas are weather conditions, head-on and side collisions, and special road infrastructure and driver behavior measures should be focused for those particular traffic situations.

The limitations in this research, constituting the open issues for further research are also highlighted. Since only accident severity was examined in this research, it would be interesting to co-examine the probability of the occurrence of the crash as indicated in the literature review. Furthermore, the under-reporting level of non fatal injuries should be taken into consideration while reading this study. The results of a study in Greece (Petridou et al., 2009) identifying different under-reporting levels for different road user types and their ages demonstrated that not reported slight injuries are more than double than not reported killed and serious injuries and consequently, severity of accidents involving PTW riders and cyclists, as well as younger road users (found less reported) may be lower than that indicated in this research.. Finally, the effect of some factors such as weather conditions (precipitation and temperature) and occupants' age require further investigation, especially in conjunction with the road type and the related traffic conditions.

REFERENCES

1. Aarts L., Schagen I.V. Driving speed and the risk of road crashes: a review. *Accid Anal Prev*. 2006;38(2):251-224.
2. Abdel-Aty M. Analysis of driver injury severity levels at multiple locations using ordered probit models. *J Saf Res*. 2003;34:597-603.
3. Al-Ghamdi Ali S. 2002. Using logistic regression to estimate the influence of accident factors on accident severity. *Accid Anal Prev*. 2002;34:729-741.
4. Boufous S., Finch C., Hayen A., Williamson A. The impact of environmental, vehicle and driver characteristics on injury severity in older drivers hospitalized as a result of a traffic crash, *J Saf Res*. 2008;39:65–72.
5. CARE: EU road database or national publications. Road accidents statistics in Europe.2010. http://ec.europa.eu/transport/road_safety/index_en.htm. Accessed June 18, 2010. Updated June 18, 2010.
6. Chang L.-Y., Mannering F.L. Analysis of injury severity and vehicle occupancy in truck- and non-truck-involved accidents. *Accid Anal Prev*. 1999;31:579–592.
7. Chang L.-Y., Wang H.W. Analysis of traffic injury severity: An application of non-parametric classification tree techniques. *Accid Anal Prev*. 2006;38:1019-1027.

8. Cheung I., McCartt A.T. Declines in fatal crashes of older drivers: Changes in crash risk and survivability. *Accid Anal Prev*. 2011;43:666-674.
9. Chimba D., Sando T. Neuromorphic prediction of highway injury severity. *Adv Trans Stud*. 2009;19:17-26.
10. Christoforou Z., Karlaftis M., Cohen S. Vehicle occupant injury severity on highways: An empirical investigation. *Accid Anal Prev*. 2010;42(6):1606-1620.
11. Dupont E., Martensen H., Papadimitriou E., Yannis G., Risk and protection factors in fatal accidents. *Accid Anal Prev.* 2010;42:645-653.
12. Hellenic Statistical Authority *(EL.STAT.)*.2010. http://www.statistics.gr/portal/page/portal/ESYE. Accessed June 17, 2010. Updated June 17, 2010.
13. Hosmer D.W., Lemeshow S. A goodness-of-fit test for the multiple logistic regression models. *Com Stat*. 1980;A10:1043–1069.
14. Hosmer D.W., Lemeshow S. *Applied Logistic Regression*. New York: Wiley;1989.
15. Islam S., Mannering F.L. Driver aging and its effect on male and female single-vehicle accident injuries: Some additional evidence. *J Saf Res*. 2006;37(3):267–276.
16. Khorashadi A., Niemeier D., Shankar V., Mannering F.L. Differences in rural and urban driver-injury severities in accidents involving large-trucks: An exploratory analysis. *Accid Anal Prev*. 2005;37(5):910–921.
17. Kim H.S., Kim J.H., Son B. Factors associated with automobile accidents and survival. *Accid Anal Prev.* 2006;38:981–987.
18. Kockelman K.M., Kweon Y.J. Driver injury severity: an application of ordered probit models. *Accid Anal Prev*. 2002;34(3):313-321.
19. Lee C., Abdel-Aty M. Comprehensive analysis of vehicle-pedestrian crashes at intersections in Florida. *Accid Anal Prev*. 2005;37:775-786.
20. Lee J., Mannering F.L. Impact of roadside features on the frequency and severity of run-off-roadway accidents: an empirical analysis. *Accid Anal Prev*. 2002;34(2):149-161.
21. Malyshkina N.V., Mannering F.L., 2009. Markov switching multinomial logit model: An application to accident-injury severities. *Accid Anal Prev*. 2009;41(4):829-838.

22. Milton J.C., Shankar V.N., Mannering F.L. Highway accident severities and the mixed logit model: An exploratory empirical analysis. *Accid Anal Prev*. 2008;40:260-266.

23. O' Donnell C.J., Connor D.H., Predicting the severity of motor vehicle accident injuries using models of ordered multiple choice. *Accid Anal Prev*. 1996;28(6):739-753.

24. OECD. Young Drivers: The roads to safety. *OECD*, Paris, France.2006.

25. Official Statistics. *Annual Statistics for the Period 1971-1997*. Ministry of Interior, Riyadh, Saudi Arabia.1997.

26. Pai C.-W. Motorcyclist injury severity in angle crashes at T-junctions: identifying significant factors and analyzing what made motorcyclists fail to yield to motorcycles. *Saf Sci*. 2009;47:1096-1106.

27. Pai C.-W., Saleh W. Modelling motorcyclist injury severity by various crash types at T-junctions in the UK. *Saf Sci*. 2008;46:1234-1247.

28. Petridou E., Yannis G., Terzidis A., Dessypris N., Germeni E., Evgenikos P., Tselenti N., Chaziris A., Skalkidis I. Linking Emergency Medical Department and Road Traffic Police casualty data: a tool in assessing the burden of injuries in less resourced countries. *Traf Inj Prev*. 2009;10(1): 37-43.

29. Quddus M.A., Noland R.B., Chor C.H. An analysis of motorcycle injury and vehicle damage using ordered probit models. *J Saf Res*. 2002;33:445-462.

30. Robertson R., Vanlaar W. Elderly drivers: Future challenges? *Accid Anal Prev.* 2008;40:1982–1986.

31. Savolainen P., Mannering F. Probabilistic models of motorcyclists' injury severities in single- and multi-vehicles crashes. *Accid Anal Prev*. 2007;39(5):966-963.

32. Shankar V., Mannering F.L., Barfield W. Statistical Analysis of accident severity on rural freeways. *Accid Anal Prev*. 1996;28(3):391-401.

33. Shankar V., Milton J., Mannering F. Modeling Accident Frequencies as zero-altered probability processes: An empirical inquiry. *Accid Anal Prev*. 1997;29(6):829-837.

34. Sjogren H., Bjornstig U., Eriksson A., Ostrom M. Differences between older and younger drivers; characteristics of fatal car crashes and driver injuries. *Saf Sci*. 1996;23(1):63-77.

35. SPRSMM. Strategic Plan for the Road Safety of Motorcycles and Mopeds. General Traffic Directorate (DGT), Madrid, January 2010
36. Sze N.N., Wong S.C. Diagnostic analysis of the logistic model for pedestrian injury severity in traffic crashes. *Accid Anal Prev.* 2007;39:1267-1278.
37. Thomas P., Bradford M. The nature and source of the head injuries sustained by restrained front-seat-car occupants in frontal collisions. *Accid Anal Prev*. 1995; 27(4): 561-570.
38. Ulfarsson G., Mannering F.L. Differences in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents. *Accid Anal Prev*. 2004;36(2):135-147.
39. Valent F., Schiava F., Savonnito C., Gallo T., Brusaferro S., Barbone F. Risk factors for fatal road accidents in Udine, Italy. *Accid Anal Prev*. 2002;34(1):71-84.
40. Viano D. Limits and challenges of crash protection. *Accid Anal Prev.* 1988;20(6):421-429.
41. Viano C.D., Culver C.C., Evans L., Frick M., Scott R. Involvement of older drivers in multivehicle side-impact crashes. *Accid Anal Prev.*1990;22(2):177-188.
42. Voas R.B., Tippetts S., Fell J.C. Assessing the effectiveness of minimum legal drinking age and zero tolerance laws in the United States. *Accid Anal Prev*. 2003;35:579–587.
43. Washington S.P., Karlaftis M.G., Mannering F.L. *Statistical and Econometric Methods for Transportation Data Analysis*. 1st ed. London: Chapman & Hall / CRC Press; 2003.
44. Yamamoto T., Shankar V.N. Bivariate ordered response probit models of driver’s and passenger’s injury severities in collisions with fixed objects. *Accid Anal Prev*. 2004;36:869-876.
45. Yannis G., Golias J., Papadimitriou E. Driver age and vehicle engine size effects on fault and severity in young motorcyclists’ accidents. *Accid Anal Prev*. 2005;37:327-333.
46. Yannis G., Laiou A., Vardaki S., Dragomanovits A., Papadimitriou E., Kanellaidis G. *Seat belt and helmet use in Greece.* In: Proceedings of the 4th Pan-Hellenic Road Safety Conference, Hellenic Institute of Transportation Engineers, National Tehnical University of Athens: November, 2009.
47. Yannis G., Papadimitriou E., Dupont E., Martensen H., Estimation of fatality and injury risk by means of in-depth fatal accident investigation data *Traf Inj Prev.* 2010;11(5):492-502.
48. Yau K.K.W. Risk factors affecting the severity of single vehicle traffic accidents in Hong Kong. *Accid Anal Prev*. 2004;36:333–340.

49. Zajac S., Ivan J.N., 2002. Factors influencing injury severity of motor vehicle-crossing pedestrian crashes in rural Connecticut. *Accid Anal Prev*. 2002;35:369-379.

Tables and Figures

Table 1: Description of the study variables

Number	*Name*	*Values*	*Abbreviation*
1	Severity	0 = SI 1 = KSI	SEVERITY
2	Age of person	1 = less than 18 2 = 18-30 3 = 31-45 4 = 46-60 5 = more than 60	AGE
3	Category	0 = passenger 1 = driver	OCCUPANT
4	Location	0 = Non-intersection 1 = Intersection	LOCATION
5	Type of vehicle	1 = car 2 = bicycle 3 = motorcycle less than 50cc 4 = motorcycle more than 50cc 5 = truck 6 = bus 7 = other	VEHICLE TYPE
6	Age of vehicle	1 = less than 5 years 2 = six to ten years 3 = more than ten years	VEHICLE AGE
7	Type of area	0 = Built up 1 = Non-built up	AREA TYPE
8	Weather	1 = Good 2 = Rainy 3 = Other	WEATHER
9	Time of the accident	0 = night/dusk 1 = day	TIME
10	Accident type	1 = Run off road 2 = Pedestrian 3 = Fixed object/Parked car 4 = Head-on collision 5 = Sideswipe 6 = Side 7 = other 8 = Rear end	ACCIDENT TYPE

Table 2: Example of coding a variable with more than 2 levels

WEATHER	Design variable		
	D1	D2	D3
Good	1	0	0
Rainy	0	0	0
Other	0	1	0

Table 3: Parameters' estimates of the model outside urban areas

Predictor	**Estimated coefficient**	**Standard Error**	**Sig.**	**Odds Ratio**	**95% C.I.**	
					Lower	**Upper**
Weather	0.210	0.087	0.016	1.234	1.039	1.464
Time	-0.203	0.062	0.001	0.816	0.722	0.922
Less_50cc	0.914	0.210	0.000	2.495	1.653	3.765
More_50cc	0.243	0.078	0.002	1.275	1.093	1.486
Sideswipe	-1.162	0.160	0.000	0.313	0.229	0.428
Rear-end	-0.958	0.113	0.000	0.384	0.307	0.479
Head-on	-0.244	0.090	0.007	0.784	0.658	0.934
Bus	-1.166	0.615	0.058	0.312	0.093	1.040
Side	-0.747	0.076	0.000	0.474	0.408	0.550
Older_than_sixty_years_old	0.387	0.143	0.007	1.473	1.112	1.950
Constant	-0.592	0.093	0.000	0.553		

Table 4: Probability ratios between killed/severely injured and slightly injured, outside urban areas

KSI/SI	Good weather	Bad weather
Bus,Head-on	*0,167*	*0,135*
Bus,Side	*0,101*	*0,083*
Bus,Rear-end,Day	*0,082*	*0,066*
Bus,Sideswipe,Day	*0,054*	*0,044*
Less_50cc,Night	*1,702*	*1,389*
More_50cc,Night	*0,869*	*0,706*
Day	*0,558*	*0,451*
Night	*0,684*	*0,553*
Bus,Night	*0,213*	*0,172*
More_50cc,Head-on	*0,682*	*0,553*
Less_50cc,Side,Day	*0,658*	*0,534*
More_50cc,Day	*0,710*	*0,575*
Less_50cc,Day	*1,389*	*1,126*

Table 5: Parameters' estimates of the model for urban areas

Predictor	Estimated Coefficient	Standard Error	Sig.	Odds ratio	95% C.I.	
					Lower	Upper
Less_50cc	0.539	0.141	0.000	1.714	1.301	2.259
More_50cc	0.560	0.067	0.000	1.751	1.534	1.998
Bicycle	0.585	0.229	0.011	1.794	1.146	2.810
Eighteen_to_thirty_years_old	0.244	0.065	0.000	1.277	1.124	1.451
Older_than_sixty_years_old	0.541	0.109	0.000	1.717	1.388	2.125
Time	-0.323	0.063	0.000	0.724	0.641	0.819
Location	-0.574	0.065	0.000	0.563	0.495	0.640
Sideswipe	-0.812	0.122	0.000	0.444	0.349	0.564
Fixed_object/Parked car	0.788	0.083	0.000	2.199	1.870	2.587
Rear-end	-0.629	0.117	0.000	0.533	0.424	0.670
Bus	-1.005	0.425	0.018	0.366	0.159	0.843
Constant	-1.786	0.076	0.000	0.168		

Table 6: Probability ratios between killed/severely injured and slightly injured inside urban areas

KSI/SI	Non-intersection	Intersection
Less_50cc,Night,Eighteen_to_thirty_years_old	*0.367*	*0.207*
More_50cc,Night,Older_than_sixty_years_old	*0.504*	*0.284*
Bicycle,Night	*0.301*	*0.169*
Sideswipe,Day	*0.054*	*0.030*
Fixed_object	*0.369*	*0.208*
Rear-end	*0.089*	*0.050*
Less_50cc,Day,Older_than_sixty_years_old	*0.357*	*0.201*
Bicycle,Day	*0.218*	*0.123*
More_50cc,Day	*0.212*	*0.120*
Bicycle, Fixed_object, Night,Older_than_sixty_years_old	*1.135*	*0.640*
Bus,Day,Rear-end	*0.024*	*0.013*
Bus,Night,Fixed_object, Eighteen_to_thirty_years_old	*0.172*	*0.097*

Figure 1: Odds ratios (ψ) of the various significant predictors outside urban areas

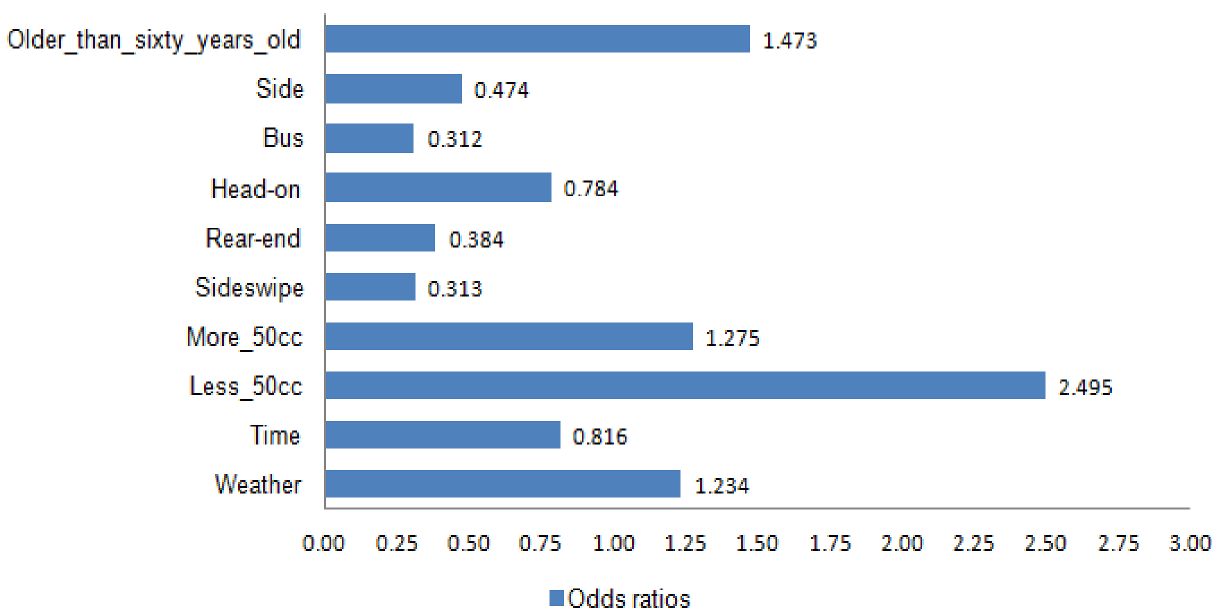

Figure 2: Odds ratios (ψ) of predictors inside urban areas