5. OVERVIEW OF THE MODELING SYSTEM

5.1 INTRODUCTION TO MODELING

Our approach to projecting the number of fatalities involving older drivers used four distinct components:

1. projection of non-institutionalized population,
2. projection of the percentage of that population that drives,
3. the projected average miles driven (VMT) per driver, and
4. the projection of fatal crash risk associated with each mile driven.

The non-institutionalized portion of the elderly population that actually drives represents the population exposed to driving risks. We chose not to use the number of drivers' licenses issued to represent the number of older drivers because many older drivers retain their licenses only for identification purposes. Our methodology for projecting the percentage of elderly drivers is given in Chapter 6. VMT was used to represent the exposure to crash risk as a likelihood of being involved in highway crashes per mile of vehicular travel. Our methodology for projecting VMT is given in Chapter 7, while our methodology for projecting the fatal crash risk is given in Chapter 8.

The population projections provided by the U.S. Census Bureau indicate that nearly all subgroups of the elderly population will increase from 50 to 150 percent by the year 2025. This enormous increase will have a substantial effect on the projection of the absolute number of casualties, even when crash risk per mile driven decreases or remains constant.

Beginning with non-institutionalized population as a foundation, we used the four components to derive the fatality projections. First, we took the non-institutionalized population projections, sorted by age groups, gender, and region, and multiplied those estimates by the projected percentages of the elderly population that will be driving, which gives us a total number of older drivers in the future.

Then, we took the projected number of older drivers and multiplied it by the average VMT per driver to obtain a total number of miles driven by elderly drivers. Finally, we multiplied the estimated total number of miles driven by the crash risk per mile driven to find our ultimate goal, the projected total number of fatalities (either driver fatalities or total fatalities, as explained in Chapter 8).

The behavior of the individual elderly person was taken into account in projecting two of the final three components of the system – when we look at how much an individual with certain characteristics will drive or whether he or she will be a driver. The behavior underlying each of these components is modeled empirically with regression analysis of data from the NPTS, the NHIS, and several other related sources. The estimated equations serve as predictive models used to project our components into the future as the independent variables of those regressions change predictably. NPTS survey data on individual drivers are available to estimate VMT and the probability that an older person will continue to drive. Our projections, however, are in the form of averages for particular subgroups of drivers (e.g., southern male drivers between 65 and 69 years old). The effects of our independent variables in these two models are also estimated at the individual level, while projections are in the form of averages for subgroups. We make the step from characteristics of the individual to characteristics of the group assuming that such groups will be homogeneous over time, and representative of the “average”driver.

5.2 BRIEF EXPLANATION OF DEMAND THEORY

Demand theory underlies the projection of the demand for VMT and the need a person has to be a driver. This theory predicts that the quantities of goods or services chosen by a consumer will respond to changes in the consumer’s income and to changes in the prices, or acquisition costs, of the goods or services and of closely related goods and services, referred to as substitutes and complements. Consumption will increase in response to decreases in the acquisition cost of the goods or services in question, to decreases in the costs of complementary goods and services (items used in conjunction with the good or service under study), and to increases in the costs of substitute goods and services. Abstracting from changes in acquisition costs, increases in income will lead consumers to increase their consumption of a good or service. In each of the predictive models, data on the acquisition costs of the directly concerned consumption item, and on substitutes and complements, are not readily available. Thus, we rely on surrogates for these costs.

Demand theory for items that involve binary choices (i.e., to buy or not to buy) has been developed in the context of the demand for durable goods for which different varieties are available and for transportation modal choice (automobile or public transit primarily). When several choices are available, a system of demand equations is used, with the dependent variables measuring the probability of choosing one item relative to the probability of choosing another, estimated typically in a logit framework. When there is only one choice, the actual choice is "yes" or "no," with "no" being the default choice and the equation system collapsing to a single equation. In the case of the "driver" model, the choice is whether to drive.

5.3 POPULATION PROJECTIONS

Population growth, by 5-year age groups and by gender, was projected by the U.S. Census Bureau (Tables B.1-B.4 in Appendix B, and Figure 5.1). The U.S. Census Bureau also projected population growths at individual Census regions. The regional distribution of population is projected by considering the migration into the South and West over the next twenty-five years. This movement results in large percent increases in total older driver fatalities in 2025 over 1995 in those regions. Since elderly persons in nursing homes are much less likely to be drivers than those living alone or with their families in non-

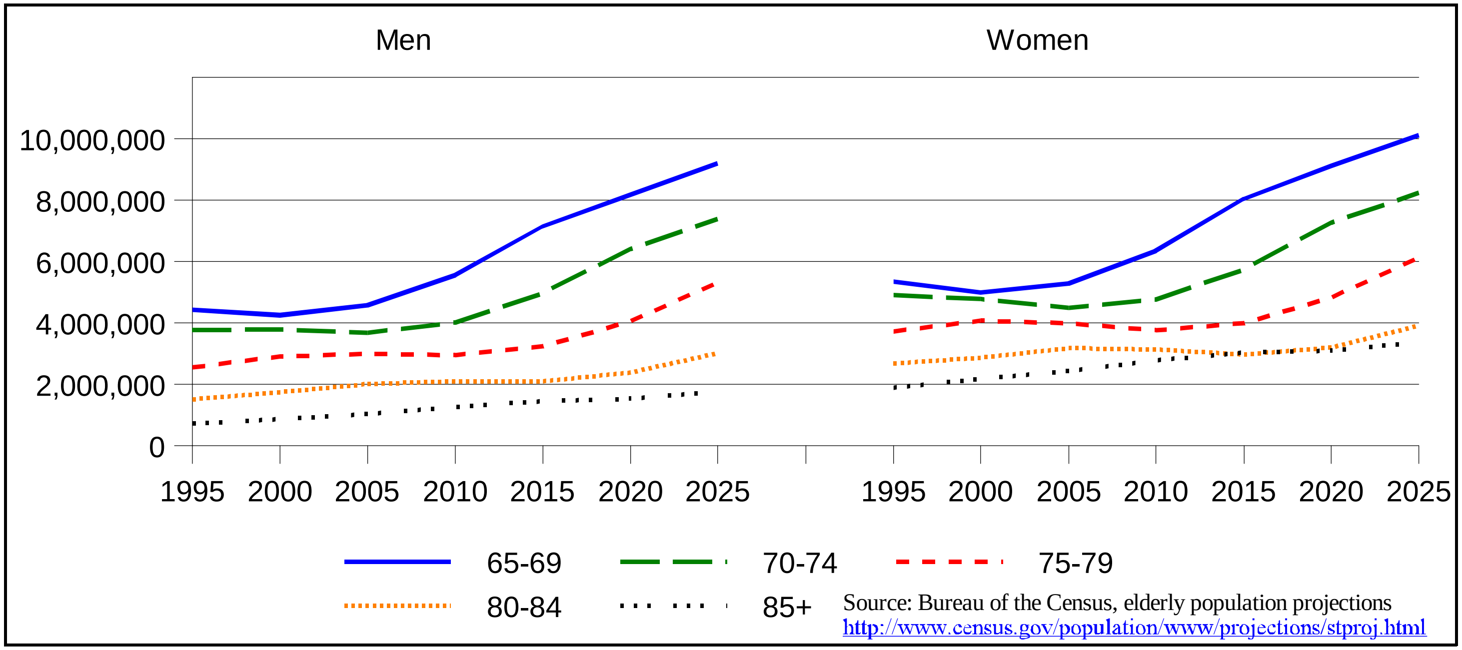

Figure 5.1. Projected National Non-Institutionalized Population

institutionalized settings, our projections of future elderly populations who will drive exclude those who are projected to be institutionalized (American Health Care Association, 1997, p.7).

The number of people aged 65-69 projected in 2000 is smaller than the actual number of people who are 65-69 in 1995, and some readers may find that decrease in cohort size odd and discomforting. That projection may be a statistical bump to some extent, but it also may reflect the effect of the Great Depression on the number of people born between 1931 and 1935, the group represented by the 65-69 year olds in 2000. Both birth rates and the numbers of births dropped below their trend levels between 1931 and 1941. Of course, the birth rate in the United States had been declining for decades as part of the demographic transition, but even that falling birth rate fell below its trend rate of decrease between 1931 and the beginning of World War II. In a growing population we would expect subsequent cohorts to be larger than earlier ones, but a peak of live births occurred in 1930, a level that was not recovered until 1941 (U.S. Census Bureau, 1975, p. 49). This smaller size of the 1931-35 cohort is likely attributable to a combination of delayed marriage, delayed childbearing within existing marriages, possible reductions in births outside of marriage (likely to have been a small effect in either direction in the 1930s), higher infant and child mortality, and possibly increased adult morbidity rates later in life. Whether immigration could have compensated for this reduction in natural increase among this cohort is an open question, but with tighter immigration restrictions dating from the 1920s, even post-World War II immigration by persons born between 1931 and 1935 may not have compensated for the reduction in U.S. births.

5.4 HEALTH STATUS DATA

The health index included in the probability of continuing to drive and VMT models is constructed using relationships found in the NHIS data. The NHIS' Activity Limitation Status (ALS) variable assumes values 1 to 4, which indicate a subject's health:

1. Unable to perform major activity
2. Limited in kind/amount of major activity
3. Limited in other activities
4. Not limited

Using this original variable, a continuous index was computed where an ALS of 1 remained 1, an ALS of 2 became 3, an ALS of 3 became 25, and an ALS score of 4 was given a value of 40. The log of this continuous predictor was then used as a dependent variable and regressed on factors including region, gender, age, race, marital status, education attainment, and income. Three regressions were run for the years 1983, 1990, and 1995. The coefficients of these models can be found in Table 5.1.

Once these models were complete, we took these coefficients into the NPTS data and calculated predicted ALS scores (on the 1, 3, 25, 40 scale) using the information found in NPTS. At this point, the scores were then broken down by percentiles with each NPTS respondent given a score on a 1 to 4 scale so that the distribution of the total population's scores matched the distribution found in the original NHIS measure. The natural log of these scores on the 1 to 4 scale were then used in the "driver" and VMT demand models as an approximate measure of an individual's health.

While we acknowledge that our method is for integrating health and driving has limitations, we nonetheless feel that an association is necessary. Even for non-institutionalized persons, health limitations curtail driving. Additional research in the relationships between health and the decision to drive and between health and VMT is recommended in Chapter 10.

Table 5.1. NHIS Activity Limitation Status Regression Coefficients and T-Values

		1983*		1990		1995	
R-squared			0.117095		0.102946		0.109982
		ß	t	ß	t	ß	t
Intercept		2.0628	24.07	2.0933	26.29	1.9726	23.56
Northeast		0.0493	2.16	0.0401	1.90	0.0946	4.24
Midwest		0.0269	1.21	0.0111	0.54	0.0817	3.69
South		0.0147	0.71	-0.0331	-1.68	-0.0034	-0.16
Gender (male=1)		0.1552	1.72	0.3392	3.73	0.3337	3.40
Ageclass	45-49	0.5070	8.62	0.5929	10.26	0.7606	11.88
	50-54	0.4058	6.76	0.5261	8.98	0.6702	10.48
	55-59	0.3017	4.93	0.5111	8.67	0.6218	9.80
	60-64	0.3031	4.94	0.4102	7.00	0.4601	7.22
	65-69	0.3241	5.28	0.3128	5.34	0.4622	7.22
	70-74	0.8146	13.22	0.8080	13.51	0.9262	14.21
	75-79	0.7290	11.36	0.6748	10.93	0.7318	10.77
	80-84	0.3620	5.29	0.4982	7.47	0.5847	8.09
Gender-ageclass	M 45-49	-0.2082	-2.18	-0.3631	-3.78	-0.3939	-3.77
	M 50-54	-0.2028	-2.09	-0.4254	-4.37	-0.3884	-3.71
	M 55-59	-0.1971	-2.00	-0.4962	-5.07	-0.5508	-5.28
	M 60-64	-0.3692	-3.71	-0.5885	-6.01	-0.5979	-5.70
	M 65-69	-0.5231	-5.24	-0.5892	-5.99	-0.7184	-6.78
	M 70-74	-0.2198	-2.17	-0.3658	-3.64	-0.3196	-2.96
	M 75-79	-0.2785	-2.64	-0.2835	-2.71	-0.1937	-1.72
	M 80-84	0.0264	0.23	-0.2865	-2.54	-0.1765	-1.42
Race		0.0673	2.98	-0.0025	-0.12	0.0921	3.81
Married		0.0843	4.67	0.0470	2.76	0.0151	0.83
years of school**		0.2962	12.89	0.3012	14.67	0.2718	14.18
Income	< $10K	-0.9489	-29.97	-0.8093	-28.53	-0.7569	-22.92
	$10-20K	-0.5825	-23.64	-0.4976	-21.93	-0.4001	-13.07
	$20-30K	-0.3718	-15.81	-0.2668	-11.75	-0.1583	-5.13
	$30-40K	-0.1542	-6.23	-0.1524	-6.29	-0.0880	-2.68
	$40-50K	-0.0985	-3.82	-0.0601	-2.32	-0.0518	-1.41

*****(Note that the 1983 NPTS does not contain a region variable and thus predictions for that year were obtained omitting the region portion of the model).

** The years of school variable is computed by taking the number of years schooling a given individual had, adding one, and taking the natural log of the result.

5.5 ASSUMPTIONS

As noted in Section 4, our search for data sets on the elderly which include information on health, income, education, driving status and practices (e.g., seat belt use, driving while intoxicated, use of transit), and VMT/person, while also including variables for gender, age group, and region, concluded that these data sets simply do not exist. In addition, it is not generally possible to form direct links between data sets. Therefore, for the critical missing link of health status, we constructed the surrogate described in the preceding section.

In other cases, we simply excluded a particular projection from our model. One example of such an exclusion is the future impact of technology on the elderly driver. Improved infrastructure, in-vehicle devices, and ITS innovations (see also Sections 3.3.5 and 10.3.1) may have a tremendous impact. We simply do not know how these innovations will impact the older driver (or how they will influence the elderly decision to choose alternative transportation or to walk). We also do not know how public policy will affect the building of new highways or affect our transportation modes. Finally, because of the tremendous influence of this large elderly population in the future, changes in public transportation will almost certainly occur. Predicting the extent of these changes, however, would be mere speculation.

Using traditional regression models on historical data is an effective way to estimate the effects of predictive variables on components of our modeling system. However, assuming that the magnitude of historical trends will continue indefinitely into the future sometimes yields absurd results. In situations where this problem presents itself in our projections, we logically modify coefficients of variables to provide more reasonable estimates of our modeling system components. For example, the probability of driving must be between 0 and 1 and projections of VMT by elderly drivers should not exceed a logical maximum, explored in Chapter 7.