Quantifying the Influence of Social Characteristics on Accident and Injuries Risk: A Comparative Study Between Motorcyclists and Car Drivers

Lyckegaard, Allan; Olesen, Morten N.; Hels, Tove

Published in:
Young Researchers Seminar

Publication date:
2011

Document Version
Publisher's PDF, also known as Version of record

QUANTIFYING THE INFLUENCE OF SOCIAL CHARACTERISTICS ON ACCIDENT AND INJURIES RISK: A COMPARATIVE STUDY BETWEEN MOTORCYCLISTS AND CAR DRIVERS

Allan Lyckegaard, DTU Transport, Kgs. Lyngby, Denmark

ally@transport.dtu.dk

Morten N. Olesen, Metroselskabet I/S, København S, Denmark

mnolesen@gmail.com

Tove Hels, DTU Transport, Kgs. Lyngby, Denmark

ths@transport.dtu.dk

ABSTRACT

In the recent years many European countries have experienced an increase in the number of fatal traffic accidents with motorcycles. Bos *et al.* (2008) reports an increase from 17.4% to 21.1% of the total number of fatalities on powered two-wheelers in the European traffic. Several reasons for this have been suggested, among the most common is the hypothesis that during the last decade or so, the typical motorcyclist has become older, and as a result of the increase in age, the loss of physical ability in driving and orientation has resulted in the increase in the number of accidents (Værø 2008, SafetyNet 2009). In Denmark in the period 2002 to 2007, the average age of motorcycle owners increased from 42.3 to 45.2 years. In the same period, the average age of injured motorcyclists increased from 35.9 to 38.4 years, meaning that the average injured motorcyclist has become younger in this period.

In this analysis we establish relationships between social and demographic characteristics and the probability of being in an accident and being injured in an accident. Logistic regression was applied to both motorcyclists and car drivers with the purpose of calculating the odds ratio with the car drivers as the control group. The available data for the regression consisted of accident and injury data for motorcyclists and car owners in the period of interest as well

the social and demographic parameters: age, gender, income, educational level and family status. The odds ratio calculations showed that the risk of being in an accident or in an injury accident decreased with age, educational level, and income. Furthermore, the risk of being in an accident was 1.72 to 1.96 times higher and the risk of being in an injury accident was 1.38 to 1.44 times higher for men compared to women. For motorcyclists compared to car drivers, the risk of being in an accident was 1.44 to 1.78 times higher and the risk of being in an injury accident was 2.29 to 3.16 times higher. Singles showed an increased risk of 1.25 to 1.87 times higher for being in an accident and 1.50 to 2.25 times higher risk for being in an injury accident when comparing to person a couple with children.

INTRODUCTION

In Denmark, as well as in the rest of Europe, the period 2002 to 2007 had a high economical growth enabling more people to buy a motorcycle. The common observation expressed in the media was that these motorcycles were bought as a leisure time toy and not with purpose of transportation though there is no evidence that this should be the case. An analysis of the net growth between 2002 and 2007 showed that the largest growth happened in the group of people with an income in the highest 40% percentile. A further analysis of the age distributions revealed a shift in these distributions towards older owners. Figure 1 and 2 shows the age distribution of motorcycle owners and owners of both motorcycle and car, respectively.

Figure 1. Age distribution of owners of motorcycles in 2002 and 2007.

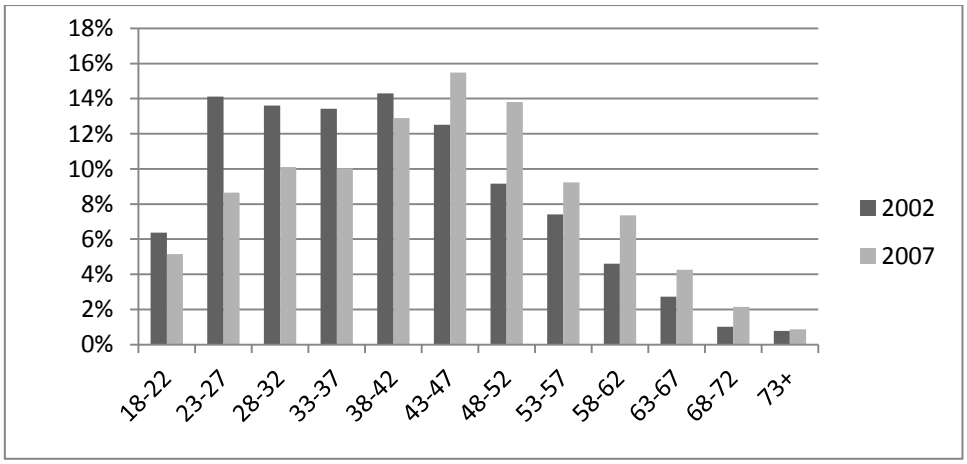
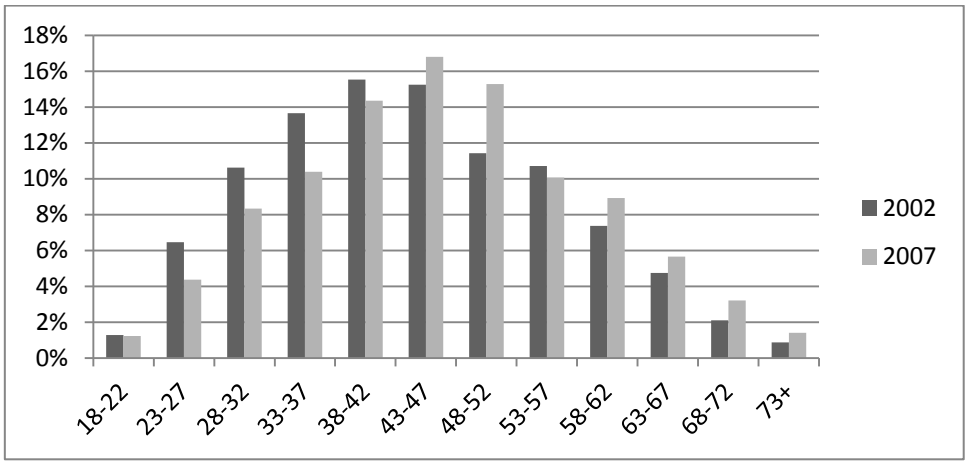


Figure 2. Age distribution of owners of motorcycles and cars in 2002 and 2007.

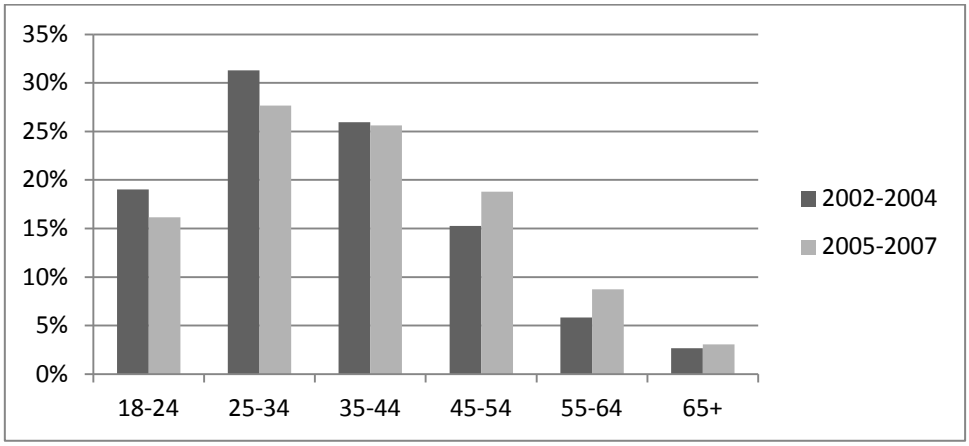


The analysis of the income and age distributions suggests that there has been a change in the social and demographic characteristics of the motorcycle owner in Denmark from 2002 to 2007.

In the same period of time, the number of killed and seriously injured motorcyclists has increased. The fact that more motorcyclists are getting killed and seriously injured is opposite the trend seen in Danish traffic in total. Since the mid 1970's, there has been a downward trend in the number of killed and seriously injured persons in traffic in Denmark, but the trend for the group of motorcyclists is clearly opposite in the period from 2002 to 2007. Up until 2005 the number of killed motorcyclists has been stabile around 5-6% but in 2006 the number was more than 7% and in 2007 the number was almost 9%. These numbers should be compared to the estimated 2% of the traffic that motorcycles and mopeds represent (ETSC 2007). In the rest of Europe, the same pattern with more people buying a motorcycle and more people getting in accidents with motorcycles can be seen. From 2000 to 2005 the number of fatalities on powered two-wheelers in the EU-15 countries went from 3739 to 4047 corresponding to 17.4 % and 21.1 % of the total number of people killed in traffic (Bos *et al.* 2008).

In Denmark, an analysis of the numbers given in DRTAIB (2009), shows that the age of the motorcyclists increased from 42.3 to 45.2 years in the period 2002 to 2007 (N_{2002}=19548 and N_{2007}=38266). In the same period, the average age of injured motorcyclists increased from 35.9 to 38.4 years (N_{2002}=222 and N_{2007}=272), meaning that the average injured motorcyclist has become younger when comparing 2007 to 2002. Figure 3 shows the age distribution of two three-year periods, 2002-2004 and 2005-2007, for killed and seriously injured motorcyclists. The distribution shift towards the older segments when comparing the period 2002-2004 with 2005-2007.

Figure 3. Age distribution of killed and seriously injured motorcyclists in the two periods 2002-2004 and 2005-2007.



The Danish Road Traffic Accident Investigation Board made an in-depth analysis of 41 motorcycle accidents (Værø 2008). The analysis showed that the main causes of accidents were speeding and willingness to take risks, i.e. factors which relates to the behaviour of the motorcyclist. As only 41 accidents were analysed and they only cover the year 2008, it is not possible to draw strong statistical conclusions or say something about development over time.

This paper aims at building on top of the in-depth analysis of the 41 accidents and the simple analysis of motorcycle accident data performed by the Danish Road Traffic Accident Investigation Board (Værø 2008, DRTAIB 2009) by focusing on a quantitative analysis of a large number of accidents with data on individual level. The analysis is an epidemiologic analysis of which factors contribute to increased risk based on socio-demographic data on all motorcyclists and car drivers in Denmark in the period 2002 to 2007.

METHOD

The overall goal with the analysis is to compute the odds ratio for being in an accident or in an accident with personal injury (fatalities are considered as injuries in this study) based on a logistic regression of a number of socio-demographic factors. The design of the study was made as a case/control study with three groups.

Sample

The data used in the analysis comes from Statistics Denmark's collection of register data and are linked to individuals such that various registers of interest can be merged. For this

analysis, all owners of a car or a motorcycle were extracted from the central vehicle register and split into three subgroups: Car, motorcycle (MC) or car/motorcycle owner, see table 1.

Table 1. Number of observations in the individual groups used for the analysis.

Ownership	2002	2003	2004	2005	2006	2007
Car	1.076.127	1.566.716	1.577.078	1.602.615	1.630.135	1.652.454
MC	19.548	27.446	29.032	31.070	34.449	38.266
Car & MC	31.698	48.219	52.536	59.204	67.198	75.807

This study is designed as a case/control study. The car owners are included in the analysis as a control group since this group of road users share the same space on the road as motorcycles and are subject to almost the same type of legislation. The sample is almost the complete population of motorcycle and car owners in Denmark except for 2002 where almost 500.000 vehicles were not registered correctly and therefore missing in the data. Chi-square tests on gender, income and family type showed that there was a small but significant difference (95%) from year to year for all years. For education, a T-test showed a significant difference (99%) in the length of education when comparing 2002 and 2007. Due to the large number of observations even small differences of up to 0.1 year in education length was significant in the T-test. Based on these tests it was decided that the observations from 2002 should be included in the analysis as 2002 was no more different from the other years than any other year.

For each of the individuals included in one of the three groups (Car, MC, Car & MC), gender (male/female), age(in years), income (in DKK before taxation), education length (in years), family relationship (in a couple with children/in a couple without children/single with children/single without children) and presence in the police accidents register is extracted. In the police accident register, accidents with material damage, personal injury and fatal accidents are registered making it possible to distinguish between any type of accident and accidents involving personal injury. It is known that the accident data is suffering from underreporting (Elvik and Mysen 1999, Amoros *et al.* 2006), but as almost all severe accidents are recorded in the police accident register and there is no systematic difference in the underreporting between case and control groups, the data quality is good enough to use in the study of risk. Furthermore, persons involved in an accident as a passenger on either a motorcycle or in a car are left out as they have no or very little influence on how the car or motorcycle is controlled so it is not relevant to include them in the analysis of risk.

Data analysis

The basis of the risk analysis is the logistic regression which is known from many epidemiologic studies. The purpose is to determine a person's probability p_i of being in an accident or an accident with personal injury given a number of socio-demographic factors $\beta_{vehicle}$, β_{gender}, β_{age}, $\beta_{education}$, β_{income} and β_{family} for each individual. The relationship between the probability p_i and the socio-demographic factors β is modelled by logistic regression (Madsen and Thyregod 2010):

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_{vehicle}x_{vehicle} + \beta_{gender}x_{gender} + \beta_{age}x_{age} + \cdots + \beta_{family}x_{family} \quad (1)$$

With this model it is assumed that the underlying unobserved behaviour is equally distributed over both the categorical and the continuous variables.

The odds ratio is an approximation to relative risk when the event under investigation is a rare event. As accidents in general are rare, the assumption that the odds ratio approximates the relative risk is valid. When the odds ratio is estimated from (1), it is possible to separate out how each factor influence the odds ratio with respect to a reference group, as the individual contribution of the β-factors is linear and without interactions between the factors. The chosen reference group will be incorporated into the intercept β_0.

The data analysis and model building was performed in SAS 9.1 using PROC LOGISTIC to compute the parameters of the logistic regression and the corresponding odds ratios.

RESULTS

The results of the logistic regression for both data on all accidents and data on accidents with personal injury are given in this section.

Odds ratio for all accidents

In table 2, the parameter estimates of the logistic regression over all accidents are given. As the data contains categorical variables it is necessary to define a reference group. The reference group (female car owners in a couple with children) is incorporated into the intercept and forms the baseline which the other groups are compared to. The estimated parameters are therefore negative when a group has lower probability compared to the reference group and positive parameter estimates when a group has higher probability compared to the reference group.

Table 2. Parameter estimates for the logistic regression for all accidents. Parameter estimates marked with * is estimated to level of significance lower than 95%. Cells marked with – could not be estimated.

	2002	2003	2004	2005	2006	2007
Intercept	-3,95	-3,91	-4,04	-4,20	-4,09	-3,71
MC	0,19	0,24	0,12*	0,16*	0,12*	0,24
Car&MC	0,10*	0,09*	0,13	0,17	0,17	0,04*
Gender (male)	0,32	0,33	0,28	0,27	0,34	0,33
Age	-0,02	-0,03	-0,03	0,03	-0,03	-0,03
Single w.o. children	0,21	0,11	0,16	0,17	0,15	0,07
Single w. children	0,16	0,36	0,31	0,28	0,22	0,27
In a couple w.o. children	-0,19	-0,21	-0,24	-0,19	-0,17	-0,20
Income, 10^{-7}	-4,57	-2,49	-3,27	-	-2,65	-
Education	-0,05	-0,05	-0,05	-0,06	-0,07	-0,09
R^2	0,0008	0,0007	0,0006	0,0005	0,0006	0,0006

From table 2 it can be seen that the accident probability for both MC and Car&MC is higher than the reference group Car as the parameters estimates are greater than zero. Not all the estimates are significant to a level of 95% significance but the overall trend is that groups MC and Car&MC have a higher probability compared to the group Car. Furthermore, the factors for males and singles increase the probability of being in an accident and the factors age, in a couple with children, income and education decrease the probability of being in an accident. These results do not contradict what has been seen before.

The R^2 is very small so the model only explains very little of a person's accident probability which is what could be expected as the socio-demographic factors only to some degree explain something about a person's behaviour in traffic or level of aggression. The goal is to find a relationship between risk and demographic factors, not building an accident model, so the small R^2 only show that there are more factors which also influence this risk.

The odds ratios for all accidents are given in table 3. For the categorical variables, the groups are compared to the reference group, and for the continuous variables age, education and income, the slope of the linear relationship is evaluated by comparing two points on the line, e.g. Age and Age+1. The conclusion is similar to the one based on table 2, namely, that the groups MC and Car&Mc have a higher risk of being in an accident. Males have almost twice the risk compared to women and singles have a higher risk than persons in a couple. Higher

age and longer education is related to a lower risk of being in an accident, whereas higher income does not seem to play any role.

Table 3. Odds ratio estimates for all accidents based on logistic regression. Cells marked with – could not be estimated.

	2002	2003	2004	2005	2006	2007
MC vs. Car	1,61	1,78	1,44	1,64	1,52	1,68
Car&MC vs. Car	1,47	1,53	1,45	1,66	1,59	1,37
Male vs. Female	1,89	1,92	1,75	1,72	1,96	1,93
Age vs. Age+1	0,98	0,97	0,97	0,98	0,98	0,97
Single w.o. children vs. In a couple w. children	1,47	1,46	1,49	1,53	1,42	1,25
Single w. children vs. In a couple w. children	1,40	1,87	1,73	1,71	1,52	1,52
In a couple w.o. children vs. In a couple w. children	0,99	1,06	1,00	1,07	1,03	0,95
Income vs. income+1000	1,00	1,00	1,00	-	1,00	-
Education vs. education+1	0,95	0,95	0,95	0,94	-	0,91

The odds ratio calculations show that the groups MC and Car&MC, i.e. individuals containing at least one motorcycle, have a higher risk of being in an accident than the control group Car. Likewise, young persons and males have a higher risk compared to older and women which is consistent with what is reported in Brems and Munch (2008). Furthermore, singles show to have higher risk and the more education a person have, the lower the risk.

Odds ratio for accidents with personal injury

In table 4, the parameters estimates of the logistic regression over accidents with personal injury are given. As in the analysis above, the reference group incorporated into the intercept was female car owners in a couple with children.

Table 4. Parameter estimates for the logistic regression for accidents with personal injury. Parameter estimates marked with * is estimated to level of significance lower than 95%. Cells marked with – could not be estimated.

	2002	2003	2004	2005	2006	2007
Intercept	-4,14	-4,10	-4,11	-4,20	-4,25	-3,92
MC	0,38	0,44	0,32	0,35	0,43	0,47
Car&MC	0,29	0,28	0,20	0,31	0,25	0,18
Gender (male)	0,16	0,18	0,18	0,16	0,16	0,17
Age	-0,03	-0,04	-0,04	-0,04	-0,03	-0,04
Single w.o. children	0,25	0,22	0,20	0,26	0,20	0,19
Single w. children	0,20	0,41	0,33	0,32	0,19*	0,33
In a couple w.o. children	-0,19	-0,23	-0,26	-0,21	-0,17	-0,28
Income, 10^{-6}	-1,04	-0,52	-0,62	-,062	-0,35	-
Education	-0,05	-0,05	-0,07	-0,07	-0,07	-0,10
R^2	0,0004	0,0005	0,0004	0,0004	0,0003	0,0004

Again it is clear that the groups MC and Car&MC have a higher probability of being in an accident with personal injury. Similarly, males and singles have a higher probability compared to the reference group and age, income and education length decrease the probability of being in an accident with personal injury.

Table 5. Odds ratio estimates for accidents with personal injury. Cells marked with – could not be estimated.

	2002	2003	2004	2005	2006	2007
MC vs. Car	2,86	3,16	2,29	2,75	3,05	3,09
Car&MC vs. Car	2,60	2,71	2,04	2,65	2,56	2,30
Male vs. Female	1,39	1,44	1,43	1,38	1,37	1,41
Age vs. Age+1	0,97	0,97	0,96	0,97	0,97	0,96
Single w.o. children vs. In a couple w. children	1,67	1,86	1,61	1,86	1,53	1,53
Single w. children vs. In a couple w. children	1,59	2,25	1,82	1,98	1,50	1,76
In a couple w.o. children vs. In a couple w. children	1,08	1,19	1,01	1,17	1,05	0,96
Income vs. income+1000	1,00	1,00	1,00	1,00	1,00	-
Education vs. education+1	0,95	0,95	0,94	0,94	0,93	0,90

Table 5 shows the odds ratio for being in an accident with personal injury compared to the reference group. The pattern in the result is the same as for all accidents, i.e. the groups MC and Car&MC have a higher risk than the group Car. Singles and males again show higher risk, but where a person in a couple without children had a lower risk before, such a person now have a higher risk compared to the reference group. Length of education and age again decrease the risk and income still does not play any role.

Compared to the results for all accidents, it is worth noticing that the risk for motorcyclists being in an accident with personal injury is higher compared to motorcyclist being in any kind of accident. This indicates that when a motorcyclist in involved in an accident, then there is a higher risk for the motorcyclist to be injured compared to a person driving in a car. Similar results have been found in SWOV (2009).

DISCUSSION AND CONCLUSION

It is evident that the risk of motorcyclists being involved in an accident and getting injured in an accident is higher compared to car drivers. Furthermore, it was observed that there in the period 2002 to 2007 were no change in the relation between age and risk of being involved in an accident or the risk of being in an accident with personal injury. This means, that the increase in the number of accidents with motorcyclists in the period 2002 to 2007 could not be ascribed to older motorcyclists. In the whole period from 2002 to 2007 the motorcyclist with the highest risk of being involved in an accident and in an accident with personal injury is a young, single male without children, low income and only a short education.

Even though the models constructed above were significant, it should be noted that they only explain a very small amount of the variance seen in the data, i.e. R^2 is small. This suggests that the regression models could be extended further to give a more precise picture of the risk of being involved in an accident. More socio-demographic factors could be included such as the criminal records of the individuals or the medical history. The criminal records could indicate if a person perhaps is more likely to break the law in general and therefore maybe also break the traffic law. The medical history could give access to the list of medical drugs that a person may be under influence of or the person's health, in particular if the person suffer from an illnesses which may influence the behaviour in traffic. Also the drivers experience would be of interest and the type and motor size of the bike.

In the models presented in this paper, only linear expressions without interactions have been used. As there is a large data material, it may be possible to include interactions in the models since there may exist correlations between some of the factors included.

One important issue that has not been discussed in this paper is exposure in terms of the yearly travelled distance. It is assumed that all individual owning a car or a motorcycle have the same exposure which is not the case. Unfortunately, yearly access to this kind of data is not easy on an individual level, although some data exist from vehicle inspections. To be able to develop a more precise model of the data, estimates for the yearly travelled distance on an individual level is needed.

REFERENCES

Amoros, E., Martin, J.L. and Laumon, B. (2006) Under-reporting of road crash casualties in France. *Accident Analysis & Prevention*. **38**, 627-635.

Bos, N., Reurings, M., Hoeglinger, S., Leitner, T., Yannis, G., Evgenikos, P., Broughton, J., Lawton, B., Walter, L. Andreu, M., Pace, J.-F. and Sanmartín, J. (2008*) Traffic Safety Basic Facts 2007 – Motorcycles and Mopeds*. European Road Safety Observatory. http://ec.europa.eu/transport/road_safety/specialist/knowledge/pdf/bfs2007_sn-swov-13-motorcyclesmopeds.pdf

Brems, C., and Munch, K. (2008) *Risiko i trafikken 2000-2007*. DTU Transport.

DRTAIB (2009). *Motorcycle accidents*. Report 6, 2009. Danish Road Traffic Accident Investigation Board. Copenhagen.

Elvik, R. And Mysen, A.B. (1999) Incomplete accident reporting: meta-analysis of studies made in 13 countries. *Transportation Research Record: Journal of the Transportation Research Board*. **1665**, 133-140.

ETSC (2007) *PIN Flash 7: Reducing Motorcyclist deaths in Europe.* European Transport Safety Council. Brussel.

Madsen, H. and Thyregod P. (2010) *Introduction to General and Generalized Linear Models.* CRC Press. Boca Raton.

SafetyNet (2009) *Powered Two Wheelers*. Retrieved 0303-2011 from www.erso.eu.

SWOV (2009) *Motorcyclists.* SWOV Factsheet. SWOV, Leidschendam.

Værø, H. (2008). *Note on Motorcycle accidents*. Danish Road Traffic Accident Investigation Board.