

Effects of Vehicle Mass and Other Parameters on Driver Relative Fatality Risk in Vehicle-Vehicle Crashes

Yibing Shi and Guy Nusholtz
Chrysler Group LLC

ABSTRACT

Regression models are used to understand the relative fatality risk for drivers in front-front and front-left crashes. The field accident data used for the regressions were extracted by NHTSA from the FARS database for model years 2000-2007 vehicles in calendar years 2002-2008. Multiple logistic regressions are structured and carried out to model a log-linear relationship between risk ratio and the independent vehicle and driver parameters. For front-front crashes, the regression identifies mass ratio, belt use, and driver age as statistically significant parameters (p-values less than 1%) associated with the risk ratio. The vehicle type and presence of the ESC are found to be related with less statistical significance (p-values between 1% and 5%). For front-left crashes the driver risk ratio is also found to have a log-log linear relationship with vehicle mass ratio.

The driver risk ratio relationship to the vehicle mass ratio is consistent with conservation of momentum and can be used with the literature reported relationship between absolute risk and velocity change, to predict the dependence of the driver fatality risk on its own mass when crashing into other vehicles in a fleet. For front-front crashes, the model indicates that when the subject vehicle mass is at half the fleet mean mass, the risk for the driver is about 3 times that of a driver in a vehicle at the fleet average mass, and it is about 50%, when the vehicle mass is twice the fleet average.

The difficulty in defining independent effects from size and mass using existing field data is discussed for the front-front crash data. Results show that out of the vehicle mass ratio, wheelbase ratio, and footprint ratio, the mass ratio is the best single predictor of the risk ratio: combining these "size" ratios with the mass ratio degrades the fit due to the high correlations between them. Examination of conditional risk shows that, for front-front crashes, the driver fatality does not show a significant dependence on vehicle wheelbase when the mass ratio of the two vehicles is already taken into account.

CITATION: Shi, Y. and Nusholtz, G., "Effects of Vehicle Mass and Other Parameters on Driver Relative Fatality Risk in Vehicle-Vehicle Crashes," *SAE Int. J. Trans. Safety* 1(1):2013, doi:10.4271/2013-01-0466.

INTRODUCTION

The relationship between occupant fatality risk and characteristics of the vehicle, occupant, and crash environment has been investigated through examination of the field accident data [1,2,3,4,5,6,7,8,9,10,11,12,13] or combined field data and crash mechanics models [14-15]. Using the Fatality Analysis Reporting System (FARS) data, Evans [1] found the following relationship for vehicle to vehicle frontal crash.

$$\frac{r_1}{r_2} = \left(\frac{m_2}{m_1}\right)^\alpha \quad (1)$$

Equation (1) does not by itself provide the relationship between the *absolute* risk for each of the vehicles and the masses. Partially based on estimations of the fatality risks in crashes between vehicles of equal masses, Evans extended

the work in [1] and hypothesized the following relationship of absolute risk as a function of mass [3]:

$$r_1 = \frac{k}{m_1 + m_2} \left(\frac{m_2}{m_1}\right)^{\alpha/2} \quad (2)$$

Separately, using the National Automotive Sampling System (NASS) data, Joksche [16] found the following relationship between the fatality risk and the velocity change (v) of the vehicle:

$$r = \left(\frac{v}{v_0}\right)^\alpha \quad (3)$$

where v_0 is a normalizing velocity of 70.6 mph or 83.4 mph depending on how v is calculated/estimated, while α is a constant with a value in the range from 3.88 to 4.08 (The empirical fit of Eq. (3) clearly needs to be capped at 1.0 as r

is a probability. This does not create a practical issue as few accidents take place as velocities higher than 70 mph).

In a vehicle-vehicle crash, the closing velocity is an important quantity in determining the fatality outcome. Velocity change of each of the vehicles can be related to the closing velocity Δ , which can be approximated using conservation of momentum and the assumption that the collision is fully plastic:

$$v_1 = \frac{m_2}{m_1 + m_2} \Delta \quad (4)$$

Substituting Eq. (4) into Eq. (3), we have

$$r_1 = \left(\frac{2m_2}{m_1 + m_2} \right)^\alpha \left(\frac{\Delta}{2v_0} \right)^\alpha \quad (5)$$

It is noted that Eq. (5) and Eq. (2) show a different dependence of fatality on the masses of the vehicles, although both are consistent with Eq. (1).

Beyond vehicle mass, other vehicle parameters may have an influence on the fatality outcome. However, what these parameters are and how they quantitatively influence the fatality are not obvious. Wheelbase of the vehicle, being a well-defined and readily available parameter has been the main indicator of the “size” of the vehicle used in the literature. Through an estimation of the relationship between the wheelbase and mass of vehicles, Evans [3] proposed that the empirical relationship of Eq. (2) should be rewritten as below to reflect both the size (represented by l_1 and l_2 , the wheelbase of the vehicles) and mass effects:

$$r_1 = \frac{k}{l_1^{2.45} + l_2^{2.45}} \left(\frac{m_2}{m_1} \right)^{\alpha/2} \quad (6)$$

Equations (1) and (3) were established through single variable regression with aggregated data.

Kahane [4, 5, 6] and van Auken et. al. [7, 8, 9, 10, 11] used multi-variable regression to model fatality risk with respect to mass and wheelbase. The results of these studies in some aspects are not fully consistent or easily reconciled, due to difference in the field datasets, the details of the regression methods, and the complexity in the regression. Kahane's work [3, 4, 5, 6] (and most part of van Auken's work [7, 8, 9, 10]) studied number of fatalities per “non-culpable” accident exposure; the results cannot be directly compared with Eq. (2) and Eq. (6).

The work presented here is towards further understanding of the dependence of the fatality risk on the vehicle's mass and other vehicle and driver parameters, and analysis of the methodologies and conclusions from the different studies discussed above. It focuses on regression based evaluation of the risk ratio and conditional risk using fatal vehicle-vehicle crashes in the field. The regression is based on disaggregated data and over multiple variables. This study has elements of most of the studies in [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12], but

also has differences from these. Front-front and front-left (i.e., driver-side) crash types are examined separately. In the following, the field data, methodology, and regression results are presented. Discussions are also given based on the effects of mass, vehicle type, and vehicle size on fatality risk and risk ratio. The current results are also compared with findings in the existing studies mentioned above.

DATA AND GENERAL TREND OBSERVATIONS

The field accident data in this study is based on a dataset extracted and publicly released by Kahane [6]. This study only focuses on using the FARS part of the full set of data in [6]. Due to the potential for fatality risk to depend on the parameters differently in front and side crashes, this study examines these two crash modes separately. To facilitate this, supplemental information on the principal impact force clock position for the vehicles was extracted from the FARS database. In addition, the seating position, safety belt use, and airbag deployment condition were also extracted from the FARS database for the occupants of the vehicles contained in the original Kahane dataset.

Kahane's FARS dataset contains cases for vehicles ranging from Model Years 2000-2007, during Calendar Years 2002-2008. From these, cases with Crash Type codes 5 (“hit passenger car of known mass”, 20,930 vehicle records) and 6 (“Hit LTV of known mass and GVWR <10,000 lb”, 16,138 vehicle records) are selected. Out of these, 2,631 “front-front” and 1,747 “front-left” crashes were identified with which information on both vehicles is available. Here, “front” means the principal impact position is between 11 to 1 o'clock, and “left” means 8 to 10 o'clock. For the front-left crashes, the label “bullet” vehicle is used to refer to the vehicle whose front end impacted the left side of the other vehicle which is noted as the “target” vehicle. The front-front and front-side data are presented separately below, and in cases where the two sets show similar characteristics, only the front-front data are shown.

A range of parameters which potentially have an influence on the fatality ratio of the two vehicles were extracted from the dataset. As a preliminary step, the correlation between the driver fatality outcome and these parameters were examined. Figure 1 shows the coefficient of correlation between the driver fatality outcome and these parameters (detailed definition of the variables shown in Figure 1 are given later).

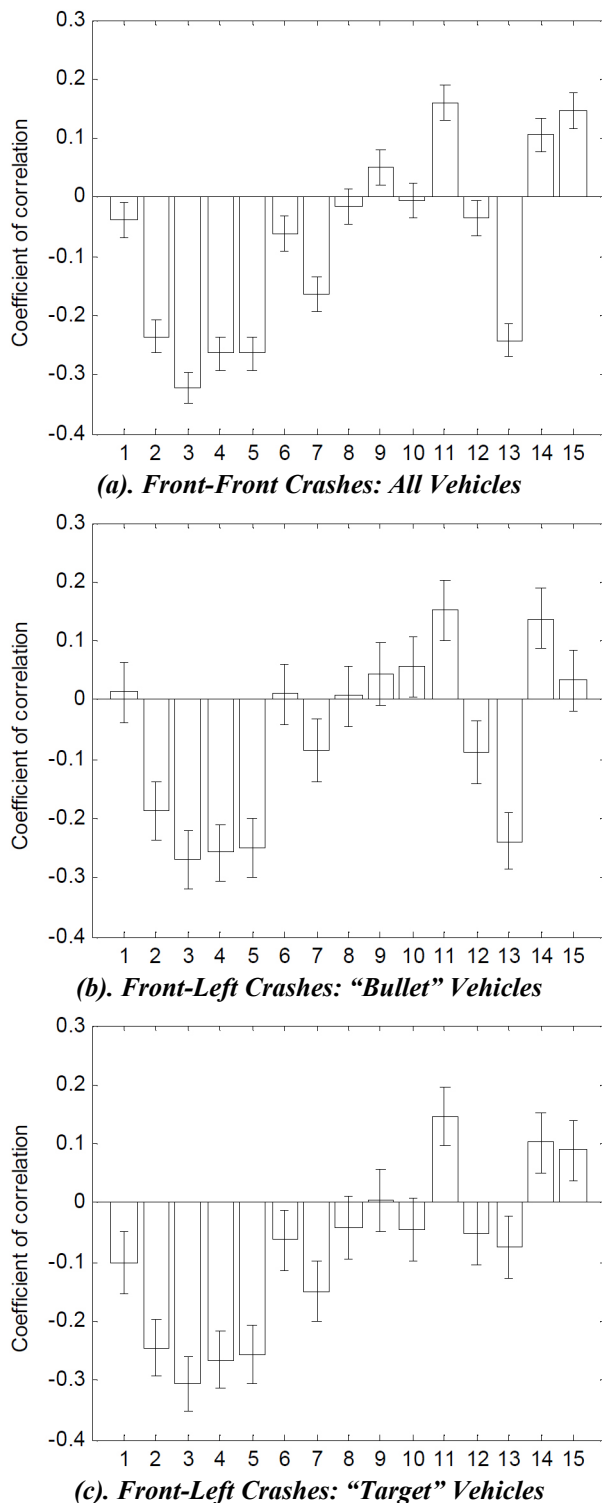


Figure 1. Coefficient of correlation between driver fatal outcome (0 for nonfatal, and 1 for fatal) and some selected variables: 1. Model Year, 2. Truck-SUV-CUV, 3. Mass, 4. Footprint, 5. Track width, 6. ESC, 7. ABS, 8. Torso bag, 9. Combo bag, 10. Curtain bag, 11. Driver age, 12. Driver male, 13. Belt used, 14. Front airbag deployed, 15. Alcohol used (error bars show 95% confidence interval of the coefficient of correlation).

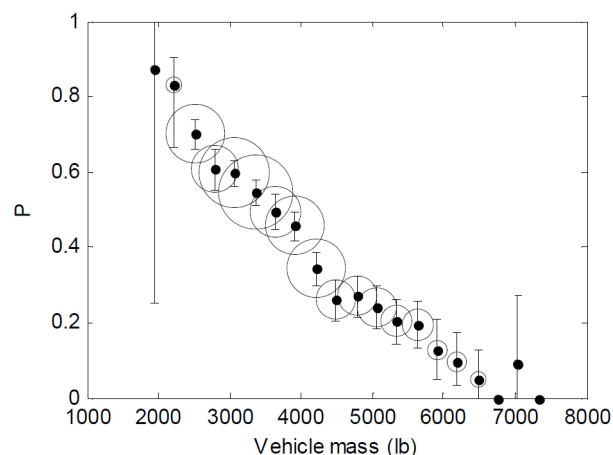
Figure 1 suggests that probability of being fatal is lower for the driver when his/her vehicle is a truck (referring to pickup trucks throughout this paper), SUV, or CUV instead of a car, and when the vehicle has a larger mass, footprint, or trackwidth; this is also the case when the safety belt is used. Conversely, the driver fatality risk is higher when he/she is older, or alcohol was involved. The fatality appears to be positively correlated to the deployment of the front airbag (note that the deployment may be influenced by the severity of the crashes).

Calculating the fatality risk requires not only the fatal accident data, but also non-fatal data. However, no non-fatal census corresponding to the FARS exists. In this study, the fatal rate of the vehicle-vehicle crashes of the extracted datasets, which can be viewed as a conditional risk under the condition that there is at least one fatality involved in the accident (the admission criterion of the FARS), is used to gain some understanding. Noting the fatal risks for the vehicles as r_1 and r_2 , the probability of having at least one of the two drivers being fatal is then $r_1 + r_2 - r_1 r_2$. The conditional probability of Vehicle 1 being fatal is then:

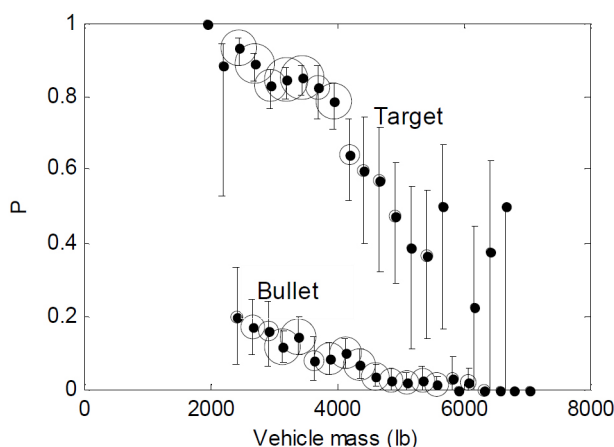
$$p_1 = \frac{r_1}{r_1 + r_2 - r_1 r_2} \quad (7)$$

This conditional risk can be estimated by the proportion of fatal cases in aggregates (or bins) created by dividing the total number of cases per the variable of interest. Figure 2 shows this conditional risk for the drivers per the vehicle mass. Similarly, Figure 3 shows the apparent dependence of the conditional risk on the driver age, separately for belted and unbelted drivers. Figure 4 shows the conditional risk for different vehicle types for front-front crashes. The front-left data show similar information that the trucks have the lowest conditional risk, and SUV, CUV, and minivan generally have values that are between the trucks and passenger cars (for detailed definitions of 'truck', 'SUV' and 'CUV', see [6]). Figure 5 shows the conditional risk for different model year vehicles for the front-front crashes.

Apparent correlation with individual independent variables does not necessarily provide an accurate information of the dependence, especially when there are correlations between the independent variables. Figure 6 shows the coefficient of correlation between the vehicle mass and the other parameters in Figure 1. The relatively high correlation between the vehicle mass and footprint and trackwidth are clearly seen, consistent with data in other studies [6,12].



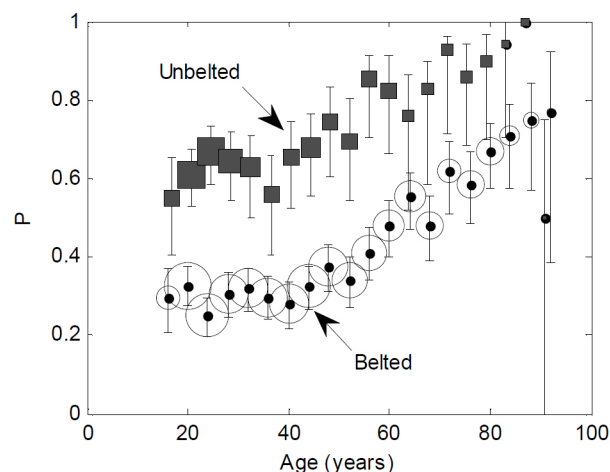
(a). Front-Front Crashes: All vehicles



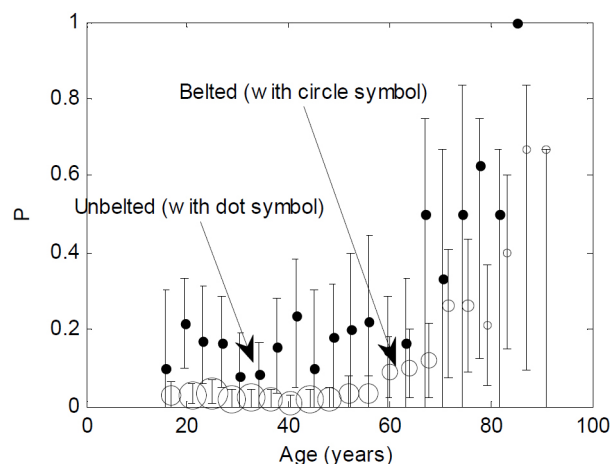
(b). Front-Left Crashes: "Bullet" and "Target" Separated

Figure 2. The average conditional risk (P) v.s. vehicle mass over 20 equally spaced bins. The area of each circle is proportional to the number of vehicles in the bin. The error bar is 95% confidence bounds obtained through bootstrap sampling of the cases in the bin.

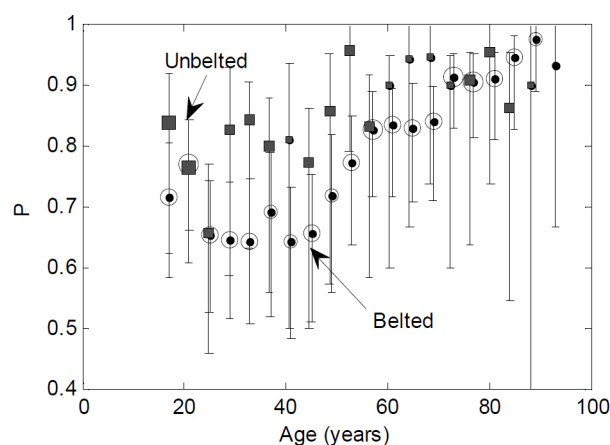
Figure 5, and to some degree Figure 1, suggest a possible correlation between fatality risk and the Model Year of the vehicle. Most likely, any Model Year correlation will be a surrogate to other more fundamental vehicle parameters. For example in Figure 7, some correlation is seen with newer technologies, such as ESC and curtain bag, which were increasingly installed in the model year range in this study.



(a). Front-Front Crashes: All Vehicles



(b). Front-Left Crashes: "Bullet" vehicles



(c). Front-Left Crashes: "Target" Vehicles

Figure 3. The average conditional risk (P) v.s. driver age over 20 equally spaced bins in age. The area of each circle is proportional to the number of vehicles (drivers) in the bin. The error bar is 95% confidence bounds obtained through bootstrap sampling of the cases in the bin.

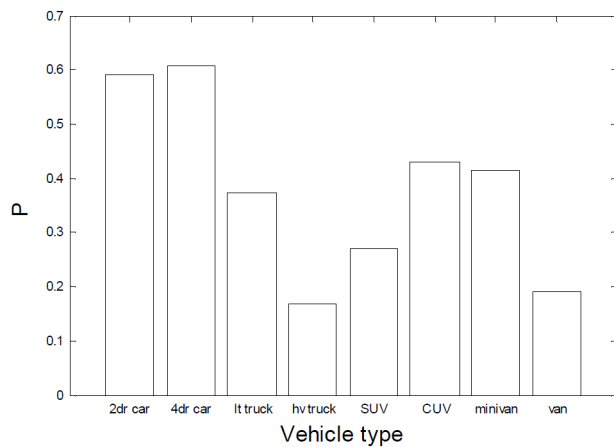


Figure 4. Average conditional risk calculated for different types of vehicles (front-front crashes).

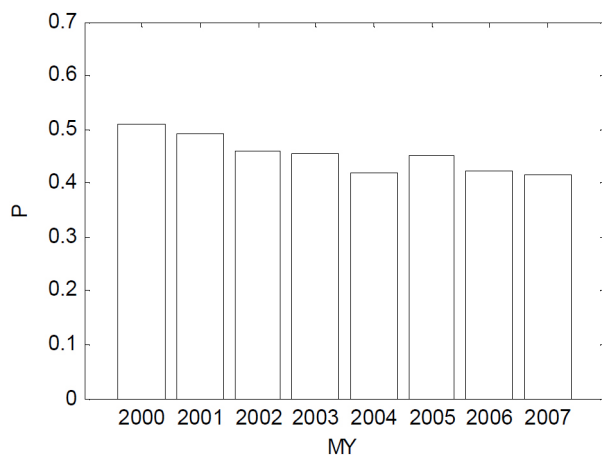


Figure 5. Average conditional risk calculated per vehicle Model Year (front-front crashes).

The relationship between the risk ratio and mass ratio (ρ) is examined similarly to [1]. The crashes are binned based on mass ratio, and in each bin, the ratio of the total numbers of fatalities of the two vehicles is used to estimate the ratio of the fatality risks. One way to examine the data is to calculate the ratio of the heavier vehicle over the lighter one in the crash, i.e., maintaining the mass ratio to be greater than (or equal to) 1.0. Figure 8 plots the data with this vehicle order assignment for the front-front cases. Two different schemes of aggregating the data into bins were used. The first (Fig. 8(a)) was keeping the number of crashes in each bin the same; and the second (Fig. 8(b)) maintaining an equal width in mass ratio for the bins. Figure 9 displays the data in a similar manner, but with the order of the two vehicles randomly assigned.

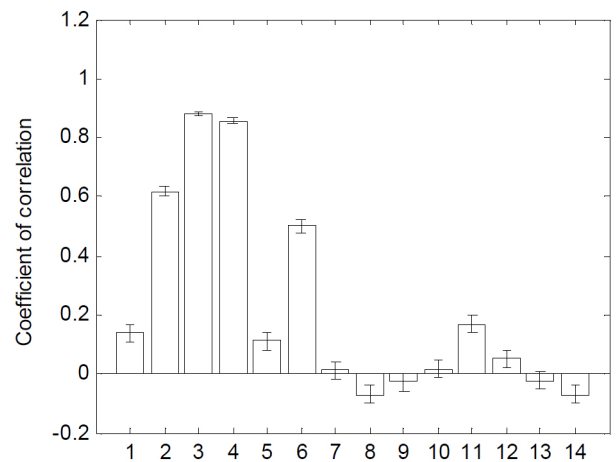


Figure 6. Coefficient of correlation between vehicle mass and selected variables: 1. Model Year, 2. Truck-SUV-CUV, 3. Footprint, 4. Track width, 5. ESC, 6. ABS, 7. Torso bag, 8. Combo bag, 9. Curtain bag, 10. Driver age, 11. Driver male, 12. Belt used, 13. Front airbag deployed, 14. Alcohol used.

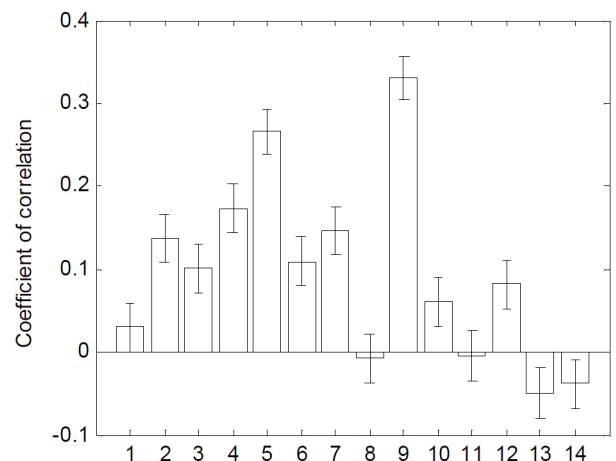
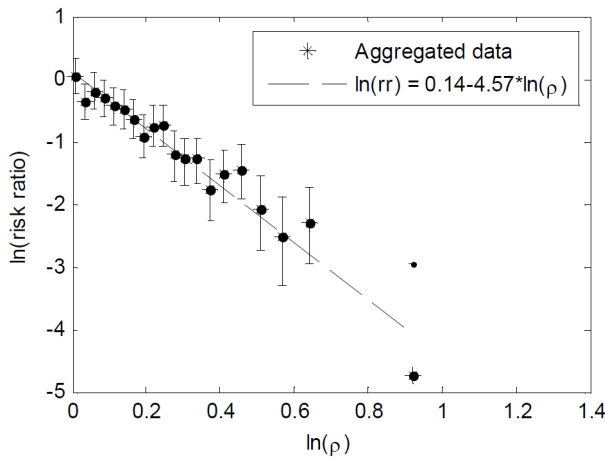
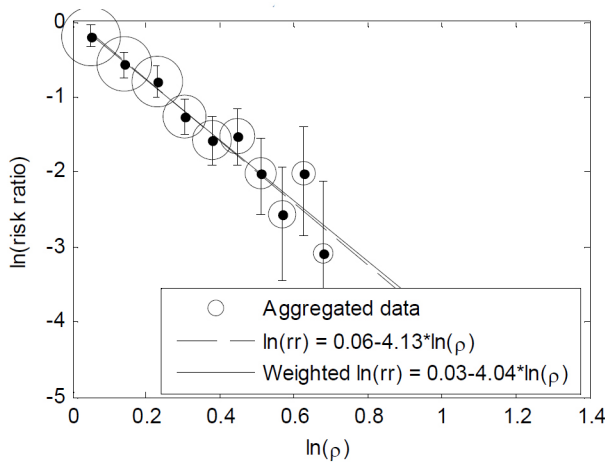


Figure 7. Coefficient of correlation between MY and selected other variables: 1. Truck-SUV-CUV, 2. Curb weight, 3. Footprint, 4. Track Width, 5. ESC, 6. ABS, 7. Torso bag, 8. Combo side bag, 9. Curtain bag, 10. Driver age, 11. Driver Male, 12. Belted, 13. Frontal airbag deployed, 14. Alcohol use

For the front-left crashes, the risk ratio v.s. the mass ratio was also examined and the results are shown in Figures 10 and 11. In Figure 10, the ratios are calculated with the “bullet” vehicle over the “target” vehicle; whereas in Figure 11, the distinction between “bullet” and “target” is ignored and the vehicle order in calculating the ratios is randomly assigned.

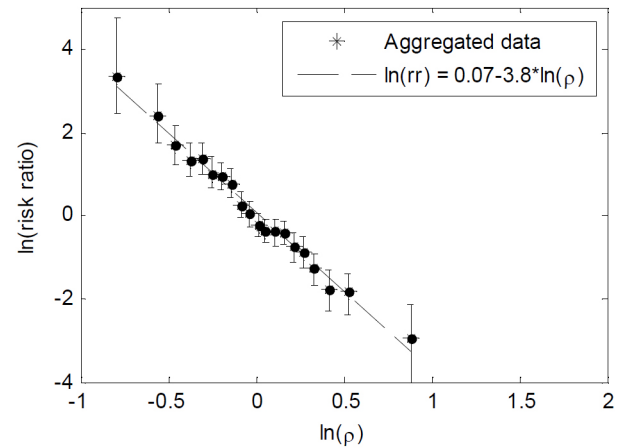


(a). Uniform Bin Size

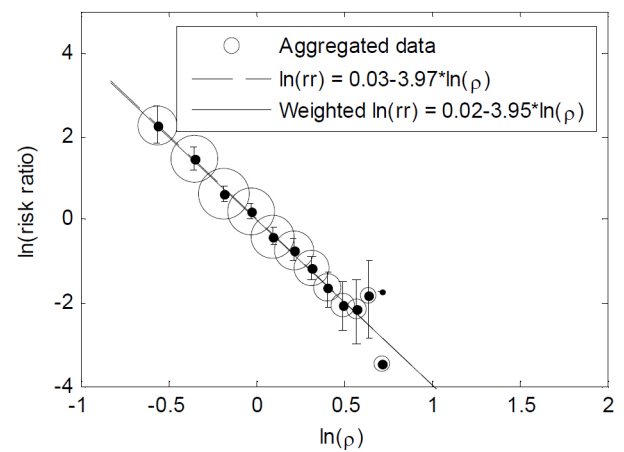


(b). Uniform Bin Width

Figure 8. Front-front crash, aggregated risk ratio v.s. mass ratio. Order of two vehicles assigned to maintain mass ratio greater than or equal to 1. Error bar of the aggregated data points is 95% confidence interval of the risk ratio estimate through bootstrap sampling of the data in the given bin (this also applies to Figures 9, 10, 11).



(a). Uniform Bin Size

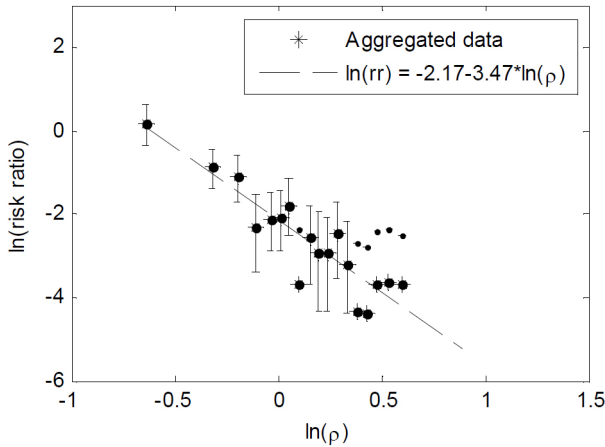


(b). Uniform Bin Width

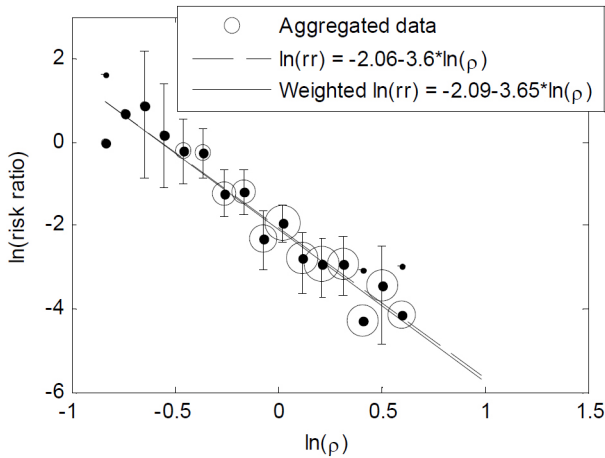
Figure 9. Front-front crash, aggregated risk ratio v.s. mass ratio. Order of two vehicles randomly assigned.

The slopes from the linear regressions in Figures 8 and 9 form a range that generally is consistent with existing data [1, 2, 3] for front-front crashes. A first observation through comparing the front-left case (Figs. 10 and 11) with the front-front case (Figs. 8 and 9) is that a near-linear trend also exists, although there is more scatter, possibly due to the reduced number of data points in each bin. Another observation from Figure 10 is that a significant intercept value of about -2 is indicated by the linear regressions, in contrast to Figure 8. This means that independent of vehicle mass ratio, there is a baseline risk asymmetry in the front-left crashes - the relative risk ratio of the bullet vehicle driver to the target vehicle driver is about $1/8$ ($\sim e^{-2.09}$) to $1/9$ ($\sim e^{-2.06}$).

Figures 8, 9, 10, 11 show an issue with the result from aggregated linear regression: it is not unique, and is influenced by factors that are related to the aggregation setup but external to the field data. The aggregated scheme will also present practical difficulty when a multi-variable regression is desired, because a very large sample size is necessary to create multiple aggregates each having a specific value combination of the independent variables. Therefore, to avoid these complexities and problems, a non-aggregated regression is necessary. In the following, disaggregate logistic regressions are carried out to model the risk ratio.

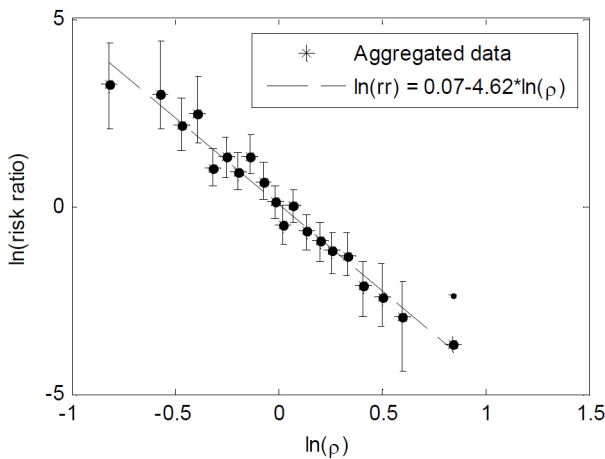


(a). Uniform Bin Size

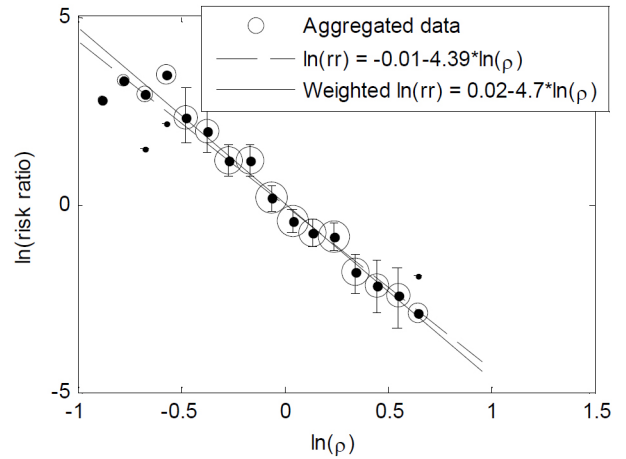


(b). Uniform Bin Width

Figure 10. Front-left crash, aggregated risk ratio v.s. mass ratio. The ratios are bullet (front impacted) over target (left impacted).



(a). Uniform Bin Size



(b). Uniform Bin Width

Figure 11 (cont). Front-left crash, aggregated risk ratio v.s. mass ratio. Order of the two vehicles for the ratios is randomly assigned for calculating the ratios.

MULTI-REGRESSION ANALYSIS OF RISK RATIO

Guided by the linear trend estimated from the data between risk ratio and mass ratio, the following log-linear model between risk ratio and the independent variables (x_{i1} for Vehicle 1 and x_{i2} for Vehicle 2.) is assumed:

$$\ln \frac{r_1}{r_2} = \beta_0 + \sum \beta_i (x_{i1} - x_{i2}) . \quad (8)$$

Eq. (8) includes Eq. (1) as a special case, if it only includes the vehicle mass as the independent variable, and lets $x_{1j} = \ln m_j$.

In this study Equation (8) is turned into a logistic form by introducing:

$$f = \frac{r_1}{r_1 + r_2} . \quad (9)$$

Therefore,

$$\ln \frac{f}{1-f} = \beta_0 + \sum \beta_i (x_{i1} - x_{i2}) . \quad (10)$$

This means that if the risk ratio has log-linear relationship with the independent variables as shown in Eq. (8), then the variable f has a logistic relationship with the difference of the independent variables. From the definition of f , for a logistic regression, the assignment of the binary outcome of f , based on the binary driver fatality outcome of the vehicle-vehicle cases in the FARS database, is as follows: