8. FATAL CRASH RATES

8.1 MODELING FATALITY RATES FOR OLDER DRIVERS

Data limitations impeded direct, reliable observation of all accidents, but accidents in which a fatality occurs are reported considerably more thoroughly and consistently than crashes not involving a fatality. Accordingly, our principal measure of fatalities is older driver fatalities, meaning the exact number of older drivers killed in accidents. We also developed another fatality measure which captured the total number of deaths involved in crashes involving an older driver, but its interpretability lacks clarity because of insufficient information on assignment of fault in crashes.

The first of the two fatal crash rate concepts we call "driver risk," or the driver fatality rate. It is the fatality rate, per mile driven, of older *drivers* alone, regardless of any other deaths that occur in a crash involving an older driver. The driver fatality risk rate is measured by dividing the number of older drivers in a given age-gender-region group killed in automobile accidents by the number of miles driven in a year for the particular group of drivers. This rate is essentially the number of annual driver fatalities per mile driven. Because of the VMT magnitude, these rates are presented in terms of number of annual fatalities per 100 million miles driven.

Figure 8.1 shows that almost all of the rates of elderly "driver risk" have been declining over time, with the exception of men between 80 and 84 years old. The biggest declines are found in the groups with the highest historical risk – those persons over the age of 85. Figure 8.2 illustrates that those elderly persons in the South generally have a higher risk than those in other regions of the country.

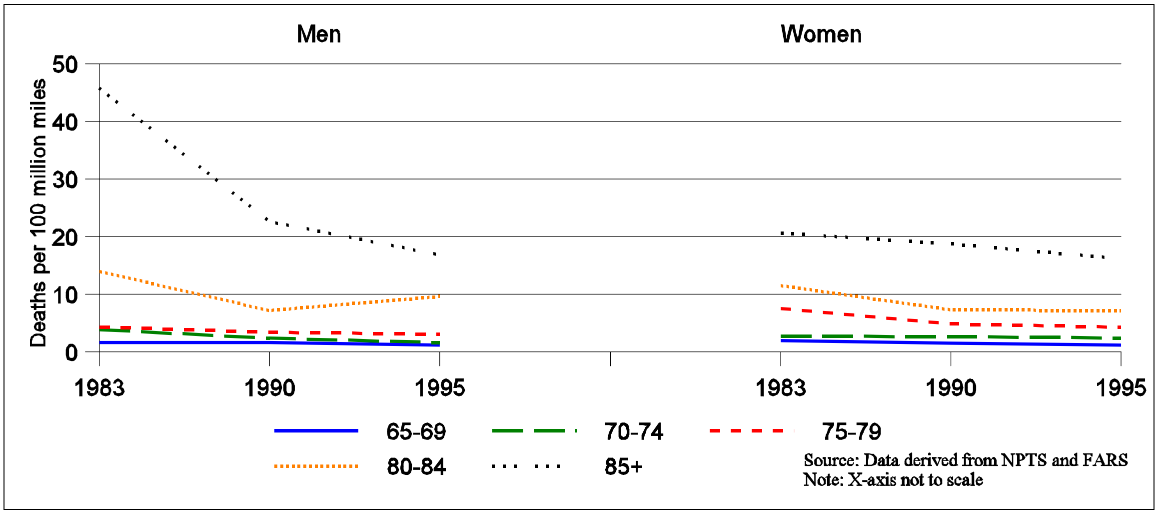

Figure 8.1. Historical Elderly Driver Fatality Rate By Age
(deaths per 100 million miles)

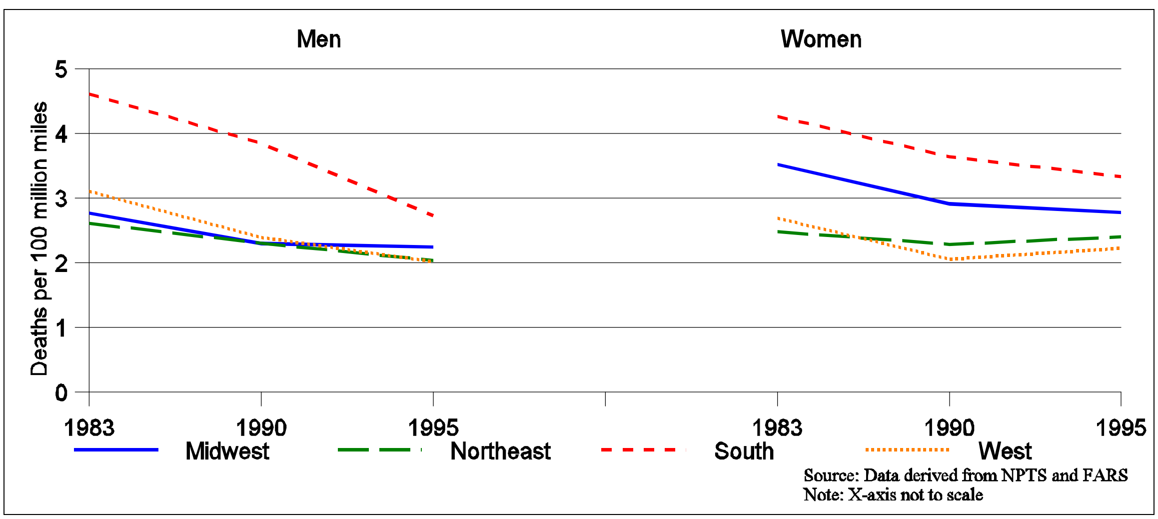

Figure 8.2. Historical Elderly Driver Fatality Rate By Region
(deaths per 100 million miles)

The other fatality risk concept we project is the "total fatality crash rate," which is the total number of fatalities, regardless of age or occupant status, associated with a crash in which at least one older driver is involved. The rate of total fatality risk is measured in a more complex fashion than the driver crash rate. Each driver in a fatal crash which involves an older driver is assigned equal responsibility for each fatality. Then, the sum of the fatalities attributed to each driver is found for a driver's age-gender-region group. For example, if a crash involving three fatalities occurs between a man aged 65 and a woman aged 72, 1.5 fatalities would be attributed to the 65-69 age group for men, and 1.5 fatalities would be attributed to the 70-74 age group for women, both in the respective region of the accident. If a two-vehicle crash involving a 67 year old woman and a 45 year old man resulted in four deaths, two of them would be attributed to the 65-69 female age group, and two would be excluded from the measure as attributable to an age group outside the study. In assigning responsibility in general, an individual older driver's contribution to age/gender/region total equals the number of fatalities divided by the number of drivers involved. The total number of fatalities attributed to each age-gender-region group is then divided by the annual number of aggregate miles driven in that group, in similar fashion to the driver fatality risk measure. Total fatality risk is also presented in terms of number of annual fatalities per 100 million miles driven. Historical trends in our measure of "total risk" mirrored trends in the "driver risk" measure.

The casualty data from the 1983, 1990, and 1995 FARS were used. FARS reports information on fatal crashes by state, disaggregated by age in one-year increments and by gender. We aggregated the individuals to 5-year age groups and the states to Census regions. Corresponding variables on income, health status, VMT, etc., were created as age/gender/region averages from NPTS and NHIS data.

The availability of automobile seat belts, beginning in the 1960s and growing relatively slowly until well into the 1970s, represents one independent variable used in this equation. Seat belt use can be viewed as an indicator of technology that shifts the degree of safety available, given income and an array of related prices. This effect on safety, as proxied by the seatbelt usage variable, was the same for all people at any given time. The income effect on safety choices had a differential effect across individuals both at any point in time, as individuals (actually groups in the aggregation required for this regression) with higher incomes purchased vehicles that provide greater safety at a given time, and over time, as the general level of incomes for all people rose, leading them to choose greater levels of safety across the board.

Empirical evidence indicated that regional effects were important as well, so presence in one of the four Census regions was also used as a set of binary variables. Seat belt use and time were highly correlated, as is to be expected with the market penetration of this innovation, along with the passage of legislation mandating seatbelt use. We experimented with combinations of one of the two variables and the residual of the regression of the other variable on the variable entered in untransformed version. The final version of the model interacted seatbelt use with age,[1] but the residual of the regression of time on seat belt use as a variable was never significant. Consequently, we excluded time from both estimated models. Interacting seatbelt use with age permits the identification of differential effectiveness of seatbelt use in preventing fatalities across age groups.

Thus the full array of independent variables for this equation included income, health status, seat belt use, age, gender, and time. As we explain below, we did not use all of these variables in the final estimation used in the projection model. The smaller sample size available for these models forced several choices between variables that, in a more perfect world, would have been retained in the regression model. We estimated this regression for all age groups and both genders, but we accounted for individuals' 5-year age groups and gender with binary variables instead of estimating separate regressions. The effect of seat belt use was virtually identical between men and women, so the gender-seat belt interaction was

[1] That is, a separate variable was created for seatbelt use for each of our 5-year age groups: e.g., a variable for seatbelt use among 65-69 year olds, another for 70-74 year olds, etc.

dropped. Since these measures have asymptotic minima at zero (in practice, probably well above zero: the fatal crash rates cannot fall below zero and are unlikely in practice to reach zero), we again chose the logistic specification for the regression equations.

Income had a highly statistically significant relationship with the fatality rate (significant at 1%), indicating that people with higher income were purchasing more safety. One such route is in the form of newer vehicles, which are more likely to have the most current safety technology incorporated into them. Other routes, such as more careful driving and ability to drive at less dangerous times of the day, are more open to debate and should be subjected to more direct, empirical examination. Income also affects the fatality rate in that individuals with higher incomes are generally healthier to begin with, and may receive better health care after a crash. This point is expanded upon later in the report.

The *health status* variable, either as a direct variable or as a residual from the regression of health status on income, had a positive relationship with the older driver fatality rate, and it substantially reduced the significance of income. The expectation for that variable was that it represented a measure of capacity for driving and should have a negative relationship to the crash rate. The only apparent explanation for a positive relationship is that people with greater physical limitations drive more carefully and actually overcompensate for their disabilities. While this explanation could be the case, it seemed like any extraordinarily conscientious driving could easily be mitigated by slower reaction time. The *health status* variable has been rather problematic, in terms of the data within the NHIS being able to explain only small proportions of the variance in that measure, through its transfer via model coefficients from the NHIS to the NPTS, and finally to its aggregation to Census region averages with the values calculated from the NPTS. Thus, we prefer to suggest that we simply do not understand the negative health coefficient. The crash rates projected with and without the health status variable are not vastly different, so we preferred to use the model without health status for the projections.

We did not include employment status in the crash rate models because it was not expected to exert an influence on crash *rates*, or fatal crashes per mile driven. Of course, being in the labor force is likely to put a person on the road more than not being employed, but that should have affected the total number of fatal crashes, not the rate. Accordingly, the employment status effect exerted its influence on total fatal crashes through its effect on VMT.

The final logistic regression specification is as follows:

Prob (driving a mile with an older driver fatality) = $(1+e^{-Z})^{-1}$,

where Z= constant + a_1 log (*income*) + a_2 (*age*) + a_3 (*seat belt use*age*) + a_4 (*region*).

(3)

The regression equation used as independent variables *income*, the categorical 5-year *age* variable; *seat belt use* interacted with *age* (meaning that the effect of seat belt use varies with age), and Census *region* dummy variables. The regression also included dummy variables to account for specific age/gender/region groups that did not adhere to general age, gender, or region effects. We did not estimate separate regressions for age, gender, and region because of the small sample size available for this model—120 observations in contrast to several thousand for the VMT and driver models.

Table 8.1 reports the regression model for the elder driver crash rate. All variables but two region dummy variables and one region/gender/age dummy variable indicated strong statistical relationships. The *income* coefficient of the driver crash rate was negative, as theory leads us to expect, and was of substantial size at 0.46. As in the logistic regressions modeling the probability of continuing to drive, this 0.46 means that for each additional one unit of log*(income)*, the logit function, Z in equation (3), increases by 0.46. The *seatbelt* effectiveness variables also had the anticipated effect of decreasing the fatality rate, but do not have a regular pattern over the age groups. The effectiveness of the protection given by seatbelt usage was greatest for people in the 85+ age group, as might be expected, with a coefficient of -1.36. This coefficient of -1.36 means that, for those 85 and older *only*, each one unit increase in the seatbelt rate will decrease the logit function by 1.36. The coefficients for the 65-69, 75-79 were virtually the same size, -0.63, and that for the 80-84 group is quite close in magnitude, -0.57. The effect for the 70-74 group was intermediate in magnitude at -1.01. The *age* dummy variables indicated that as age increases, so does the risk of being involved in a fatal crash. Regional variables showed no relationship between the Midwest and Northeast regions and the fatality rate, but the fatality rate in the southern Census region was somewhat higher than in other regions. The southern regional dummy variable added 0.30 to the overall intercept term of -10.01, not an especially large effect but a statistically significant one. Several other interacted dummy variables also captured differential fatality rates: the rate for western males in the 85+ age group rose over time, 85+ women in the South and West had a slightly lower fatality rate, and 80-84 men in the South a slightly higher rate. Women in the 80-84 age group in the Northeast had a small, statistically weak, negative differential in their fatality rates.

Table 8.1. Driver Crash Rate Regression

	ß	Prob ?^{2}
Intercept	-10.0123	0.0001
log (income)	-0.4620	0.0001
Sb*age65	-0.6255	0.0001
Sb*age70	-1.0080	0.0001
Sb*age75	-0.6279	0.0001
Sb*age80	-0.5729	0.0001
Sb*age85	-1.3596	0.0001
Age65	-3.0419	0.0001
Age70	-2.4238	0.0001
Age75	-2.1552	0.0001
Age80	-1.4633	0.0001
Midwest	-0.0305	0.3974
Northeast	-0.0601	0.1148
South	0.3005	0.0001
W 85 M * time	-0.0303	0.0023
NE 80 F	-0.1714	0.1567
S 85 F	-0.4718	0.0001
W 85 F	-0.5550	0.0003
S 80 M	0.4098	0.0001

Table 8.2 reports the regression results for the model of the total fatal crash rate. The total fatal crash rate is far more complex than the total driver fatality rate. Thus, while a factor such as income of the driver in the elder driver fatal crash rate is directly relevant, it is not clear whose income would affect the total crash rate: the elder driver, who may or may not have died in the crash, a passenger who died in any of the vehicles involved, or the non-elder driver of another vehicle. The only information we have is the income of the elder driver involved in the crash, and he or she may not be the fatality. Consequently while the income of the elder driver involved in the crash may affect his or her safety, other people are involved, and we do not have the data to model the choices that led them to be on the scene. Thus the appropriateness of the income of the elder driver in one of these crashes is open to question. Not surprisingly, income performed oddly in these regressions, obtaining positive and significant regression coefficients. Our lack of understanding as to why higher income would lead to a higher probability of total fatal crashes in an age/gender population led to our omission of the variable from the regression specification. The *health status* variable also had a positive significant regression coefficient when used in the same regression with income of the elder driver involved. Not understanding why superior *health status* would lead to a higher fatal crash rate, we also omitted that variable from the specification. While higher levels of income and better health status may increase the total number of fatalities through increased VMT and increased probability of driving, there is no directly logical effect of these two variables on the total fatality rate *per mile driven*.

Altogether, it was not clear that variables thought to influence individual choices should be in this regression. Accordingly, the regression model for total fatal crash risk contains only age and region dummy variables, the *seatbelt* variable interacted with age to account for the differential frailty of elderly persons of different ages, and some more specific age/gender/region interacted dummy variables to account for idiosyncratic effects in the aggregate crash statistics. Nevertheless, no differentials on the basis of gender alone were found, so the only gender dummy variables used were those interacted with the occasional age group and region.

Table 8.2. Total Risk Crash Rate Regression

	ß	Prob ?^{2}
Intercept	-14.6373	0.0001
Age65	-2.8306	0.0001
Age70	-2.1929	0.0001
Age75	-2.0658	0.0001
Age80	-1.1724	0.0001
Midwest	-0.1186	0.0002
Northeast	-0.1213	0.0008
South	0.2505	0.0001
Sb*age65	-0.7053	0.0001
Sb*age70	-1.2005	0.0001
Sb*age75	-0.6562	0.0001
Sb*age80	-0.7641	0.0001
Sb*age85	-1.5809	0.0001
NE 80 M	-0.2824	0.0013
S 80 85 F	-0.5416	0.0001
W 85M*Time	-0.0329	0.0018
W 65 F	-0.102	0.1840
W 75 80 F	-0.2289	0.0018
W 85 F	-0.7147	0.0001

The great majority of the variables had an extremely highly significant effect on the total fatal crash rate. The age dummy variables indicate higher fatality rates for this measure in older age groups. As in the elder driver crash risk model, this measured rate was somewhat higher in the southern Census region, with a coefficient on the southern region dummy variable of 0.25 to be added to the general constant term of -14.64. Also as in the case of the driver crash risk, no clear age pattern emerged in the effect of seatbelt use, but the greatest effectiveness of seatbelt use also emerged for the 85+ and 70-74 age groups, as in the driver crash risk regression. And as in that other crash rate, the time trend on 85+ men in the western Census region was negative.

We acknowledge that the use of simple crash per mile measures to characterize the risk facing drivers has been criticized on the grounds, among others, that the crash rate per mile does not appear to be constant for drivers who average substantially different annual mileages (Janke 1991). Janke notes that drivers with low annual VMT tend to have higher crash rates per mile than do drivers with high annual mileage, to a considerable extent because the low-mileage drivers are driving disproportionately on city streets as opposed to expressways. For example, crash rates per mile from California in the 1980s were 2.75 times as high on open-access streets as on expressways. One of the important implications of this empirical finding is that mileage by itself may not be a satisfactory measure of the exposure to crashes. For example, Janke notes that part of the linearly measured crash risk of, say, elder drivers should be attributed to where they drive, and only part of it to their age. Stratification of driver populations according to various criteria is one recommended strategy for reducing this nonlinearity in the mileage-accident relationship.

Our analysis of crash risk begins with a stratification of the elderly as opposed to all age groups and continues the stratification with the interacted age-seatbelt use variable, which permits crash risk to vary within elderly age groups. Data on the predominant use of one type of roadway were not available in observational units compatible with the state-wide FARS data used in the crash rate regressions. Even if such roadway data had been available, projecting the values of those variables (possibly as the percent of driving on one or the other type of roadway) to 2025 would have been a major challenge. Overall, we believe the combined effect of driver age and roadways driven in our measure of age-specific crash risk to be satisfactory: if older drivers tend to drive proportionally more on city streets than do mid-career drivers (ages 35-55) and consequently have higher crash rates per mile, that is an acceptable indicator of the crash rates expected for older drivers, even if some of the differential between their rates and those of younger drivers is attributable to driving location.

8.2 FATAL CRASH RATE PROJECTIONS

Two variants of older driver fatality rates , the older driver fatality rate and the total fatality rate from crashes involving an older driver, are projected with the originally estimated regression equation (3). Since the values both of these dependent variables can take are constrained (neither fatality rate can go below zero), it was necessary to retain in the projection the asymptotic behavior of the dependent variable imposed by the form of the equation. For projections of future years' values of the dependent variables, the projected levels of the independent variables were substituted into the respective logistic equations and the new values of the dependent variables calculated. We did not use the time trend in the projection of older driver fatality rates.

The projection of household income to support projection of both fatal crash rates was identical to its projection for the previous components. The age and region variables were dummy variables, and they simply took values of 1 in each projection year, to be multiplied by the estimated regression coefficients. This is also the case with the interacted dummy variables age/gender/region and West/men. The doubly interacted dummy variable West/male/year was projected as the time trend variable was projected, but only for the group of males in the western Census region. The seatbelt use variable was projected using predicted values of a regression on time of 1991-1996 NHTSA rates and the U.S. DOT's 2025 expected rate of use at 85%, on time. Data were available from 1983-1996, but inconsistent data due to a change in methodology in 1991 required us to use only the 1991-1996 data and 2025 projection in the regression (Appendix B.2.5). The interaction of these projections with age used the same percent use with each age group but multiplied that usage by a separate regression coefficient in the projection.

Finally, the projection of the numbers of fatal crashes required the use of population projections, which were furnished by U.S. Census Bureau projections. The modification of total projected population by the fraction projected to be institutionalized (AHCA, taken from *Statistical Abstract of the United States*) yielded the projected non-institutionalized population. Projections of all other independent variables were embodied in the projections of the other components.

Figure 8.3 illustrates the projected decline in fatal crash rates among elderly drivers. One can observe that the oldest age group (85+) has the highest historical crash rates per 100 million miles driven, with the 1995 national averages being 16.83 deaths per 100 million miles driven for men and 16.23 per 100 million miles driven for women. Additionally, the positive relationship between age and fatal crash risk holds for both genders and all regions. The oldest age groups have also had much sharper rates of historical decline in these rates than some of the younger elderly groups. Our projections reflect this information, with the oldest age groups declining at a much sharper rate (to 66% of their observed 1995 rates in 2025) than the younger groups, which decline to around 85% of their observed 1995 rates in the 2025 projections. These projected national trends are virtually identical between the two genders at the youngest (65-69) and oldest (85+) age groups, but vary to some degree among the middle groups, with men generally declining at a slower rate. Fatal crash rates among drivers were fairly consistent among regions, except for those drivers in the South, who had higher fatality rates across virtually all age groups and both genders. Tables A.3.1-A.3.4 in Appendix A show these age and regional differences in greater detail.

Our total fatality rate from crashes involving an older driver follows similar trends, with the oldest age groups having rates that start higher and drop relatively more dramatically than the youngest elderly age groups. We also see the same positive relationship between age and risk as we do in the fatal crash rate of drivers, as well as the similarity between genders. As with the driver fatal crash rate, southern men have higher absolute rates per 100 million miles than men in other regions. Tables A.5.1-A.5.4 in the Appendix show these differences in greater detail.

Income growth and projected growth in seatbelt use contribute roughly equally to these projected declines in driver risk ranging from around 60%-40% to 50%-50%, depending on age/gender group. Tables 8.3 and 8.4 show these contributions at the national and regional levels. Differences are more substantial across age groups than between genders within any particular age group, primarily because the regression coefficient on the

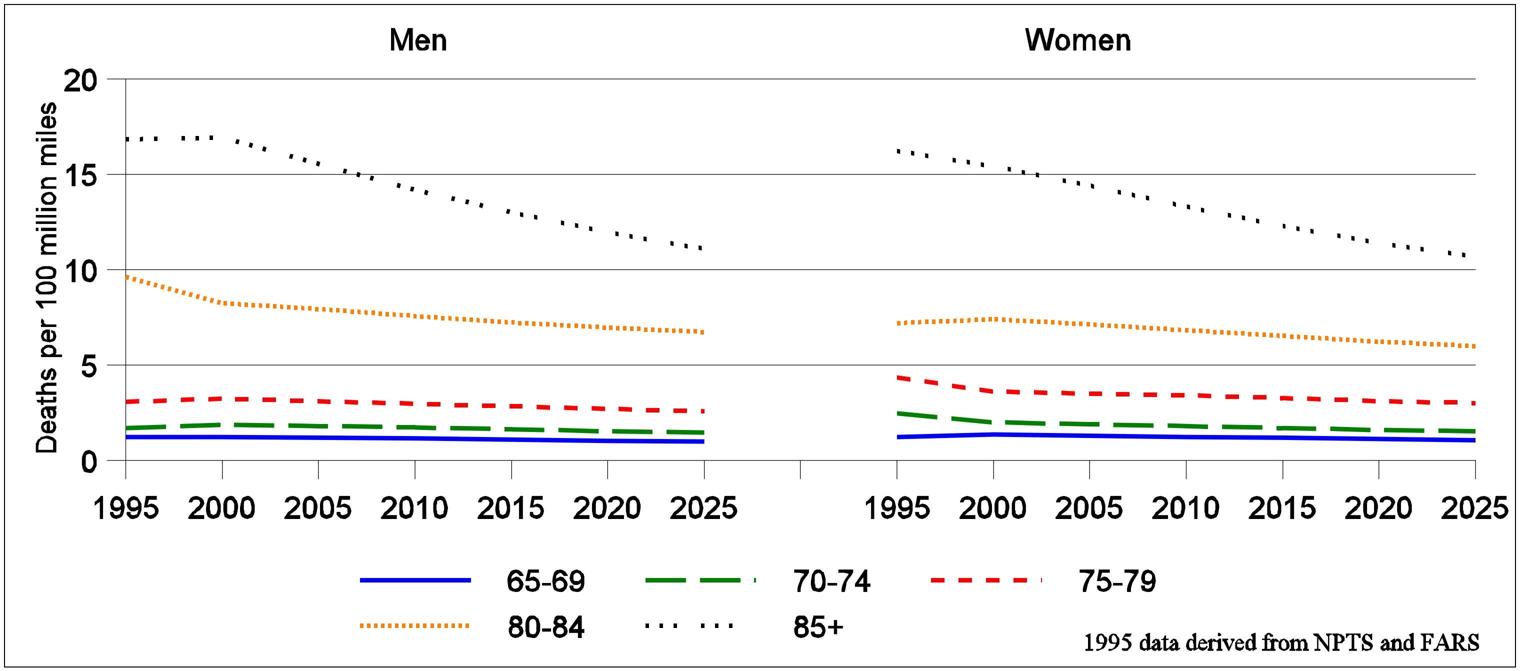

Figure 8.3. Projected Driver Risk By Age (deaths per 100 million miles)

seatbelt use variable was invariant across gender. There is no smooth pattern of increase or decrease in the contribution of either variable as we move up the age groups. Among men, seatbelt use contributes the most among the 70-74 and 85+ groups, 60% to 63%, and the least to decreases in driver crash risk among 65-69, 75-79, and 80-84 year olds, from 40% to 42%. Growth in seatbelt use contributes the most to decreases in women's driver crash risk among the 65-69 year olds, around 37%, the least among the 80-84 group, and at an intermediate level among the 70-74 and 75-79 year olds, at 55% and 49%.

Table 8.3. The Determinants of Projected Changes in Driver Risk, National Level

	Men		Women	
Age	Income	Seatbelt	Income	Seatbelt
65-69	58.09%	41.94%	62.46%	37.44%
70-74	39.92%	59.90%	44.13%	55.78%
75-79	57.71%	42.19%	50.54%	49.37%
80-84	59.19%	40.68%	62.40%	37.40%
85+	37.17%	62.93%	38.69%	61.42%

While seatbelt use is a traditional focus of concern, the substantial contribution of income growth to decreasing crash risk in these projections is important as 40% to 60% of the decrease in this risk indicator is attributable to increasing income. Our modeling has not specified the routes of effect of higher real, elderly income, but we have pointed to the most likely possibilities as ability to afford safer equipment and generally higher valuation of safety which may spill over into driving practices as well as equipment purchases. Income and health are generally positively correlated, although in the regressions underlying these projections, better health, as measured by our indicator, would have elevated crash rates. The health-income-crash rate relationship needs further research.

Regional variations in the different variables' contributions to the crash rate decline are not as striking as they have been in some of the other projections, as Table 8.4 clearly shows. The seatbelt contribution to crash rate decline is largest among western women 85+, at 83% (followed by western men in the same age group, at 74%), and the lowest is among

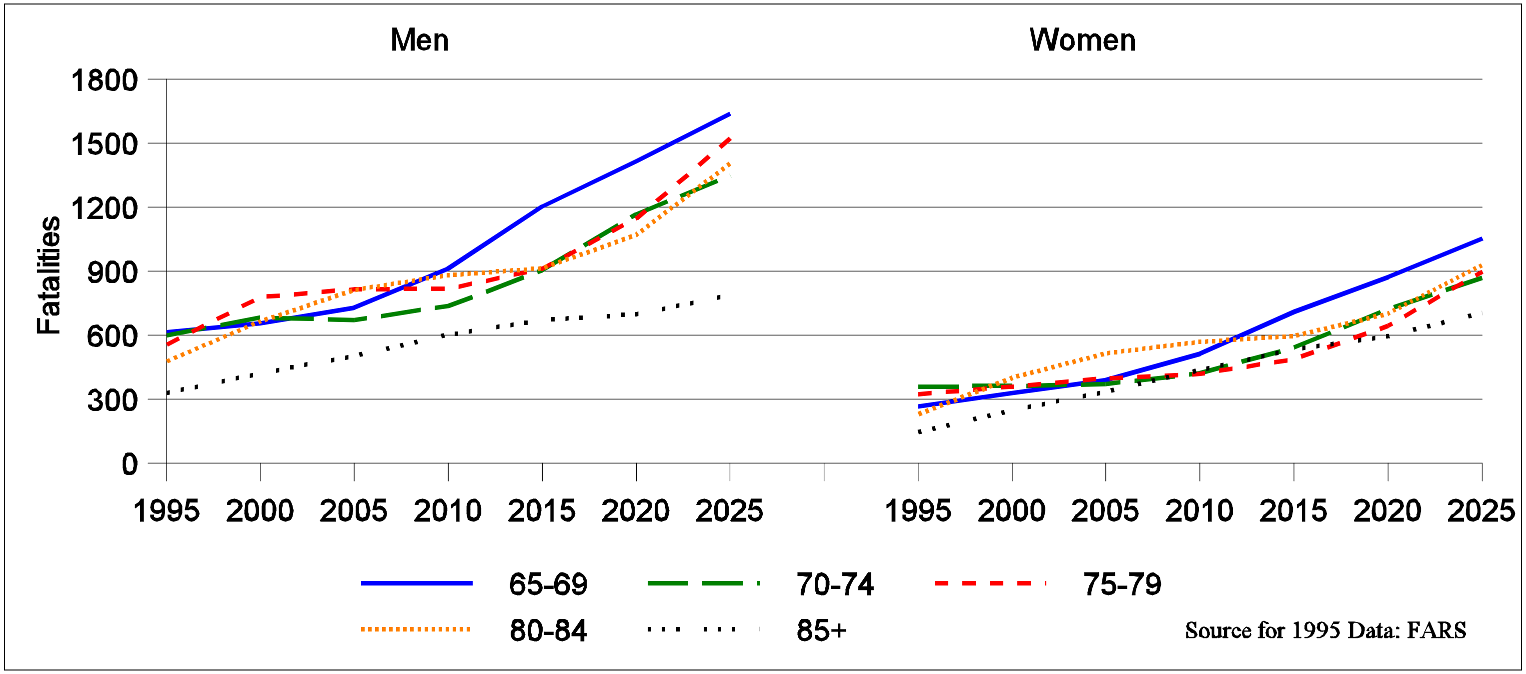

Figure 8.4. Projected Driver Fatalities By Age

80-84 men in the South, at 30%, accompanied by southern women 80-84, at 31%. Of course, these groups have the highest contribution of income to crash risk decrease.

Table 8.4. The Determinants of Projected Changes in Driver Risk, Regional Level

	Men		Women	
Midwest				
Age	Income	Seatbelt	Income	Seatbelt
65-69	61.78%	38.22%	62.62%	37.38%
70-74	35.70%	64.30%	38.93%	61.07%
75-79	45.01%	54.99%	39.55%	60.45%
80-84	53.41%	46.59%	61.59%	38.41%
85+	39.32%	60.68%	39.71%	60.29%
Northeast				
Age	Income	Seatbelt	Income	Seatbelt
65-69	57.05%	42.95%	57.48%	42.52%
70-74	33.33%	66.67%	43.96%	56.04%
75-79	50.98%	49.02%	51.75%	48.25%
80-84	51.51%	48.49%	51.63%	48.37%
85+	35.65%	64.35%	39.48%	60.52%
South				
Age	Income	Seatbelt	Income	Seatbelt
65-69	55.81%	44.19%	61.79%	38.21%
70-74	47.64%	52.36%	50.95%	49.05%
75-79	67.31%	32.69%	58.91%	41.09%
80-84	69.89%	30.11%	68.94%	31.06%
85+	43.63%	56.37%	48.99%	51.01%
West				
Age	Income	Seatbelt	Income	Seatbelt
65-69	58.89%	41.11%	68.60%	31.40%
70-74	37.82%	62.18%	37.70%	62.30%
75-79	62.52%	37.48%	47.11%	52.89%
80-84	55.58%	44.42%	63.81%	36.19%
85+	25.73%	74.27%	16.70%	83.30%

8.3 DRIVER FATALITY PROJECTIONS

The total *number* of driver fatalities, represented in Figure 8.4 and presented in greater detail in Tables A.4.1-A.4.4 in the appendix, is projected to increase in a less stable way. This is due to the conflicting influences of increased population, VMT, and percentages of people driving, and decreases in the fatal crash rate of drivers. For male older drivers aged 65-69, the number of annual driver fatalities is projected to increase 166.6%, from 614 to 1637. For females of the same age group, fatalities are projected to increase by 294%, from 267 to 1051. This greater increase is due in large part to greater projected increases for women in VMT and the percent of women who drive. Note that, although the rate of increase in the number of fatalities is expected to be higher for women during 1995-2025, the absolute number of female driver fatalities will remain lower than that of male drivers.

In the regional breakdown of driver fatalities, the greatest number occurs, as expected, in the South. In 1995, 42% of all driver fatalities, nationally, occurred in the South. In 2025, our projections indicate that this number will rise to 51%, mostly due to higher than average expected population growth of the elderly in this region.

The projections of our total fatalities measure mirrors the trends found in driver fatalities, from the higher rate of increase in total fatalities attributed to women, to the increase in the already high proportion of fatalities occurring in the South. We do not present the numbers for those projections in either tables or graphs since they were so similar to the driver fatality results. Additionally the implicit attribution of fault in that measure lends those numbers to easy misinterpretation.