Risk and protection factors in fatal accidents

Emmanuelle Dupont[1a], Heike Martensen[a], Eleonora Papadimitriou[b], and George Yannis[b]

[a] Belgian Institute for Road Safety, Behaviour and Policy Department,
1405 Haachtsesteenweg, B-1130
Brussels, Belgium

[b] National Technical University of Athens, Department of Transportation Planning and Engineering, 5
Iroon Polytechniou str. GR-15773
Athens, Greece

ABSTRACT

This paper aims at addressing the interest and appropriateness of performing accident severity analyses that are limited to fatal accident data. Two methodological issues are specifically discussed, namely the accident-size factors (the number of vehicles in the accident and their level of occupancy) and the comparability of the baseline risk. It is argued that - although these two issues are generally at play in accident severity analyses - their effects on, e.g., the estimation of survival probability, are exacerbated if the analysis is limited to fatal accident data. As a solution, it is recommended to control for these effects by (1) including accident size indicators in the model (2) focusing on different sub-groups of road-users while specifying the type of opponent in the model, so as to ensure that comparable baseline risks are worked with. These recommendations are applied in order to investigate risk and protection factors of car occupants involved in fatal accidents using data from a recently set up European fatal accident investigation database. The results confirm that the estimated survival probability is affected by accident-size factors and by type of opponent. The car occupants' survival chances are negatively associated with their own age and that of their vehicle. The survival chances are also

[1] Corresponding author: Dr. Emmanuelle Dupont,
Belgian Road Safety Institute – Behaviour and Policy Department -Chaussée de Haecht 1405 – 1130 Brussels
Phone: +32 -2-244.15.40, fax: +32-2-216 43.42
e-mail: Emmanuelle.dupont@ibsr.be

lower when seatbelt is not used. Front damage appears to be associated with increased survival probability, as compared to other damaged area, but mostly in the case in which the accident opponent was another car - or a light goods vehicle - rather than a heavy goods vehicle or than in single vehicle accidents. The interest of further investigating accident-size factors and opponent effects in fatal accidents is discussed.

Keywords: Accident severity; fatal accidents; logistic regression; opponent type; accident size

1. INTRODUCTION

The goal in collecting accident data is to learn from the past and gain information that can help preventing future accidents from occurring (crash prevention), or mitigating their consequences (crash protection). The ultimate objective of road-safety management is the reduction of the number of fatalities. As a consequence, road-safety targets are expressed and quantified as number of casualties (or as the desired reduction thereof). Well-maintained fatal accident databases are thus necessary to monitor the evolution of road-safety and the effects of the measures implemented. Detailed information on fatal accidents, on the other hand, is also sought after with the aim of increasing knowledge of fatal crashes and of developing fatal crash prevention measures. It is well-known that fatal accident data, as compared to data recorded from less severe accidents, are the most reliable. “Not only are fatalities the most serious and permanent consequence of traffic crashes, but fatality data are vastly more reliable and readily interpretable than data for any other level of harm” (Evans, 2004, p.19). As a result, databases are developed that focus on fatal accident exclusively. This is the case of the Fatal Accident Reporting System in the U.S, and more recently of the Fatal Accident Investigation database (Reed & Morris, 2009), created under the impetus of the European commission.

Such fatal accident databases cannot be used to perform analysis focusing on fatal crash prevention, unless they are linked with data from other accident severity levels.

Indeed, in order to determine which features are *specific* to fatal accidents, these have to be compared to non-fatal accidents. Yet, as Evans notes: “the majority of people involved in fatal crashes are not themselves killed” (Evans, 2004). Consequently, differentiating the survivors from the fatalities *in fatal crashes* - and thereby identifying protection factors *within* those severe crashes - is a legitimate and interesting step to take to improve existing knowledge of fatal crashes. Of course, observations that are limited to fatal accidents can only provide information that is restricted to this high-end of the accident-severity continuum. The conclusions that can be derived from such an analysis in terms of protection factors will similarly be limited to the “worst case scenarios”. But identifying the properties of the road users, vehicles, or of the accident itself that play a protective role in those extreme situations is all the more important.

Despite their potential interest, few investigations have been conducted so far on risk and protection factors in fatal accidents (Evans, 1983; 1986; Evans & Frick, 1993). Yet, limiting a severity analysis to fatal accident data also raises important methodological considerations. Although these considerations are generally at play in all accident severity analyses, they are seldom explicitly discussed, even when appropriately addressed in the models developed (e.g.: Lui; McGee; & Pollock, 1988). Below, two of these issues – here labelled “the accident size bias”, and “the comparability of the baseline risk” - are discussed: Their general effects on severity analyses are described, as well as the reasons to expect these effects to be exacerbated in the case of data limited to fatal accidents. In this paper, the risk and protection factors of road-users involved in fatal accidents are investigated on the basis of the Fatal Accident Investigation database (Reed & Morris, 2009) using. The aim is both to stress the importance of accident size of the comparability of baseline risks in this kind of investigation and to propose practical solutions to control for these factors when working with fatal accident data.

1.1 The Accident-Size bias

The size of an accident (i.e., the total number of road-users involved) is a joint function of the number of participants involved in the accident (pedestrians, passenger cars, powered 2-wheelers) and of the level of occupancy of the vehicles. Using all-severity crash data, Kockelman & Kweon (2002) have shown, for example that vehicle occupancy had no effect on driver injury risk in 2-car crashes, but that higher occupancy levels were associated with a lowered driver injury risk in single-car crashes. Chang & Mannering (1999) predicted the most severe injury sustained by car occupants in all-severity crashes using vehicle occupancy as a nesting factor. The results showed that the occupancy level corresponding respectively to property damage only, injury, and fatal accidents were 1.31, 1.53, and 1.63, suggesting a positive association between the number of occupants and the worst consequence of the accident (the most severe injury). Finally, Khorashadi et al. (2005) observed a negative relation between the level of occupancy of vehicles and the probability that the driver will be left uninjured, while the relation between the number of vehicles in the accident and the probability for the driver to be uninjured was found to be positive. The severity of an accident is consequently affected by its size. The lack of consistency in the results summarized above indicate that the overall relation between accident size and accident consequences is far from being simple, and probably depends on a number of other factors, to begin with a likely interplay between the accident size factors themselves, namely the number of participants and the vehicle occupancy levels.

Matters are different when the accident size-severity relation is examined within the restricted context of fatal accident data. In this case, the relationship between the size of an accident and its outcomes can be considered as a bias. Indeed, it results mainly from the selection criterion applied during the data collection: Each accident recorded in a fatal accident database generated one fatality at least. As a consequence, the presence of survivors in the same accident most crucially depends on whether or not more than one person was involved. The probability to survive will inevitably be estimated as 0 for single car-occupants in single-vehicle accidents, and steadily increase with the number of occupants in vehicles, as well as with the number of accident participants.

Although they are seldom explicitly discussed in the literature, the effects of accident-size factors are usually dealt with in accident-severity models. The effects of vehicle occupancy on estimates of accident severity are often controlled for by selecting drivers as units of observation (e.g., O'Donnell & Connor, 1996; Kockelman & Kweon, 2002; Shibata & Fukuda, 1994; Martin & Lenguerrand, 2008), or by including occupancy as a predictor in the model (Kockelman & Kweon; Chang & Mannering, 1999; Khorashadi et al., 2005). The number of accident participants is, on the other hand, usually maintained constant by selecting accidents with a given number of participants and focusing, for example, on two-car crashes or single-car crashes (e.g.: Savolainen & Mannering, 2007; Yau, 2004; Khorashadi et al., 2005; Martin & Lenguerrand, 2008). The main disadvantage of most of these methods, however, is that the desired level of control is attained at the costs of data losses. Some of them, such as the selection of drivers as units of observation may remain problematic when working exclusively with fatal data: It does not allow to fully control for the effects of occupancy levels on the dependent variable ("severity"), since the risk for the driver to sustain, say, a fatal injury still depends on the level of occupancy of his/her vehicle. To avoid this problem, only cars occupied solely by the driver have to be selected (e.g., Evans, 1984), meaning even less data available for the analysis and further restrictions imposed to the generalisability of the results. Finally, working on the basis of driver data often poses problems in interpreting results related to the individuals' characteristics. As an example, drivers who wear a seatbelt are known to be less often involved in severe accidents, so that it is difficult to determine with certainty whether lower severe injury probability for belted drivers reflect their lower involvement in severe crashes, or whether unbelted drivers are indeed more at risk for severe injuries (Evans, 2004). When the results are based on all car occupants, accident risk is not confounded with injury risk any more.

1.2 The comparability of baseline risks

Accident severity models focus on the risk ran by road users to sustain one or several types of injuries, once involved in an accident. Whatever the particular injury risk that is focused upon (i.e.: the fatality risk, or the risk to sustain fatal vs. severe vs. slight vs. no injury, and so on), the initial risk, or the “baseline risk” ran by each accident protagonist strongly depends on their respective modes of transport, and on that of the road-user they collide with in the course of the accident: The injury risk ran by a pedestrian differs strongly from that of a car driver, and so does the risk for a car driver vary a lot depending on whether he/she collides with a motorcycle or with a heavy goods vehicle (HGV). When the aim of the analysis is to compare the survivors and the fatalities in accidents so as to identify risk/protection factors, it is important to ensure that one and the others have comparable “baseline risks”. Most often, the road-users’ and their opponents’ modes of transport are controlled for through the selection of well-defined accident types (e.g.: Martin & Lenguerrand, 2008). Other models have been developed that include the respective transport modes of the road-users and of their opponents as predictors (e.g.: Kockelman et al., 2002; Khorashadi et al., 2005). Quite often, however, the road-users’ transport mode is the only one that is controlled for (Yau, 2004; 2006; Shibata & Fukuda, 1994)

The importance of the comparability of baseline risks is exacerbated when dealing exclusively with fatal accident data. Figures 1 to 3 were calculated on the basis of the Fatal Accident Investigation data. These figures show - each for a different type of road-users - the variation in the proportion of fatalities, of severe or slight injuries, and of uninjured people associated with different opponent types. Two things can easily be noticed on the basis of these graphs: Firstly, there are road-user groups for which the outcomes of the accident hardly vary namely HGV occupants (Figure 1), and pedestrians (Figure 2). The second important point is that, for other types of road-users – such as car occupants (Figure 3) – the type of *opponent* appears to exert a strong influence on the fatality risk. Indeed, while car-occupants survived the vast majority of collisions with vulnerable road-users, they most frequently deceased as a result of collisions with heavy and light good vehicles. The picture is different for car occupants who collided with another car. Among the latter, the various types of consequences are more evenly distributed.

Insert Figure 1 to 3 about here

The analysis presented here focuses on car occupants and integrates the type of collision opponent into the model. Car-car accidents – the most frequently encountered accident type in the Fatal Accident Investigation database - are taken as the reference category to which car-LGV; car-HGV; and single car accidents are compared. Car accidents with vulnerable road users (pedestrians, 2-wheelers) are not included in the analysis, given the very high proportion of car occupants who survived those accidents.

2. METHOD

2.1 Data

The “Fatal Accident Investigation” database (or FAI database) was created during the 6th framework programme of the European Commission, with the aim – among others – of boosting knowledge acquisition about fatal crashes in Europe. Seven countries took part in the data collection (United Kingdom; France; Sweden; Finland; Germany; Italy; and the Netherlands). The data collected were intended to be representative of the national populations of fatal accidents within these countries. The FAI database was developed on the basis of existing accident investigation infrastructures, working with retrospective investigation methods: Information derived from the police documentation of fatal accident investigations in each country were complemented with information from hospital, insurance companies, and prosecution records (Morris and Reed, 2006). It concerns the accident itself (e.g.: the number, type and sequence of events, the type of infrastructure, etc), as well as the vehicles (weight, width, length, age, the manoeuvre executed by the driver, and so on) and the road-users involved (level of impairment, familiarity with the road, etc). Given the method of investigation adopted, the information stored in the database for each accident case is detailed and various, as compared to national accident records, for example. The total dataset contains some 1,300 fatal accident cases involving around 3,500 road users in total. To ensure that comparable risk estimates will be computed for the present analysis, only data for car-occupants were selected, reducing the dataset to a total of 1638 individual car-occupants (both drivers and passengers).

2.2 Assignment of an opponent to the car occupants:

In the Fatal Accident Investigation database, the information about each accident is recorded along the different “events” that made up the accident (up to 6 events per accident). As a consequence, within one accident, one given road user can happen to have interacted with two different vehicles. Adequately selecting the opponent that will be assigned to each road-user is most important for the objectives of the analysis. This was done using the chronological development of the accident as a reference: The earlier opponents (those with whom the road-user interacted first) were considered as the most “significant” ones, and consequently designated as opponents over those the road-user collided with in the later stages of the accident. In most cases (91.5%), the event which described the (first) interaction in the accident proved to have been rated by the investigation team as being the “most harmful event” for the road-user as well.

The pattern of some accidents was very complicated and could even prove ambiguous. This was typically the case of accidents involving parked vehicles. Parked vehicles can be both vehicles that are parked along the road at some place where they are to be expected, but also vehicles stopped for one reason or another at some unexpected place (e.g.: defective vehicle). In such cases, the status of the road-users themselves can change in the course of the accident (the driver who steps out of the car to check the motor and gets hit by a truck is no car driver anymore, but a pedestrian). Extra-attention was devoted to accidents involving parked vehicles so as to ensure that the opponent had been rightly coded for all road-users. When there was too much ambiguity about the exact nature of the interaction, the accident was left out of the analysis.

2.3 Model Specification

The model proposed consists of an attempt at predicting the survival probability of car occupants involved in fatal accidents depending on a number of crash, vehicle, and personal characteristics. Survival is here defined in the broad sense of the term: Road-users who suffered serious to slight injuries, as well as uninjured road-users are said to have survived the accident. By contrast, a fatality is defined in accordance with the common European definition, namely: as someone who died within the 30 days following the accident. Survival was coded as 1, while fatality was coded as 0. The binary (0,1) response for the i^{th} unit (here, car occupant) is denoted by y_i. The probability that $y_i = 1$ is denoted by π_i.

Given the binary nature of the response variable, a binomial logistic regression model was used, which can be specified as:

$$f(\pi_i) = \beta_0 + \beta_1 x_i$$

Where the chosen link function $f(\pi_i)$ was the logit link, so that:

$$f(\pi_i) = \log it\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 x_i$$

Generally speaking, accident data are hierarchically organised: At the lowest level individual road-users – who eventually sustain the consequences of accidents – are «nested» within vehicles, which form the second level of data. The vehicles are in turn nested within the accidents. This makes accidents the third level in the data. The presence of these different levels is likely to introduce correlations between the observations, leading to violations of the assumption of independent errors. For these reasons, hierarchical structures in accident data are receiving increasing attention from the road safety research community (Lenguerrand, Martin, et Laumon, 2006; Dupont & Martensen, 2007; Jones & Jorgensen, 2003), and endeavours are made to apply statistical models that allow taking the hierarchical relations among the observations into account. These models are known as multilevel models, (e.g., Goldstein, 2003; Hox, 2002; Kreft and de Leeuw, 2002; Snijders and Bosker, 1999), random-coefficient regression models (Longford 1993), or mixed effects models (Pinheiro & Bates, 1995). Random variation in the observations is partitioned and assigned to the different levels identified, and the significance of the corresponding variance estimates is tested. Significant random variation at higher levels in the data (i.e.: second, third level) indicates that multilevel models should be applied.

The presence of random variation at the country, accident, and vehicle levels in the fatal accident data was consequently tested in a preliminary test, by fitting the following three-level empty model:

$$\log it(\pi_{ijk}) = \beta_{0ijkl} + \beta_{1_i} x_i + \varepsilon_{ijkl}$$
$$\beta_{0ijkl} = \beta_0 + f_{0l} + v_{0kl} + \mu_{0jkl}$$

It is further assumed that:

$$f_{0l} \sim \mathrm{N}\ (0,\ \sigma^2{}_{f0kl})$$
$$v_{0kl} \sim \mathrm{N}\ (0,\ \sigma^2{}_{v0kl})$$
$$\mu_{0jkl} \sim \mathrm{N}\ (0,\ \sigma^2{}_{\mu 0jkl})$$

Where π_{ijk} corresponds to the probability that y_{ijk} will be one, so that road user i in car j in country k will survive. f_{0l}, v_{Okl}, and μ_{0jkl} refer to the random variation at the country, accident, and vehicle levels, respectively. None of these variance estimates was significant, indicating no serious correlation problems between the observations for car occupants in the same car ($\sigma^2 \mu_{0jkl} = 0.01, n.s.$) and in the same crash ($\sigma^2 v_{0kl} = 0.01, n.s.$), or for observations within the same countries ($\sigma^2 f_{0l} = 0.05, n.s.$). Yet, compared to other data hierarchies, accident data present some peculiarities that prevent any straightforward application of multilevel models for their analysis.

The most important is the fact that there are usually few road-users per vehicle. For this reason, computational problems may be encountered, or the random effects can be erroneously estimated as null (Lenguerrand et al., 2006). For this reason, a three-level version of the final, adjusted model was also fitted. The values and significance of the coefficients in the single and in the multilevel version of the model were similar. On the basis of these results, the choice was made to leave the model in its single-level state, and to perform a standard logistic regression analysis.

3 RESULTS

Table 1 presents the different variables that have been tested; their means and standard deviations, as well as the coding scheme that was applied to each (the reference category is indicated in italics).

Insert Table 1 about here

Table 2 provides the estimation results (odds ratios) for the different predictors. The estimates obtained when variables were entered alone in the model as predictors of the survival probability (univariate tests) are presented in the left-hand columns of the table. Estimates resulting from the final model – that are thus adjusted on the basis of the effects of other predictors in the model - are provided in the right-hand columns. Generally speaking, estimates below 1 indicate a negative association between the corresponding predictor and the fatality risk, while estimates above 1 reveal a positive relationship.

The univariate estimates for the Opponent effect indicate that the survival chances of car occupants involved in car-car accidents are higher than that of car occupants who collided with any of the three other opponent types (the survival chances of car occupants in accidents with light and heavy goods vehicles, or in single car accidents are respectively 3.45, 5, and 2.27 times lower than those of car occupants in car-car accidents). Yet, on the basis of the adjusted estimates, it can be concluded that only the survival chances of car occupants involved in accidents with HGV significantly differ from those of car occupants involved in car-car accidents. This change from the univariate to the adjusted estimates is attributable to the inclusion of the term representing the interaction between the type of opponent and the area of the vehicle that was damaged mostly as a result of the accident. This interactive effect will be further discussed later, when examining the results for the “damaged area” variable.

A positive relation was also observed between car occupants' survival probability and the two accident-size indicators: The higher the number of vehicles involved in the accident and the level of occupancy of these vehicles, the higher the probability for each car occupant to survive.

As far as road-user characteristics are concerned, the adjusted estimates indicated that the survival probability of car occupants decreases with their age. This result is in line with the conclusions of studies conducted on all-severity accident data: (e.g., O'Donnell & Connor, 1996; Kockelman et al., 2002; Martin & Lenguerrand, 2008). Skyving, et al. (2008) used fatal accident data to compare – for different, pre-defined accident patterns – the proportions of those crashes that proved fatal to older vs. to younger drivers. The results showed that the "baseline proportion" was 60%, thus showing that "when older drivers are involved in fatal road traffic crashes, they are often the one who dies" (ibid, p..

The results of the univariate tests also indicated that being female; being passenger or being seated at the back of the car was associated with increased survival chances. However, none of these effects remained significant after adjustment for other predictors in the final model. Actually, the change in the value and significance of all three estimates is attributable to the inclusion of the number of occupants in the model. The sample distributions indeed show a strong overlap between each of these factors and occupancy level. For example, while 81% of the single car occupants were male this percentage dropped to 68% in cars with two occupants, and then kept on decreasing as the number of occupants increased. Evans (2004) came to similar conclusions on the basis of the FARS data: "The number of fatalities in a seat is determined mainly by the occupancy of that seat (…) Occupants in different seats have different distributions by gender and age, factors that influence fatality risk in a crash" (p. 53). With respect to previous research, it is difficult to conceive that being female consists of a protection factor for car occupants. The conclusion that the univariate relationship observed between gender and survival is spurious, and has to be attributed to a confounding between gender and occupancy thus seems logical. Matters are less straightforward as far as seat position is concerned: Although in the present case the seat position effect was fully accounted by occupancy level, it should be mentioned that rear seats have been found to be safer, and this on the basis of fatal accident data (Evans & Frick, 1984, quoted in Evans, 2004, see also Shinar, 2007). However, these conclusions are based on analyses that allowed disentangling – among others - occupancy, and gender effects, since it was based on the comparison of the fatality rate of same age-and-gender drivers and passengers. On the basis of the adjusted model, seatbelt also appeared to be a protection factor for road-users involved in fatal accidents, since the survival probability of occupants who did not wear a seatbelt was estimated to be two times lower than that of unbelted car occupants[2]. This is an important finding, given the overall high level of severity that characterises the accidents examined here. The protective effect of seatbelt that is observed here is also in line with previous results based on all-severity accident data, which led to the conclusion that failure to wear a seatbelt increases the probability of death (O' Donnell & Connor, 1996; Shibata & Fukuda, 1994; Martin & Lenguerrand, 2008). Evans (2004) reached similar conclusions on the basis of fatal accident (FARS) data.

The age of the vehicle appeared to be negatively related to the occupants' survival chances, an effect that remains significant after adjustment for other predictors in the model. This result is coherent with previous findings based on all-severity accident data, where vehicle age appeared to be positively related with probability of death (O'Donnell & Connor, 1996), with the probability for one occupant in the vehicle to sustain fatal or severe injury (Yau, 2006); or with the probability for the driver to be killed (Martin & Lenguerrand, 2008).

[2] The « seatbelt » variable was initially made up of 4 categories in the database. For the present analysis, two categories "Used" and "Used claimed" have been merged. The "Used Claimed" observations correspond only to a total of 2% of the observations. It might be the case, however, that a substantial part of people having claimed to have used their seatbelt actually haven't done so. This should not be a major problem, however, given the few number of observations in the "Use Claimed" category. Moreover, should the "used claimed observations" indeed correspond to people that did not wear a seatbelt, then their inclusion in the "Used" category should only weaken the chances that the "Used" – "Not Used" comparison leads to the conclusion of a protective effect of seatbelt wear.

Finally, the fact that the vehicle driver did – or did not – brake at the moment of the accident also affects the survival probability of the occupants: The results of the final model indicate that the survival probability of occupants in vehicles whose driver did not brake is 1.35 times lower than that of the occupants of cars whose drivers braked. This finding offers no straightforward interpretation, however, as it may also indicate that drivers have had more time to react by braking in accidents occurring at lower speed, which might consequently also be less severe accidents. In a similar vein, speed is likely to account for the fact that car occupants involved in accidents occurring on a road junction had higher survival chances than car occupants in accidents occurring on a road section. Unfortunately, the number of missing values in the database for the vehicle's actual speed did not allowed this variable to be included in the model.

The last vehicle variable that was found to affect the occupants' survival chances is the area of the car that was mostly damaged, here coded as "front vs. other". The results showed that survival chances were higher when the car was damaged at the front mostly (as compared to any other area). To interpret this effect, it is important to note that the left and right sides of the vehicle make up the most important part of the "other damaged areas" category (59% altogether). The result obtained could thus be considered consistent with previous results, based on all-severity accident data, showing that the risk for a driver to be killed was the highest in side-impacted vehicles (Martin & Laumon, 2008). Working with the FARS data, Evans (2001) came to a similar conclusion: a driver in a left-impacted car was found to be about ten times more likely to be killed than a driver in a car with frontal damage, while a driver in a right-impacted car was between 4 and 5 times more likely to die than a driver in a front-impacted car. To a certain extent at least, this result can be attributed to improvements in vehicle crashworthiness, many of which have focused on vehicles front. Elvik & Vaa (2004) examined the effects of crashworthiness on the number of injured persons in accidents, and came to the conclusion that collapsible steering columns, laminated and better-fastened front windscreens, padding and changing the design of the instrument panel were likely to be associated with a decrease of some 12 to 22% of fatal injuries in head-on frontal impacts accidents. The analysis performed here further revealed a significant interaction between the variables “most damaged area” and “opponent type”. This significant interaction indicates that the protective effect of front damage does not hold for car occupants involved in accidents with heavy goods vehicles and in single-car accidents. This interactive effect also appears to explain a good deal of the “Type of Opponent” main effect. Indeed, after inclusion of the interaction term in the model, the estimate for accidents with heavy-good vehicles is the only one that remains significantly different from the car-car accident. This is understandable when considering that, after the inclusion of the “opponent-most damaged area” interaction term, the opponent main effect holds only for the reference category of the “Most Damage” variable, namely: “Other Damage”. In this case, the probability to survive the accident for car occupants in accidents with other cars is more elevated, and does not differ much from that of car occupants in accidents with Light Goods Vehicles or in single car accidents. For the same reason, the adjusted estimate for the “Front Damage” main effect is stronger than the univariate one (it holds only for car-car accidents, the reference category for the “Type of Opponent” variable, which is precisely the category for which the effect of front vs. other damaged area is the most distinct).

4 DISCUSSION

4.1 Main findings

The present study examined the risk/protection factors of car-occupants involved in “extreme traffic situations”, that is to say: fatal accidents. The approach adopted here stressed the influence of accident size and opponent factors, which are likely to be exacerbated when working with data limited to fatal accident. The analysis allowed deriving a series of factors that appear to make a difference for car occupants, once they are involved in a fatal accident.

As expected, accident-size factors were found to affect the survival probability: The larger the number of vehicles involved in the accident and the number of occupants inside the cars, the higher the chance for each of them to survive. Including level of occupancy as a predictor in the model led to rule out gender, back vs. front seat position, and driver vs. passenger status as protection factors, hence confirming the importance of controlling for these factors.

The results also showed that the survival probability of car occupants depends much on the type of vehicle their car collided with during the accident: Fatal accidents with other cars seem generally safer for car occupants than accidents with any of the 3 types of opponent investigated here (heavy and light goods vehicles and no opponent, i.e.: single vehicle accidents). This effect, however, was much reduced by the inclusion of the interaction with the “most damaged area” variable in the model. Car occupants involved in accidents with other cars appear indeed appear to be better-off, but only to the extent that the car was damaged at the front mostly. When some other area (in this case, the sides mainly) of the car was damaged, the survival chances of car occupants in car-car accidents were just as low as those of car occupants involved in other accident types, and only car occupants involved in crashes with HGVs suffered even lower survival chances. All in all, these results stress the vulnerability of car occupants to accidents involving HGVs, and the need to enhance side protections for car occupants.

Although exclusively based on fatal accident data, many of the results obtained here appear to be generally consistent with the conclusions of severity analyses performed with data characterised by a broader severity range. For example, the survival chances of car occupants involved in fatal accidents are lower as their age and the age of the car they are travelling in increase, and if they do not wear a seatbelt. This confirms the role of these factors in extremely severe accident situations.

4.2 Strengths and limitations of the approach

The database used for the present analysis is only of a moderate size as compared to the FARS database, to name just the most important one. A very large database offers the possibility to apply sophisticated methods – such as the double-pair comparison method (Evans, 1986) – which in turn allow a very detailed and controlled investigation of the causal role of specific factors with respect to accident

severity. Such methods, however, can only be used at the costs of (important) data losses and are not applicable with smaller databases.
The approach adopted here offers the advantage of minimizing data losses, while handling important methodological issues raised by the analysis of fatal accident data. One of its main limitations lies in the level of detailed attained, that does not fully meet up the one offered by the data themselves. The outcome investigated here, for example, distinguished only survival from death, with the result that “survival” also comprises less-than-optimal survival conditions. This is likely to result in a somewhat truncated image of the outcomes of fatal accidents, as well as to obscure the effect of the factors investigated on other types of outcomes (e.g.: severe injury, slight injury). Similarly, several potentially important and interesting predictors could not be included in the model because of the number of missing observations, despite of careful data recording from the investigation teams. One must bear in mind that the conclusions attained on the basis of this model remain mainly conditional on the particular predictors that have – or have not - been used. As already noticed, the results observed for the variables “Junction” and “Braking” might be somewhat different would it be possible to add a measurement of the vehicles’ actual speed in the model.
This lack of detail did not, however, preclude the obtaining of results that appear both sensible in the context of previous research, and informative. In particular, the approach developed here allows the investigation of the generalisability of the effects observed across several sub-groups of road users (e.g., car occupants in crashes with other cars, light or heavy goods vehicles, with one or several vehicles, characterised by different occupancy levels…). In terms of efficiency, and keeping in mind that the ultimate goal is the identification of detailed and reliable effects, this is a potentially interesting approach, even if used only as a preliminary data analysis.

Finally, it is worth mentioning that discussing the effects of accident-size and opponent effects in the restricted context of fatal accident data inevitably raises questions on these factors’ general role with respect to accident severity. Accident-size, on the one hand, is a function of several factors (occupancy, vehicle numbers) and our short review of related findings suggest that its effects are likely to be complex. The ability to isolate opponent effects in investigations of accident severity is, on the other hand, of tremendous importance with respect to our comprehension of risk and protection factors in road accidents. The type of opponent has already been shown to affect the development of the risk for subgroups of road-users over the years (Stipdonk & Berends, 2008). Previous attempts at including opponent effects in accident severity models often yielded the conclusion that a single attribute can play both a protective and a deleterious role on accident severity on the sole basis of who – the investigated road-user or his/her opponent – displays this attribute, as is typically the case with superior car mass (Martin & Lenguerrand, 2008; Evans, 1984; Evans & Frick, 1993). The two types of factors would deserve making the object of more attention in road-safety research.

Acknowledgements

This paper is based on work carried out within the SafetyNet integrated project co-financed by the European Commission. The authors would like to address special thanks to the SafetyNet WP5 research teams on in-depth accident and injury data led by VSRC of the University of Loughborough, for establishing and making available the Fatal Accident Investigation Database, as well as for their assistance and feedback throughout this research. The WP7 research team has also offered highly valuable comments and feedback on previous versions of the work presented here. Finally, we would like to thank Jean-Louis Martin (INRETS) and Rachel Elliman (VSRC) for their comments on a draft version of this paper.

References

Chang, L.Y., & Mannering, F., 1999. Analysis of injury severity and vehicle occupancy in truck- and non-truck involved accidents. Accident Analysis and Prevention (31), 579-592.

Dupont, E. and Martensen, H. (Eds.), 2007. Multilevel modelling and time series analysis in traffic research - Methodology. Deliverable D7.4 of the EU-FP6 project SafetyNet, European Commission, Brussels. Available on-line at: http://www.erso.eu/safetynet/fixed/WP7/D7_4/D7.4.pdf

Dupont, E., and Martensen, H. (Eds.), 2008. Analysing European in-depth data: Methodological framework and results. Deliverable D7.9 of the EU-FP6 project SafetyNet, European Commission, Brussels.

Elvik, R., & Vaa, T., 2004. The handbook of road-safety measures. Oxford: Elsevier Science.

Evans, L., 1983. Driver fatalities versus car mass using a new exposure approach. Accident Analysis and Prevention, 16 (1), 19-36.

Evans, L., 1986. Double pair comparison – a new method to determine how occupant characteristics affect fatality risks in traffic crashes. Accident Analysis and Prevention, 18, 217-227.

Evans, L., & Frick, M. C., 1993. Mass ratio and the relative driver fatality risk in two-vehicle crashes. Accident Analysis and Prevention, 25 (1), 213-224.

Evans, L., 2001. Causal influence of car mass and size on driver fatality risk. American Journal of Public Health, 91, 1076-1081.

Evans, L., 2004. Traffic Safety. Science Serving Society: Bloomfield Hills, Michigan.

Jones A. P., & Jørgensen, S. H., 2003. The use of multilevel models for the prediction of road accident outcomes. Accident Analysis and Prevention, 35, 59-70.

Khorashadi, A., Niemeier, D., Shankar, V., & Mannering, F., 2005.. Differences in rural and urban driver-injury severities in accidents involving large trucks: An explanatory analysis. Accident Analysis and Prevention, 37, 910-921.

Kockelman, K. M., & Kweon, Y. J., 2002. Driver injury severity: An application of ordered probit models. Accident Analysis and Prevention, 34, 313-321.

Lenguerrand, E., Martin J.L., & Laumon, B., 2006. Modelling the hierarchical structure of road crash data – Application to severity analysis. Accident Analysis and Prevention, 38, 43-53.

Lui, K. J.; Mc Gee, D.; Rhodes, P. ; & Pollock, D., 1988. An application of conditional logistic regression to study the effects of safety belts, pricipal impact points, and car weights on drivers' fatalities. Journal of Safety Research, 19, 197-203.

Martin, J.L., & Lenguerrand, E., 2008. A population-based estimation of the driver protection provided by passenger cars: France 1996-2005.

Morris A., Reed D. (Eds.), 2006. Review of the fatal accident pilot study - Brief report on methodological developments. Deliverable 5.3 of the SafetyNet project, European Commission, Brussels. Available on-line at: http://www.erso.eu/safetynet/fixed/WP5/D5.3%20Review%20of%20fatal%20accident%20pilot%20study_Fatal%20data%20methodology%20development%20report.pdf

National Highway Traffic Safety Administration, Fatality Analysis Reporting System (FARS). Web-based encyclopedia. Data files and procedure to analyze them: http://www-fars.nhtsa.dot.gov

O' Donnell, C. J., & Connor, D. H., 1996. Predicting the severity of motor vehicle accident injuries using models of ordered multiple choice. Accident Analysis and Prevention 28 (6), 739-753.

Pinheiro, J.C., & Bates, D. M., 1995. Approximations to the Log-Likelihood Function in the Nonlinear Mixed-Effects Model Journal of Computational and Graphics Statistics 4, 12-35.

Reed, S., & Morris, A., 2009. Fatal accident database development and analysis. Deliverable D5.7 of the EU FP6 SafetyNet project. Available on-line: http://www.erso.eu/safetynet/fixed/WP5/D5%207%20final%20report.pdf

Savolainen, P., & Mannering, F., 2007. Probabilistic models of motorcyclists' injury severities in single- and multi-vehicle crashes. Accident Analysis and Prevention, 39, 955-963.

Shibata, A., & Fukuda, K., 1994. Risk factors of fatality in motor vehicle traffic accidents. Accident Analysis and Prevention, 26 (3), 391-397.

Shinar, D., 2007. Traffic safety and human behaviour. Elsevier: Amsterdam.

Skyving, M., Berg, H. Y., and Laflamme, L., Older drivers' involvement in fatal RTCs. Do crashes fatal to them differ from crashes involving them but fatal to others? Safety Sci. (2008), doi: 10.1016/j.ssci.2008.09.001

Stipdonk, H., & Berends, E., 2008. Distinguishing traffic modes in analysing road safety development. Accident Analysis and Prevention, 40, 1383 – 1393.

Yau, K., 2004. Risk factors affecting the severity of single vehicle traffic accidents in Hong-Kong. Accident Analysis and Prevention, 36, 333-340.

Yau, K., Lo, H. P., Fung S., 2006. Multiple-vehicle traffic accidents in Hong Kong. Accident Analysis and prevention, 38, 1157-1161.

Figure Captions:

Figure 1: *Proportion of fatal, serious, slight, and no injuries among heavy goods vehicle occupants as a function of the type of opponent*

Figure 2: *Proportion of fatal, serious, slight, and no injuries among pedestrians as a function of the type of opponent*

Figure 3: *Proportion of fatal, serious, slight, and no injuries among car occupants as a function of the type of opponent*

Explanatory variables				
		Mean/ Proportion	S.D.	Definition
Number of participants		1.83	0.85	Total number of vehicles involved in the accident
Number of occupants		2.53	1.40	Number of occupants in the vehicle
Age		36.47	18.61	Age of the car-occupant
Vehicle Age		8.20	4.80	Age of the vehicle the road-user was travelling in
Opponent Type				= 1 if LGV, =2 if HGV, = 3 if single car accident, =0 if another car
	LGV	0.53	0.50	
	HGV	0.03	0.16	
	Single Car	0.09	0.29	
	Car	0.36	0.48	
Gender	Female	0.30	0.46	= 1 if car occupant is female, 0 otherwise
	Male	0.70		
UserClass	Passenger	0.46	0.50	= 1 if car occupant is passenger, 0 otherwise
	Driver	0.64		
Position in Vehicle	Back	0.15	0.36	= 1 if car occupant was at the back, 0 otherwise
	Front	0.85		
Seatbelt				= 1 if seatbelt not used, 2 if seatbelt wear unknown, 0 if seatbelt used
	Not Used	0.45	0.50	
	Unknown	0.18	0.38	
	Used	0.37	0.48	
Day	Day	0.57	0.49	= 1 if accident took place during night, 0 otherwise
	Night	0.43		
Weekend	Week-end	0.48	0.50	= 1 if accident took place in the week-end, 0 otherwise
	Week	0.52		
Junction	Yes	0.27	0.44	= 1 if accident took place on junction, 0 otherwise
	No	0.73		
Most damaged car area	Front	0.54	0.50	= 1 if front of vehicle sustained most damages, 0 otherwise
	Other	0.46		
Braking	No	0.79	0.41	= 1 if driver did not brake at the moment of the accident, 0 otherwise
	Yes	0.21		

Table 1: Explanatory variables (continuous predictors were centred around their mean)

		Univariate tests				Adjusted Estimates			
		Exponential	CI-	CI+	P-Value	Exponential	CI-	CI+	P-Value
Collision Opponent	LGV	0.29	0.15	0.41	<.0002	0.45	0.13	1.49	NS
	HGV	0.20	0.14	0.25	<.0001	0.41	0.22	0.75	<.004
	Single Car	0.44	0.35	0.49	<.0001	1.26	0.82	1.95	NS
	Car	1.00	1.00	1.00		1.00	1.00	1.00	
Number of participants		1.44	1.26	1.54	<.0001	1.64	1.33	2.01	<.0001
Number of occupants		1.58	1.45	1.64	<.0001	1.78	1.58	2.01	<.0001
Age of R.U		0.98	0.98	0.99	<.0001	0.99	0.98	0.99	<.0001
Gender	Female	1.31	1.05	1.46	0.013	1.08	0.82	1.41	NS
	Male	1.00	1.00	1.00		1.00	1.00	1.00	
UserClass	Passenger	1.70	1.39	1.88	<.0001	0.85	0.63	1.14	NS
	Driver	1.00	1.00	1.00		1.00	1.00	1.00	
Position in Vehicle	Back	2.18	1.60	2.55	<.0001	1.10	0.72	1.68	NS
	Front	1.00	1.00	1.00		1.00	1.00	1.00	
Seatbelt	Not Used	0.49	0.37	0.56	<.0001	0.43	0.31	0.61	<.0001
	Unknown	1.07	0.86	1.19	NS	0.85	0.66	1.10	NS
	Used	1.00	1.00	1.00		1.00	1.00	1.00	
Vehicle Age		0.96	0.94	0.97	<.0002	0.97	0.95	0.99	<.008
Day	Day	1.03	0.84	1.14	NS	1.11	0.87	1.41	NS
	Night	1.00	1.00	1.00		1.00	1.00	1.00	
Weekend	Week-end	1.16	0.95	1.29	NS	0.93	0.73	1.17	NS
	Week	1.00	1.00	1.00		1.00	1.00	1.00	
Junction	Yes	1.44	1.15	1.61	<.002	1.62	1.23	2.15	<.0006
	No	1.00	1.00	1.00		1.00	1.00	1.00	
Most damaged car area	Front	1.85	1.51	2.04	<.0001	2.86	2.03	4.03	<.0006
	Other	1.00	1.00	1.00		1.00	1.00	1.00	
Braking	No	0.65	0.51	0.73	<.0002	0.74	0.56	0.97	<0.02
	Yes	1.00	1.00	1.00		1.00	1.00	1.00	
Collision Opponent & Damaged Area	LGV	1.06	0.25	2.18	<0.01	0.92	0.20	4.25	NS
	HGV	2.26	1.01	3.38	<.0001	0.45	0.19	1.08	<.007
	Single Car	2.39	1.50	3.02	<.0001	0.35	0.21	0.60	<.0001

Table 2: Relative risk for each car-occupant to survive the accident – Binomial logistic regression model