Risk Analysis, Vol. 5, No. 1, 1985

Do Drivers of Small Cars Take Less Risk in Everyday Driving?

Paul Wasielewski[1] and Leonard Evans[1]

Received August 2, 1983; revised February 17, 1984

Previously reported observed data on risky everyday driving are brought together and reanalyzed in order to focus on the relation between risky driving and the size of the car being driven, as indicated by car mass. The measures of risky driving include separation between vehicles in heavy freeway traffic and speed on a two lane road. Observed seat belt use provides a third measure of driver risk. Confounding effects arising from the observed association between car mass and driver age are taken into account by segmenting the data into three driver age groups. Driver risk taking is found to increase with increasing car mass for each of these three aspects of everyday driving. The implications of these results with respect to driver fatality rates are discussed in terms of a simple model relating observed risky driving to the likelihood of involvement in a severe crash.

KEY WORDS: Risky driving; car size; driver age; automobile accident rates; automobile fatality rates.

1. INTRODUCTION

Recent studies[1,2] focussing on comparing the relative safety of small and large cars have provided a quantitative estimate of the relationship between car mass and the likelihood of a driver fatality, using two different approaches. Both approaches involved "non-two car" crashes; the first[1] used all car crashes except those involving exactly two cars; the second[2] used a slightly more restrictive inclusion criterion, namely single car crashes plus car–truck crashes. These studies, in common with the present study, sought relations involving car mass. However, such relations do not imply that mass, as such, is the causative factor. This variable is chosen because it is a clearly defined and measured vehicle attribute; it correlates strongly with other vehicle attributes, such as wheelbase, track, "size" in general, hood length, trunk size, engine displacement, etc. No mechanisms involving specific physical attributes of vehicles are identified or discussed in connection with the results reported here.

In the first approach,[1] the ratio of car drivers killed per registered car was examined as a function of car mass, based on driver fatalities other than those involving exactly two passenger cars. The number of drivers killed per registered car was found to be 70% higher in a 900 kg car than in an 1800 kg car. It was noted[1] that this procedure yields results that reflect both attributes of the car and possible attributes of driver performance that may affect the likelihood of being involved in a serious crash.

In the second approach,[2] the relationship between driver fatality likelihood and car mass was estimated using the ratios of the frequency of occurrence of different types of fatal crashes involving cars of specified mass categories. For the numerators of these ratios, the crash types were single car and car–truck crashes in which the car driver was killed. For the denominators, the crash types were car–pedestrian and car–motorcycle crashes in which the pedestrian or motorcyclist was killed. The likelihood for a motorcyclist or a pedestrian to be killed in

[1] Transportation Research Department, General Motors Research Laboratories, Warren, Michigan 48090.

0272-4332/85/0300-0025$04.50/1

such crashes was considered to be essentially independent of car mass. Hence the denominators were taken as a measure of exposure to serious crashes in general by the cars of a particular mass category. This procedure was interpreted to yield driver fatality likelihood estimates dependent only on the physical attributes of the car. It was concluded that the driver of a 900 kg car is 160% more likely to be killed, given a serious crash, than is the driver of an 1800 kg car.

The lesser 70% effect found in the approach based on fatalities per registered car[1] was interpreted to reflect more cautious behavior on the part of drivers of smaller cars that is partly neutralizing the estimated 160% effect that was attributed to the physical characteristics of the car using the approach based on fatality ratios.[2]

The conjecture of more cautious behavior on the part of drivers of smaller cars was suggested primarily by the difference between the 70% result and the 160% result. Here we present a reanalysis and integration of previously reported observational data from two separate studies, that directly associates more cautious behavior with drivers of smaller cars, using three different variables to indicate risk taking.

1. Speed
2. Headway (time interval between successive vehicles in a lane)
3. Seat belt use.

In previously reported analyses of these data,[3,4] car mass was included in the analysis as one of a number of explanatory variables but was not singled out for particular attention. A nonmonotonic dependence of both headway and speed on car mass was observed in univariate analyses, and car mass was therefore treated as a categorical rather than a numerical variable, by dividing the data into three mass categories. It was noted, however, in the case of speeds that a monotonic increase of speed with mass resulted when confounding effects from other variables were removed through multivariate analysis.

Correlations between headways or speeds and driver-vehicle characteristics are found to be relatively small, although statistically significant. This is expected, since much of the variance in these quantities undoubtedly results from fluctuating traffic conditions that have nothing to do with driver attitudes. It is therefore not expected that observations of high speed or short headway would have high predictive value for possible crash involvement by individual drivers, but rather that average values of risk measure observations for large groups of drivers, classified for instance by car mass or driver age, may be predictive of the average crash involvement rate of the group.

In this study, the earlier data relevant to risky driving versus car mass[3, 4] are integrated and reanalyzed, using the same analytical framework as in the fatality studies, [1, 2] thus allowing a direct comparison between the relationship of car mass to risky driving and to fatality rates.

2. THE DATA

The data used in this study were obtained in earlier studies relating everyday risky driving to various car and driver characteristics. In one study,[3] short headways in heavy freeway traffic were taken as an instance of risky driving. In the other,[4] speeds of isolated cars on a two lane suburban arterial road were used. All data consist of photographs of oncoming cars and a measurement of the car's headway or speed. From the photographs, the driver's seat belt use was observed, and the license plate was read. From the license number, information on the car, and on the owner (usually the driver) was obtained, as available, from Michigan State files. No information was received from the State that would allow the identification of individuals by name or address. The data are summarized in Table I. The age categories are the same as those used in earlier fatality studies,[1, 2] namely A_1: 16–24 years, A_2: 25–34 years, A_3: ≥ 35 years.

The data with unknown driver age result both from cars for which vehicle information but not owner information was available from the State files (71% of these data) and also from the exclusion of observations where the photograph of the driver was judged incompatible with owner age or sex (the remaining 29%). The exclusion of drivers not driving their own car and the inclusion of some (presumably small) number of misidentified drivers is not expected to have an important effect on the results, since there is no reason to believe that the relation

Table I. Data Summary

Risk measure	Number of observations: Age known		Age not known
	A_1	Total	
Speed			2810
Headway			3018

between car mass and driving risk would be different for the excluded or misidentified drivers than for those included.

3. RESULTS

The previous analyses using these data[3, 4] addressed risk taking versus car mass. No simple effects systematic in car mass were found for either speeds or headways, when considering car mass independent of other variables. A more detailed examination shows that such effects are indeed present in the data, but they were not found by univariate analyses because of confounding between effects of age and effect of mass. Specifically, younger drivers, who tend to take more risk, also tend to use lighter cars. This correlation between driver age and car mass conceals the relationship of car mass to speed and headway in the univariate analysis when all the data are included without regard for driver age. When the effect of driver age is included, as was done for speeds[4] and is done below for both speeds and headways, similar systematic effects are obtained for both variables. Here we take account of driver age by segmenting the data, as in the fatality studies,[1, 2] into the three age groups defined above.

In analogy with the corresponding relations found between fatalities and car mass,[1, 2] let us assume that a measure of risk taking, R, obtained from speed, headway, or seat belt use data, is related to car mass, m, and driver age group, y, by

$$R(m, y) = a_R(y)\exp[b_R(y)m] \qquad (1)$$

where m is taken as a numerical variable and y as a categorical variable with three categories, corresponding to the three driver age groups. The constants $a_R(y)$ and $b_R(y)$ are evaluated by a linear regression of the logarithm of the risk measure on the car mass for the data in each driver age group, as was done for corresponding expressions in the fatality studies.[1, 2]

The range of variation of R with respect to mass is sufficiently small that essentially the same results would be obtained with a linear expression for R instead of the exponential. Also, a linear variation with driver age could be assumed, as in Refs. 3 and 4, rather than using a categorical age variable as in Eq. (1). This would lead to the alternative formulation

$$R(m, y) = c_1 + c_2 m + c_3 y \qquad (2)$$

where both m and y are numerical. Trial calculations show that Eq. (2) gives a slightly better fit to the risk data than Eq. (1), but the resulting slopes of R with respect to mass obtained from the two equations do not differ statistically.

Equations 1 and 2 can also be generalized by including other explanatory variables in addition to driver age and car mass. Trial multivariate calculations using the full set of explanatory variables found significant in the original studies[3, 4] indicate that essentially the same mass dependence for the risk variables is found in these more complicated calculations as in calculations including only car mass and driver age. This result is not surprising, since driver age was found to have a substantially larger effect on the risk variables than any of the other explanatory variables. These other variables include driver sex, driver history of accidents and violations, seat belt use, car model year, presence of passengers, and frequency of observation in the study.

Hence the results of this study do not strongly depend on the choice of Eq. (1) or Eq. (2), or generalizations of them including other explanatory variables. Eq. (1) is preferred and is applied in the balance of this paper for the sake of simplicity and because it allows easy comparison with the fatality results of Refs. 1 and 2.

3.1. Speed

Average speed is plotted as a function of car mass in Fig. 1 for the three age groups. To produce the plot, the data for each age category were divided into five mass cells with boundaries determined by the requirement that each cell contain an equal number of data (to within one datum). The average cell speed is then plotted against the average cell mass. The positions of the points along the mass axis for each age group thus provide an indication of the distribution of car masses for the age group. Note that the points plotted for Age Group A_1 are shifted towards lower masses compared to A_3, reflecting the use of lighter cars by younger drivers. The data are divided into mass cells in Fig. 1 in order to provide a clear illustration of the relation between speed and mass. The analysis, however, is based on regression on the logarithms of the individual observed data, not on the averages over the mass cells.

A trend for speed to increase with mass is apparent for each age group. A general linear model

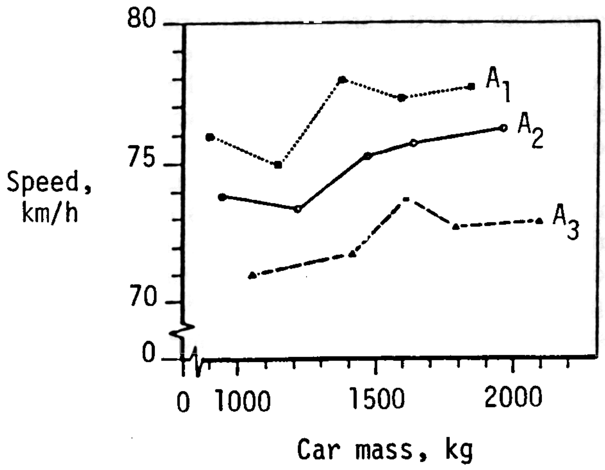

Fig. 1. Average speed versus average car mass for three driver age groups. For each group, the data are divided into five cells with equal numbers of observations in each cell, and the average values of speed and mass for the cells are plotted.

calculation, including an interaction term between age and mass, shows that the constants $b_R(y)$ of Eq. (1) for the three age groups have no significant dependence on age ($p = 0.40$). Hence a common value b_R, independent of y, was fitted to the data by repeating the general linear model calculation without the interaction term. The results are summarized in Table II, which also gives the squared multiple correlation coefficient R^2.

The constant b_R is found to be significantly different from zero at a high level ($p = 0.0004$), confirming that the observed speed tended to increase with the car mass at a similar rate for all three age groups. The differences in the age coefficients a_R are significant at $p < 0.0001$ for both speeds and reciprocal headways.

3.2. Headway

The reciprocal of the observed headway is used as a risk measure, as in Ref. 3, since it increases as the headways become smaller and thus the risk greater. Average reciprocal headway is plotted as a function of car mass in Fig. 2, using age groups as in Fig. 1 and mass cells with equal numbers of data.

The reciprocal headway data exhibit less systematic behavior with respect to mass than do the speed data. The same analysis was applied as for speeds, again showing no significant difference among the values of $b_R(y)$ for the three age groups ($p = 0.53$). In this case, the resulting common value of b_R, although positive, falls just short of differing from zero at the 5% significance level ($p = 0.056$). (See Table II.) It may be noted that fitting Eq. (2) to these data gives a similar slope with respect to mass, but in this case the significance level is higher ($p = 0.03$) because of the better fit to the data provided by Eq. (2).

Although the headway data do not by themselves provide conclusive evidence for an increase in driver risk taking with mass, taken in conjunction with the speed data they do tend to provide additional confirmation of higher risk taking in heavier cars.

For both speed and headway, the driver's risk taking can therefore be expressed as

$$R(m, y) = a_R(y)\exp(b_R m) \qquad (3)$$

because in both cases the coefficient b_R is found to be independent of driver age y.

3.3. Seat Belt Use

Seat belt *nonuse* is plotted versus car mass in Fig. 3 for the data from both the headway and the speed study. The variable nonuse was chosen in preference to use so that all three variables (speed, reciprocal headway, and seat belt nonuse) would increase with increasing risk taking, thus facilitating comparisions between the speed, headway, and seat belt variables. Three separate plots are given in Fig. 3. No confounding effects from driver age were found for seat belt nonuse, so it is appropriate to concentrate on the combined data with no segmenting by driver age, as given in Fig. 3(a). However, as a matter of interest, the data are also given segmented by driver age, as was done in Figs. 1 and 2, in Fig. 3(b) (seat belt nonuse data from speed study) and Fig.

Table II. Parameters of Equation 1

Risk measure	$a_R(A_1)$	$a_R(A_2)$	$a_R(A_3)$	b_R, kg^{-1}	$p(b \neq 0)$	R^2
		72.5	70.1	1.87×10^{-5}		
		0.661	0.588	5.50×10^{-5}		

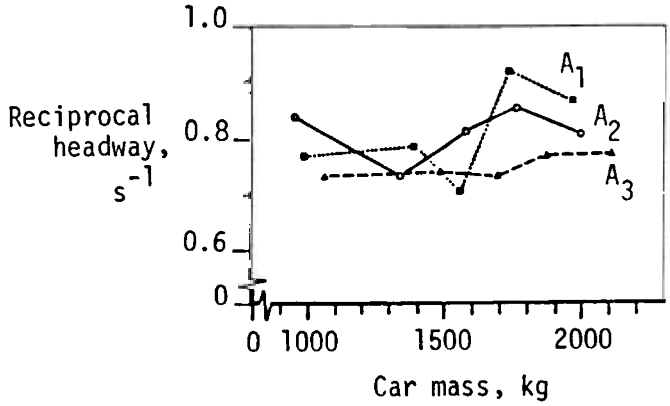

Fig. 2. Average reciprocal headway versus average car mass for three driver age groups. As in Fig. 1, the data for each group are divided into five cells with equal numbers of observations and the average values of reciprocal headway and mass for the cells are plotted.

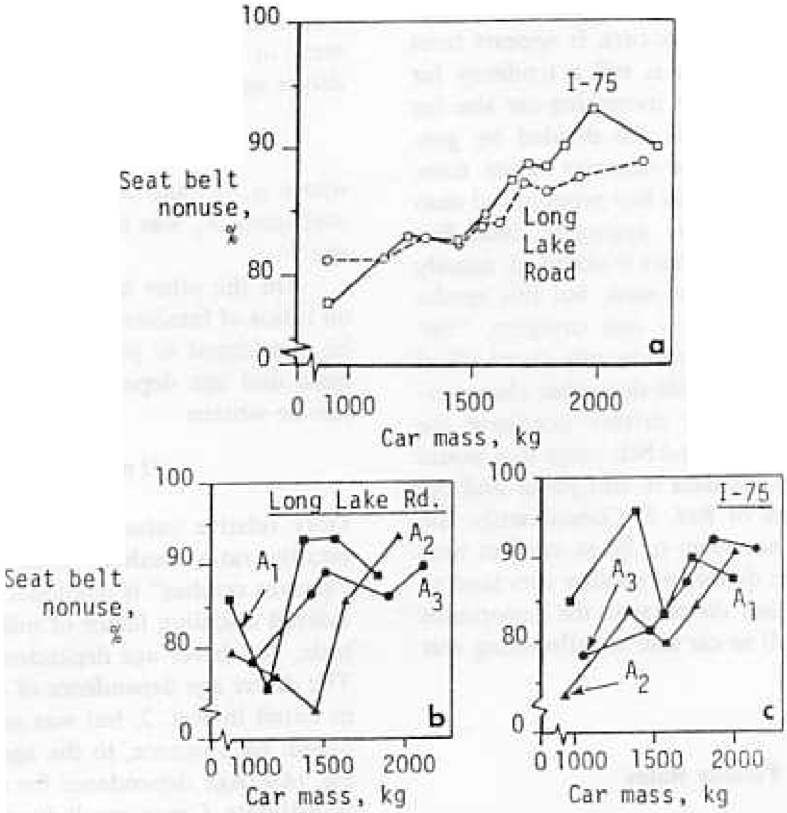

Fig. 3. Seat belt nonuse versus average car mass for the observed data of the speed study (Long Lake Road) and of the headway study (I-75). All the data from both studies are plotted in (a) combining the three age groups and including also the data with unknown driver age. For each data set, the data are divided into ten cells with equal numbers of data in each cell and the proportion of drivers not belted is plotted at the average car mass for the cell. In (b) and (c) the data are segmented by driver age and divided into five mass cells for each age group, as in Figs. 1 and 2.

3(c) (data from headway study). The division into mass cells in Fig. 3 follows the same scheme as in Figs. 1 and 2, except that ten mass cells are used in Fig. 3(a) rather than five because of the greater amount of data available. Note that the data of Fig. 3(a) include not only all the data of Figs. 3(b) and 3(c) but also the data for which driver age was not available. An increasing trend is evident for the data both from the speed study and from the headway study, as confirmed by a log linear regression [that is, fitting the proportions to Eq. (1)]. In each case, b_R differs from zero at the $p < 0.0001$ level, indicating an increase in risk taking with increasing car mass.

In a recent paper,[5] data on belt use versus car size from the National Highway Traffic Safety Administration have been reanalyzed to assess the influence of possible confounding variables. It is found that much of the effect on seat belt use that had been attributed to car size can be explained by higher seat belt use in imported cars and by geographic differences in belt use in domestic cars. It appears from the data of Ref. 5 that there is still a tendency for seat belt use to decrease with increasing car size for domestic cars when the data are divided by geographic regions, although the decrease is not completely uniform and is certainly less pronounced than the decrease observed in the aggregate data. For imported cars, the opposite effect is observed, namely an increase in usage with car mass, but this results entirely from higher usage in one category, "imported compacts," which constitute only about 1% of the car population. It is possible that other characteristics of these cars or their drivers dominate the association between car size and belt usage that would be expected based on the data of this paper and the data on domestic cars of Ref. 5. Consequently, the results of Ref. 5 do not seem to be in conflict with the interpretation that drivers of smaller cars tend to avoid risk, although they do point to the importance of other factors as well as car size in influencing seat belt use.

3.4. Driver Risk and Fatality Rates

In this section, a simple model is introduced in order to show that the reduced driver risk taking in smaller cars demonstrated above can quantitatively account for the differing mass dependence found in the fatality rates of Refs. 1 and 2. This model is admittedly speculative, and other interpretations of the data are doubtless possible. The goal of this section is to illustrate that all the data discussed in this paper can be fit into a consistent pattern based on a plausible assumption about the relation between risky driving and crash rates.

To carry out this program, a parameter in the model corresponding to the strength of the relationship is evaluated from the observed driver age dependence of the risk measures discussed in this study and the fatality rates of Ref. 1, as will be discussed below. Let us introduce the following terminology.

$L =$ likelihood of a driver fatality per "unit of exposure" (such as per registered car)
$g =$ likelihood that a driver is in a serious crash per "unit of exposure"
$f =$ probability of a driver fatality, given that the driver is in a serious crash.

In terms of these quantities, the results of Ref. 1, based on fatalities per registered car, can be considered to provide an estimate of the car mass and driver age dependence of L, which can be written

$$L(m, y) = a_L(y)\exp(b_L m) \qquad (4)$$

where $a_L(y)$ and b_L are regression coefficients. The coefficient b_L was found to be independent of driver age.[1]

On the other hand, the results of Ref. 2, based on ratios of fatalities in different types of crashes, can be considered to provide a similar estimate for the mass and age dependence of the quantity f, which can be written

$$f(m) = a_f\exp(b_f m) \qquad (5)$$

Only relative values of f can be estimated by the fatality ratio method, since the actual number of "serious crashes" is unknown. Hence a_f can be considered a scaling factor of unknown absolute magnitude. No driver age dependence was found for b_f.[2] The driver age dependence of a_f was not investigated in detail in Ref. 2, but was noted to be small [compared, for instance, to the age dependence of a_L in Eq. (4)]. Age dependence for the fatality ratios used to estimate f may result from physical effects, such as greater frailty for older drivers, or from associations between age and the relative exposure to the accident types in question. For the broad age categories used here, the latter explanation is considered more likely, and hence f is assumed independent of driver age, as indicated by Eq. (5).

In order to estimate g, we assume that the likelihood that a driver is in a serious crash is a power function of his propensity to take risk, R, so that

$$g = g(R) = CR^{\alpha} \quad (6)$$

where C is a scaling factor and α is a constant, anticipated to be much larger than one, representing the strength of the association. A small percentage change in a risk measure R corresponds to a percent change in the likelihood of a fatality that is α times as large. By substituting the expression for R from Eq. (3), we obtain an estimate of the age and mass dependence of g,

$$g(m, y) = C[a_R(y)]^{\alpha}\exp(\alpha b_R m) \quad (7)$$

From the definitions of L, g, and f, we have

$$L = fg \quad (8)$$

A direct estimate of L is given by Eq. (4). An alternative estimate of the same function can be obtained by substituting Eqs. (5) and (7) into Eq. (8) to give

$$L(m, y) = Ca_f[a_R(y)]^{\alpha}\exp[(\alpha b_R + b_f)m] \quad (9)$$

Note that this expression depends on driver age only through the coefficient $a_R(y)$. It contains two unknown quantities, the scaling factor Ca_f and the parameter α.

The parameter α is evaluated by comparing the age dependence of these two alternative expressions for L. Since there are three age groups and two unknown quantities, this could be done by a linear regression procedure after combining Eq. (9) and Eq. (4) and taking logarithms. However, in view of the illustrative nature of this computation, we prefer a simpler approach using only the extreme age categories, A_1, and A_3. From Eq. (9), we find

$$\frac{L(A_1)}{L(A_3)} \quad \left(\frac{a_R(A_1)}{a_R(A_3)}\right)^{\alpha} = \frac{a_L(A_1)}{a_L(A_3)} \quad (10)$$

The ratio $a_L(A_1)/a_L(A_3) = 3.9$ is evaluated by fitting Eq. (4) to the driver fatality data of Ref. 1, using the average value of b_L given in Ref. 1. That is, based on the analysis in Ref. 1, data from the U.S. Fatal Accident Reporting System and national car registration data indicate that a 16–24 year old driver is 3.9 times as likely to be involved in a non-two car fatal crash as is a driver older than 35 years, for both drivers in cars of the same mass. Equating these two independently estimated ratios yields

$$\left(\frac{a_R(A_1)}{a_R(A_3)}\right)^{\alpha} = 3.9$$

Substituting values of $a_R(A_1)$ and $a_R(A_3)$ from Table II gives $\alpha = 22.8$ for speeds; and $\alpha = 16.0$ for reciprocal headways. That is, an increase in observed speeds of 1% is associated with a 23% increase in fatalities, and an increase in reciprocal headways of 1% with a 16% increase in fatalities.

The particular value of α obtained is, of course, a function of the risk measure used. Had speeds been measured, say, at a site with more speed variance or a different average speed, a very different value of α might have resulted. The question of interest here is not the value of α but rather whether the association observed between risk and fatality rates as a function of driver age is consistent with the data on driver fatalities versus car mass, as discussed above. In considering the large values of α obtained in this study, note, for instance, that a 1% change in absolute speed (0.74 km/h) represents about 10% of the standard deviation of observed speeds at the data aquisition site. Thus the relative change in absolute speed is not necessarily the most appropriate measure to use in subjectively assessing the risk associated with speed difference (or similarly for headway differences), and undue emphasis should not be placed on the large value found for α.

The mass dependence of the fatality likelihood as given in Eq. (9) may be compared with the corresponding estimate of mass dependence given by Eq. (4), based on fatalities per registered car.[1] Under the assumptions introduced here, it is expected that Eq. (4) and Eq. (9) both estimate the same mass dependence, and thus that

$$b_L = \alpha b_R + b_f$$

is subject to statistical uncertainty in the various estimates.

For comparison of the results of this study with those of Refs. 1 and 2, it is convenient to rewrite Eq. (12) as

$$\alpha b_R = b_L - b_f$$

The left hand side of this equation is based on the everyday risky driving data of this study and the driver age dependence of the fatality data, whereas the right hand side is based solely on the two fatality studies, from which we have $b_K = -0.000\ 58\ \text{kg}^{-1}$ [1] and $b_f = -0.001\ 06\ \text{kg}^{-1}$.[2] Thus, based on results of Refs. 1 and 2, we expect

$$\alpha b_R = 0.000\ 48\ \text{kg}^{-1} \tag{14}$$

Using estimates of α and b_R derived in this study from the speed data, αb_R is estimated as $22.8 \times 1.87 \times 10^{-5}\ \text{kg}^{-1} = 0.000\ 43\ \text{kg}^{-1}$, in satisfactory agreement with Eq. (14). For reciprocal headways, αb_R is estimated as $16.0 \times 5.50 \times 10^{-5}\ \text{kg}^{-1} = 0.000\ 88\ \text{kg}^{-1}$. In view of the statistical uncertainty in the parameters derived from the headway data, this value is not inconsistent with Eq. (14).

The basic conclusion from these calculations is that more cautious everyday driving can account quantitatively for the difference between the physical mass effect on fatality rates estimated in Ref. 2 and the mass effect estimated in Ref. 1, which arises from both the physical effect and driver behavior patterns associated with car mass. The particular relation, namely the power function, assumed between risky driving and fatalities is, of course, speculative, and to the extent that they depend on this relation, the numerical results must be regarded as uncertain. However, it seems plausible that *some* functional relation such as Eq. (7) exists between driver risk and fatality likelihood, and the basic conclusions of this section, if not the specific quantitative details, would be borne out using any such plausible relation.

4. SUMMARY AND CONCLUSIONS

The results of this study can be summarized by graphs of the analytical forms of risk measures examined here as a function of car mass. Such graphs are given in Fig. 4, based on Eq. (3) for the three risk measures. These results, taken together, provide clear evidence that drivers of smaller cars tend to take less risk in everyday driving.

The observed more cautious driving in smaller cars is consistent in magnitude with the difference reported earlier between the mass dependence of the number of fatalities per registered car (reflecting both physical and driver behavior effects)[1] and the estimated dependence of the probability of a fatality given an accident (reflecting only physical effects).[2]

It is not claimed that this study, taken by itself, is sufficient to *prove* that more cautious driving in smaller cars is in fact substantially reducing the occupant fatality rates that might otherwise be associated with these cars. However, the available data are consistent with this interpretation. We feel that the data and analysis presented in this study indicate the need for additional research into the relations between driver behavior, car characteristics and crash rates. In fact, work performed after the completion of this study[6] provides further evidence in support of the relations inferred here and in Ref. 2 between crash rates and car size.

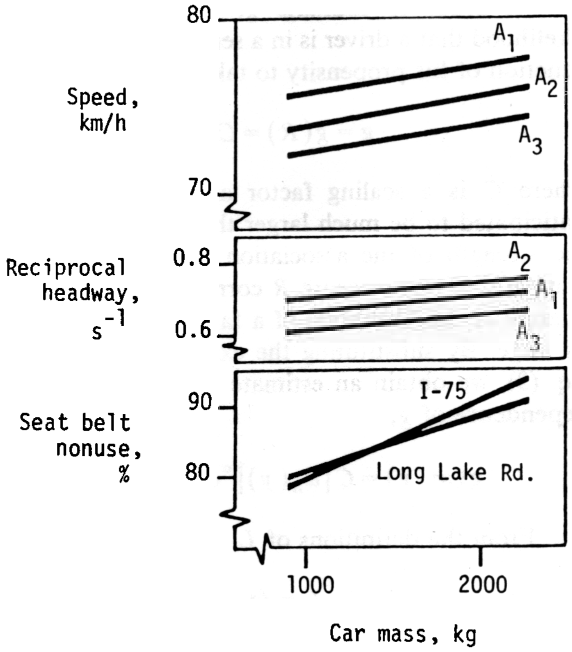

Fig. 4. Summary graphs of analytical relations between observed risk measures and car mass.

REFERENCES

1. L. Evans, "Car Mass and Likelihood of Occupant Fatality," Society of Automotive Engineers Paper #820807, Troy, Michigan (1982).
2. L. Evans, "Driver Fatalities Versus Car Mass Using a New Exposure Approach," *Accident Analysis and Prevention* **16**, 19–36 (1984).
3. L. Evans, and P. Wasielewski, "Risky Driving Related to Driver and Vehicle Characteristics," *Accident Analysis and Prevention* **15**, 121–136 (1983).
4. P. Wasielewski, "Speed as a Measure of Driver Risk: Observed Speeds Versus Driver and Vehicle Characteristics," *Accident Analysis and Prevention* **16**, 89–103 (1984).
5. B. O'Neill, A. F. Williams, and R. S. Karpf, "Passenger Car Size and Driver Seat Belt Use," *American Journal of Public Health* **73**, 588–590 (1983).
6. L. Evans, "Accident Involvement Rate and Car Size," *Accident Analysis and Prevention* **16** (3–4) (1984).