AUTOMOBILE SAFETY AND THE VALUE OF STATISTICAL LIFE IN THE FAMILY: VALUING REDUCED RISK FOR CHILDREN, ADULTS AND THE ELDERLY

Timothy Mount, Weifeng Weng, William Schulze
Cornell University
and
Laurie Chestnut
Stratus Consulting

March 2000

ABSTRACT

Little work has been done theoretically or empirically to obtain the value of a statistical life (VSL) for children or the elderly. This paper addresses both of these issues by first presenting a theoretical model of how families value risk and then examining family automobile purchases. Automobile safety is shown to be a family public good, where the marginal cost of purchasing and operating a safer automobile is set equal to the usage-weighted sum of the values of statistical life of family members. We use data on automobile purchases to estimate how much single car families of different composition (in terms of children, adults and the retired) spend on safety to impute the VSL of each age group. We find that children are valued more highly than some existing studies suggest. The VSL of the elderly is consistent with the discounted present value of life years approach. These results come, in great part, from an analysis of the fatal accident data that shows that fragility--the susceptibility to death in an accident of fixed severity--increases with age. Also, we show that an important factor for survival in two-vehicle accidents is the relative weight of the vehicles involved. The models of survival in fatal accidents are used to estimate standardized risks of mortality in different types of vehicles. These standardized risks are then used in hedonic models of the purchase price and fuel efficiency of a specified vehicle to determine the capital costs and the operating cost of reducing the risk of mortality.

This research was supported by United States Environmental Protection Agency Cooperative Agreement Number CR824393-01-1. We would like to thank Sharon Sandlan for her assistance in preparing the manuscript. All conclusions and remaining errors are the sole responsibility of the authors.

Section 1. Introduction

Little work has been done either theoretically or empirically to value morbidity and mortality either for children or retired adults (for exceptions see Blomquist, et al., 1996, and Jenkins, et al. 1999). This paper addresses both of these issues by first presenting a theoretical model of how families value risk and then examining family automobile purchases. In particular, we show that parents may value risks to their children's lives and health (the model assumes two altruistic parents) through Nash cooperative bargaining to determine how much money to invest in the health and safety of their children. To allow empirical estimation of values, automobile safety is then shown to be a family public good, where the marginal cost of purchasing and operating a safer automobile is set equal to the usage-weighted sum of the values of statistical life (VSL) of family members. We use data on automobile purchases to estimate how much families with children spend on automobile safety and how much families with retired members and no children spend on safety, for comparison to families without children or retired members. This not only allows indirect estimation of an average value of a statistical life (VSL) for each type of family, but also allows estimation of an average value of a statistical life (VSL) for each age group: children, adults and seniors.

Our research using secondary data is a preliminary effort to determine the feasibility of collecting a national data set to allow direct estimation of separate values for mortality and possibly morbidity for different family members from choices made concerning both the type of vehicle and usage pattern by family members. A major limitation of the secondary data we use here is that only the usage weighted average statistical values of life per family member can be estimated for single car families. We examine families with different compositions to impute the VSL for different family members.

The paper is organized as follows: Section 2 presents a simplified theoretical model of family automobile purchase decisions focusing on safety and how safety values for each individual are determined in a family setting. Section 3 addresses the problem of driver characteristics affecting estimates of the inherent risk of fatality of different automobiles and develops a procedure for identifying the driver independent level of risk. Section 4 describes our empirical work estimating a hedonic price function for automobiles showing a negative correlation between risk of fatal accident and price and operating costs as well as our estimates of average implied values of life for different family groups. Finally, we summarize our findings and implications for future research in Section 5.

Section 2. Theoretical Issues

How willingness to pay (WTP) for health and safety may vary with the age of the person at risk is a very important policy question for which we have little well-established empirical data. Cropper and Freeman (1991) address this question with a life-cycle consumption-saving model that they apply with a quantitative example to examine how WTP for a risk reduction in the current time period can be theoretically expected to change over a person's lifetime. This model is based on the premise that a person makes consumption and saving decisions over time to maximize personal utility. Because this model is based on the premise that utility is a function of consumption, the authors note that, if there is additional utility derived from survival per se, then the life-cycle model provides a lower bound estimate of WTP. The quantitative example depends on assumptions regarding a lifetime pattern of earnings, endowed wealth, the rate of individual time preference, and other parameters of the model. These will all vary for different individuals, and uncertainty exists empirically about population averages for many of these factors. However, using reasonable values to calibrate the model is illustrative. Cropper and Freeman note that if consumption is constrained by income early in life, the model predicts that VSL increases with age until age 40 to 45, and declines thereafter. Shepard and Zeckhauser (1982) also illustrate this point with numerical examples for the life-cycle model. When they estimate the model with reasonably realistic parameters and assume no ability to borrow against future earnings or to purchase insurance, they find a distinct hump in the VSL function with a peak at around 40 years and dropping to about 50% of the peak by 60 years. When they allow more ability to borrow against future earnings and to purchase insurance, the function flattens and at 60 years drops only to 72% of the VSL at age 40. However, the hump shape to the VSL over a person's lifetime remains.

The conclusions reached by these theoretical analyses of the effect of age on WTP for mortality risk reduction using the life-cycle model are somewhat consistent with the empirical findings obtained by Jones-Lee et al. (1985). However, the empirical findings show that WTP varies with age much less than would be predicted by the life-cycle models. In this stated preference study, respondents gave WTP estimates for reductions in highway accident mortality risk and the answers showed a fairly flat hump-shaped relationship between VSL and age, peaking at about age 40. Although the directions of the changes in WTP with age are consistent with what the life-cycle models predict, the magnitudes of the changes are smaller. The Jones-Lee et al. results show that at age 65 the VSL is about 90% of the VSL of a 40-year-old person.

It is often suggested that WTP will be lower for the elderly than for the average adult because expected remaining years of life are fewer. This expectation is based on the presumption that WTP for one's own safety declines in proportion to the remaining life expectancy. Some analysts have suggested that effects of age on WTP might be introduced by dividing average WTP per statistical life by average expected years of life remaining (either discounted or not) to obtain WTP per year of life (Moore and Viscusi, 1988; Miller, 1989; Harrison and Nichols, 1990). Such a calculation implies very strong assumptions about the relationship between life expectancy and the utility a person derives from life; namely, that utility is a linear function of life expectancy and that the value of life year remains constant.

Determining appropriate WTP values for changes in mortality risks to children poses some particular analytical challenges. Children are not the economic decision makers whose preferences can be analyzed to determine an efficient allocation of society's resources regarding their own health and safety, so both revealed and stated preference approaches must rely on parental decisions to show what WTP for children's health and safety might be. Based on the expected relationship between WTP and expected life-years lost, it may be reasonable to assume that reductions in risks to children are valued equal to or greater than risks to adults. Blomquist, et al. (1996) support this view in their analysis of seat belt use for children. On the other hand, the life-cycle consumption-saving models show increasing WTP for risk reductions between the ages of 20 and 40, reflecting the typical pattern of increasing income and productivity during this stage of life. Extending this to children might suggest lower WTP for reducing risks to children, however, this pushes beyond the theoretical constructs of the life-cycle model regarding an individual as an economic decision maker. The only theoretical model that addresses these concerns, with respect to dependent children, has been developed by Chestnut and Schulze (1998). Their work treats the case of a family with non-paternalistic altruistic parents who engage in Nash cooperative bargaining to determine health and safety expenditures on their children and the implied VSL. We use this model as a starting point for our analysis.[1]

Given the state of existing research, our first task is to develop a model that can potentially explain the behavior of households with dependent children. This model is developed in the context of automobile safety to allow empirical estimation of an appropriate family VSL, since the existing theoretical literature only considers individuals rather than families, with the exception of the work by Chestnut and Schulze mentioned above. Our work here paraphrases this earlier work and adds a hedonic market for automobile safety.

[1] It should be pointed out that some interesting revealed preference empirical approaches based on a household production function framework to analyze household expenditure decisions as they relate to children's health have been attempted (Agee and Crocker, 1996; Joyce et al. 1989). These analyses infer implicit WTP for changes in children's health as revealed by expenditure decisions of the household. Limitations in available data and analytical difficulties in properly specifying and verifying modeled relationships pose challenges for this approach; however, its basis in actual household decisions and behavior is an important strength. Estimates of WTP for changes in mortality risk for children are not

We begin by considering the case of a single individual with no family who may, or may not, survive for a single period. The following notation will be useful:

c = consumption,

w = wage income,

r = risk of a fatal automobile accident,

Π = probability of survival without automobile fatality risk,

Π-r = probability of survival with automobile fatality risk,

$H(\Pi)$ = health expenditures (increasing in Π),

$P(r)$ = automobile price (decreasing in r), and

$F(r)$ = total fuel consumption expenses

$U(c)$ = strictly concave utility function.

Note: subscripts or primes denote derivatives where appropriate.

The individual must make two choices. First, the baseline probability of survival, Π, is chosen subject to the constraint that increasing Π increases health expenditures, $H(\Pi)$, consequently reducing both consumption, c, and money available for purchasing a car, P. Similarly, the individual chooses how risky a car to drive, r, taking into account that lower r implies that both the price of the car, P(r), and operating expenses, F(r), are greater. Investments in health, Π, and automobile safety, reducing r, are chosen prior

---

directly available from these two studies, but similar approaches might be applied to obtain to realizing whether or not the individual will survive. The individual is assumed to maximize expected utility,

$$(\Pi - r)U(c), \tag{1}$$

where the death state provides no utility because the individual has no family, subject to the budget constraint,

$$(\Pi - r)(w - c) - P(r) - F(r) - H(\Pi) = 0. \tag{2}$$

This budget constraint assumes that costless insurance (available for expected value) is available both to cover the purchase price and operating costs of the automobile, P + F, and initial health and other safety investments, H. Most car loans, in fact, carry life insurance for the amount of the loan, and life insurance could presumably cover the costs of other health and safety investments. The optimal choice of Π is then determined by

$$H_{\pi} = VSL, \tag{3}$$

and, the optimal choice of r is determined by

$$-(P' + F') = VSL, \tag{4}$$

where,

$$VSL \equiv (U/U_c) + w - c. \tag{5}$$

Equation (3) sets the marginal health cost of increasing the odds of survival equal to the value of the individuals statistical life (VSL) while equation (4) sets the marginal increase in cost for purchasing and operating a

---

such WTP estimates.

safer car equal to the VSL as well. The VSL is defined in (5) for the case of perfect insurance markets and is equal to the monetized value of utility, (U/U_c), which is lost in death, plus the excess of earnings over consumption. The interpretation of this relationship is much clearer in the family setting that we treat below, so we will defer discussion.

The model developed above can readily be extended to a family setting by using the Nash cooperative bargaining between parents approach employed by McElroy and Horney (1981). Following our previous work (Chestnut and Schulze, 1998), we modify the notation used above, again considering a single car family (the case we analyze empirically), as follows:

n = the size of the family,

i = 1, 2,….,n denotes individual family members,

i = m = 1 denotes the mother,

i = f = 2 denotes the father,

i= k = 3, …..,n denotes children,

c_i = consumption of the ith family member,

w_i = wage of family member i,

r= automobile fatality risk, the same for all family members,

Π_i = probability of survival, excluding automobile fatality risk, of i,

$H(\Pi_1,\ldots, \Pi_n)$ = family health expenditures (increasing in Π_i),

P(nr) = automobile price (decreasing in total family risk, nr),

$F(nr)$ = total fuel consumption expenses (decreasing in total family risk, nr),

$U^k(c_k)$ = child's utility function,

$U^i(c_i;\ldots,(\Pi_k\text{-}r)U^k(c_k),\ldots)$ = parent's utility function ($i = m, f$), and

E^i = bargaining threat point of expected utility in divorce ($i = m, f$).

The family must decide how much to allocate to each family member for consumption, spending on the health of each (and in so doing select survival probabilities), and on the safety level of the single automobile they purchase for all. Note that the demand for driving is inelastic in this model, since the only driving choice is over the risk of the chosen automobile. The hedonic price function for the automobile is now taken as $P(nr)$ so that the total family risk level determines the price of the car where $P'<0$ so more dangerous cars are cheaper. All of the existing hedonic price analyses of automobile safety use total fatalities per year for a vehicle model divided by the total number of that model on the road as the risk variable. Thus, the risk measure is not divided by occupancy (n in this theoretical model). It is, in fact, plausible to suppose that it is more expensive to increase the safety for each of four passengers than for one, so this assumption may be reasonable. Also, since heavier cars are unambiguously safer in collisions, fuel costs and other operating expenses are positively related safety and negatively to family risk, so these expenses are defined as $F(nr)$ where $F'<0$. Thus, rather than treating fuel economy as a hedonic characteristic in the hedonic price function for automobiles, where it inevitably has the wrong sign, we treat fuel costs as an expense which is also a function of the choice of the other attributes of the car.

The utility functions of both the father and mother are assumed to depend not only on their own consumption, but also on the expected utilities of each of their children. The children's utility is assumed to be solely a function of their own consumption.

Investment in the safety and health of their children is a public good to the parents, which is the subject of negotiation, as is the level of consumption of each. The Nash cooperative bargaining model assumes that the solution maximizes the multiplication of the increase in the expected utility of the outcome over the threat point expected utility in divorce for the mother and the father. The threat points are assumed, in models of the family, to be a function of divorce laws, job opportunities, etc. Thus, in the Nash cooperative bargaining solution,

$$(\Pi_m\text{-}r)U^m(c_m;..,(\Pi_k\text{-}r)U^k(c_k),..)\text{-}E^m]\ [(\Pi_f\text{-}r)U^f(c_f;..,(\Pi_k\text{-}r)U^k(c_k),..)\text{-}E^f], \quad (6)$$

is maximized with respect to c_i, Π_i, and r, subject to the budget constraint,

$$\sum_{i=1}^{n} (\Pi_i\text{-}r)\ (w_i - c_i) - P(nr) - F(nr) - H(\Pi_1,\ldots\ldots,\Pi_n) = 0.$$

(7)

The resulting conditions for allocating health expenditures and survival probabilities take the form:

$H_i = U^i/U^i_c + w_i - c_i \equiv VSL_i \qquad i = 1,\ldots\ldots\ldots,n.$

(8)

The remarkable fact is, that, in spite of the complicated structure of the problem specified above, the implied VSL_i for each family member shown in (8) is identical in form to that for the single individual shown in (5) above. The interpretation of the VSL_i can be illustrated with the following examples. Imagine that the mother is the sole breadwinner with a stay-at-home father. In this case, assuming that the children are young, $w_i - c_i < 0$ for the other family members and $w_m - c_m > 0$ for the mother. Thus, if the mother were to die, this would be a severe financial blow to the rest of the family and the mother's VSL would reflect this relative to the VSL of other family members. For young children it is clear that $w_k - c_k < 0$ in the short run. However, in the inter-temporal version of the model, $w_k - c_k$ is replaced by its discounted present value, which may be positive. U^i/U^i_c depends solely on c_i in the single period model and on the lifetime consumption pattern in the full inter-temporal model. The important point is that the child's consumption depends in youth on the parents' income and wealth. Further, if parents find the value of their child's smile to be high enough, the child's consumption will be maintained, by them, at a high level, leading to a high VSL. A young child's utility may also be large in the parent's view from relatively small levels of financial consumption, also leading to a high VSL. These arguments suggest that the VSL of children is a purely empirical question and depends not only on their own life cycle wealth but also on their family's wealth and the beliefs of the parents regarding their children's utility.

Finally, the choice of automobile risk, r, is determined by

$$-n(P' + F') = \sum_{i=1}^{n} VSL_i. \tag{9}$$

Thus, the safety of the shared family vehicle is determined by a public good condition which sets the marginal cost of obtaining a safer vehicle for each individual equal to the sum of the VSLs of individual family members. Thus the marginal cost of a safer vehicle is the slope of the hedonic price function for automobile safety, -P′, plus the marginal fuel cost penalty, -F′, which, by (7), is set equal to the average VSL for the family, $\sum_{i=1}^{n} VSL_i / n$, to determine the choice of automobile risk, r.

Thus, if we examine P′ + F′ for different households with a single car, we can obtain estimates of the average value of life for those households. However, the average is a weighted average where the weights are determined by each family member's use of the vehicle.

Section 3. Statistical Model of Automobile Fatalities

The basic data on fatalities in automobile accidents provide a census of accidents with at least one fatality. Hence, the probability of an accident being included in the data set depends on the number of individuals involved in an accident as well as the characteristics of the vehicles and driving behavior (e.g. the use of seat belts). This can be illustrated by the following examples for a one-vehicle and a two-vehicle accident. For a one-vehicle accident, assume that the driver and one passenger have the same probability of survival $P^* = P\{\text{survival}\} = .5$. The four possible events are illustrated below, and in this example, each event has the same probability of occurring of $0.5^2 = .25$.

		Passenger	
		Fatality	Survives
Driver	Fatality		
	Survives		

Accidents in which both the driver and the passenger survive (shaded) are not included in the data set. Hence, the probability of either the driver or the passenger surviving in an accident with a fatality corresponds to the probability of one of three possible events with a probability of $P = P\{\text{survival} \mid \text{at least one fatality}\} = 0.25 / (1 - 0.25) = 0.33$. The observed probability of survival in the data set, P, is much lower than the unconditional probability, P^*. The observed probabilities of survival, P, are 0, 0.33, 0.43 and 0.47 for 1, 2, 3 and 4 occupants, respectively, and the values of P increase and get closer to P^* as the number of occupants increases.

In the one-vehicle accident with two occupants and $P^* = 0.5$, the expected number of fatalities is one (the modal type, corresponding to 91% of one-vehicle accidents in the data set). In a two-vehicle accident with two occupants in each vehicle, the same expected number of fatalities would occur if $P^* = 0.25$ (for multiple-vehicle accidents, 54% of vehicles have no fatalities, and 40% have one fatality). The probability of an accident having at least one fatality, and being in the data set, is $(1 - 0.75^4) = 0.68$. There are 16 possible permutations of survival / fatality for the four individuals and 15 of them are in the data set. For any selected individual, 7 of the 15 observed events correspond to surviving with a probability $P = 0.63$. While this is lower than the unconditional probability of survival $P^* = 0.75$, it is much larger than the corresponding probability for the one-vehicle accident $P = 0.33$. Setting the severity of the two types of accident at the same level (E [number of fatalities] = 1) makes the probability of a specific individual surviving in a fatal accident almost twice as large in the two-vehicle accident as in the one-vehicle accident. The reason is simple, for any unconditional probability of survival P^*, the expected number of fatalities is $P^* \times$ number of individuals in the accident.

Since the data set includes all accidents in which at least one fatality occurs, a fatality is more likely to occur if more people are involved.

In reality, the unconditional probabilities of survival for individuals differ by individual characteristics such as age, whether or not a seat belt was used and the location of the seat in a vehicle. In addition, these probabilities differ by the type of vehicle, and for two-vehicle accidents by the relative size and type of the other vehicle. For an individual *i* riding in vehicle *j*, the unconditional probability of survival in a two-vehicle accident, for example, can be written:

$$P_{ij}^{*} = f\left(x_i, v_{i1}, v_{i2}\right) = f\left(z_{ij}\right)$$

where x_i are the characteristics of individual *i*

v_{i1} are the characteristics of individual *i*'s vehicle ($j = 1$)

v_{i2} are the characteristics of the other vehicle ($j = 2$)

z_{ij} is the vector of all explanatory variables

The probability of observing at least one fatality in the accident is

$$\left(1 - \prod_{j=1}^{2} \prod_{i=1}^{n_j} P_{ij}^{*}\right)$$

where n_j is the number of individuals in vehicle *j*.

If $P_{ij}^{*} = f\left(z_{ij}\right)$ is specified as a logistic function, then it can be written:

$$P_{ij}^{*} = \frac{e^{z_{ij}'\beta}}{1 + e^{z_{ij}'\beta}}$$

where β is a vector of unknown parameters that are the same for all individuals and vehicles. Using this form, it would be possible to recover the unconditional probabilities of survival using the available data on accidents with at least one fatality. In the simplest case with one individual in each vehicle, for example, the probability of observing two fatalities in the data set would be:

$$\frac{1}{1+e^{z'_{11}\beta}+e^{z'_{12}\beta}}$$

and the unconditional probability of two fatalities would be:

$$\frac{1}{1+e^{z'_{11}\beta}+e^{z'_{12}\beta}+e^{(z'_{11}+z'_{12})\beta}}$$

The unconditional probability of the individual in vehicle 1 surviving would be:

$$P^{*}_{11}=\frac{e^{z'_{11}\beta}+e^{(z'_{11}+z'_{12})\beta}}{1+e^{z'_{11}\beta}+e^{z'_{12}\beta}+e^{(z'_{11}+z'_{12})\beta}}=\frac{e^{z'_{11}\beta}}{1+e^{z'_{11}\beta}}$$

An equivalent expression for P^{*}_{12} can be derived in exactly the same way. Since β could be estimated from the available data on fatal accidents, the unconditional probabilities of survival could be calculated.

The parameters in β can be estimated by maximum likelihood. estimation. The likelihood function for the probability of survival in two-vehicle accidents, for example, can be specified as:

$$L = \prod_{k=1}^{K} \frac{\prod_{j=1}^{2} \prod_{i=1}^{n_{jk}} P_{ijk}^{*\;Y_{ijk}} (1 - P_{ijk}^{*})^{1-Y_{ijk}}}{(1 - \prod_{j=1}^{2} \prod_{i=1}^{n_{jk}} P_{ijk}^{*})},$$

where K = 1, …, m, number of accidents;

n_{jk} is the number of individuals in vehicle j, accident k;

Y_{ijk} = 1 if individual i survived, else 0.

The basic structure of the model of the risk of having fatality in an accident is to distinguish between one-vehicle, two-vehicle and multiple-vehicle accidents. The expectation is that the characteristics of drivers contribute more to the probability of having a one-vehicle accident than to a two- or multiple-vehicle accident. On the other hand, vehicle characteristics, particularly the weight relative to the weight of the other vehicle, will affect the survival rate in two-vehicle accidents but may be less important for one-vehicle accidents. In addition, the earlier discussion of why survival rates are likely to differ systematically between one-vehicle and two-vehicle accidents provides another reason for modeling one-vehicle and two-vehicle accidents separately. The justification for separating multiple-vehicle accidents from two-vehicle accidents is that it is impossible to identify the "other" vehicle from the data for multiple-vehicle accidents.

If r is the overall fatality rate, then the model's components can be written as follows:

$$r = [P\{V1\}(1 - P_1 *) + P\{V2\}(1 - P_2 *) + P\{Vm\}(1 - P_m *)]M,$$

where r is the annual fatality rate per occupant;

$P\{V1\}$ is the probability of having a one-vehicle accident per 10,000 miles;

$P\{V2\}$ is the probability of having a two-vehicle accident per 10,000 miles;

$P\{Vm\}$ is the probability of having a multiple (three or more) vehicle accident per 10,000 miles;

P_1^* is the probability of surviving in a one-vehicle accident;

P_2^* is the probability of surviving in a two-vehicle accident;

P_3^* is the probability of surviving in a multiple-vehicle (three or more) accident;

M is the average annual mileage traveled (13,989 miles from the NPTS).

The units for r, $P\{V1\}$, $P\{V2\}$ and $P\{Vm\}$ are all standardized to measure the probability of having a fatal accident per 1000 vehicles.

Conceptually, all six components of the observed values of *r* may be functions of the characteristics of the driver (and the passengers) and the vehicle driven (and the other vehicle for two-vehicle accidents). For computing a hedonic price index, the characteristics of an average driver and passenger are used to predict *r* for different types of vehicle (make, model and year), and each type of vehicle is assumed to have an accident with a typical other vehicle in a two-vehicle accident. Hence, the effects of drivers' characteristics are removed prior to estimating the hedonic price equation. The effect of standardizing the other vehicle in a two-vehicle accident is relatively small because the observed combinations of vehicles in two-vehicle accidents are approximately random. Standardizing drivers' characteristics, however, matters a lot for the probabilities of being in a fatal accident. It is the primary reason for the difference in results for the value of a statistical life compared to a conventional model in which drivers' characteristics are added as additional regressors in the hedonic price equation.

The structure of the equations for the six components of r is described in an appendix which is available from the authors upon request. In summary form, they can be written as follows:

$$P\{V1\} = g_1(V_1, D_1)$$

$$P\{V2\} = g_2(V_1, D_1)$$

$$P\{Vm\} = g_m(V_1, D_1)$$

$$P_1^* = f_1(V_1, O_1)$$

$$P_2^* = f_2(V_1, V_2, O_1)$$

$$P_m^* = f_m(V_1, O_1)$$

where V_1 are the characteristics of a selected vehicle.

D_1 are the average driver's characteristics for the selected vehicle and include factors such as the use of seat belts and whether alcohol was a factor.

O_1 are the characteristics of the occupants of the selected vehicle, including the driver.

V_2 are the characteristics of the other vehicle, its weight relative to the weight of the selected vehicle being the most important.

Since all six dependent variables are probabilities, appropriate statistical models for limited dependent variables are used. P_1^*, P_2^* are specified as logistic functions and estimated by maximum likelihood in GAUSS. For P_m^*, we assume the unconditional probability P_m^* is same as the observed probability P_m, and P_m is specified as a regular logit model and estimated in SAS. $P\{V1\}$, $P\{V2\}$ and $P\{Vm\}$ are determined by a censored regression model to allow for a probability mass at zero (see the appendix for more explanation). (Appendices are available from the authors upon request.) Note that P_m^* is determined by the characteristics of the own-vehicle only because it is not possible to identify the "other" vehicle in a multiple-car accident.

The complete econometric analysis consists of the following five steps:

Step 1. Augment the FARS data on observed fatal accidents with additional characteristics about the vehicles (e.g. weight and safety features), and use these data to estimate equations for the unconditional probabilities of survival in one-vehicle, two-vehicle and multiple-vehicle accidents (P_1^*, P_2^* and P_m^*). Derive the estimated numbers of serious accidents (including accidents with no fatalities) for one-vehicle, two-vehicle and multiple-vehicle accidents.

Step 2. Calculate the average drivers' characteristics in fatal accidents from the FARS data by make, model and year of the vehicle driven, and combine with survey data on the composition of the fleet of vehicles. Use these data to estimate equations for the probabilities of having one-vehicle, two-vehicle and multiple-vehicle accidents by the make, model and year of vehicle (P{V1}, P{V2} and P{Vm}).

Step 3. Use the average drivers' characteristics from the FARS data, and the average other vehicle in two-vehicle accidents, to standardize the unconditional probability of a driver and/or passenger being killed in a fatal accident by make, model and year of the vehicle.

Step 4. Combine the standardized risk of a fatal accident (assuming two occupants) with the data on vehicle characteristics by make, model and year and use these data to estimate hedonic indices of the purchase price and standardized risk and price index the fuel efficiency for other vehicle characteristics.

Step 5. Select a subset of the survey data on vehicle ownership corresponding to families that own only one vehicle. Further subdivide these observations into 1) families with children, 2) families with no children and no seniors, and 3) families with seniors. Calculate the average value of a statistical life for each of the three types of family using the observed make, model and year of the vehicle owned by each family. Two occupants are specified for each vehicle, and for the second type of family, one occupant is a young kid. Then, assuming the VSLs of the adults in the first and second types of family are the same, the average VSL of children can be derived by decomposing the average VSL for each member of the second type of family.

The empirical results for Steps 4 and 5 are presented in the following section, and additional information about Steps 1-3 are given in an appendix which available from the authors upon request.

Section 4. The Empirical Results

4.1 The Hedonic Price and Fuel Efficiency Models

The econometric model used for Step 4 is based on the work of Rosen (1974), Atkinson and Halvorsen (1990), and Dreyfus and Viscusi (1995) on hedonic pricing. Atkinson and Halvorsen (1990) use the data for 112 models of new 1978 automobiles to obtain estimates of the VSL. Since the available fatality data is a function of both the inherent risk of the vehicle and the driver's characteristics, the drivers' characteristics are included in the regression as control variables. Their estimated VSL for the sample as a whole, based on willingness to pay, is \$3.357 million 1986 dollars.

The data used in Dreyfus and Viscusi (1995) differ from those used in earlier studies in that they reflect actual consumer automobile holdings. Dreyfus and Viscusi (1995) use the 1988 Residential Transportation Energy Consumption Survey together with data from industry sources. They generalize the standard hedonic models to recognize the role of discounting on fuel efficiency and safety. The estimates of the implicit value of life range from \$2.6 to \$3.7 million and the estimates of the discount rate range from 11 to 17 percent.

The hedonic price equation for automobiles can be written, following Atkinson and Halvorsen (1990), as follows:

$$P_{auto} = f(R, A),$$

where P_{auto} is the price of an automobile, R is the inherent risk of mortality (a similar measure for injury could also be included) associated with the automobile, and A is a vector of other characteristics. The available mortality rate, F, is a function of both R and a vector of the involved driver's characteristics D. Assuming that F is monotonic in R, the above equation can also be written as:

$P_{auto} = g(F, A, D)$,

The standard functional form used for the estimation of a hedonic price equation is:

$$\log(P_{auto}) = \beta_0 + \sum_i \gamma_i D_i + \sum_k \beta_k \log(X_k) + e$$

where X_k is a representative measured regressor (e.g. horsepower to weight ratio), D_i is a dummy variable for vehicle type, γ_k , β_k are the corresponding parameters and e is an unobserved residual.

A different approach was proposed in the previous section, and it involves predicting the inherent mortality rate using standardized driver's characteristics. In other words, the unobserved values of R are predicted directly. Since the typical number of occupants of a vehicle is two, the observed mortality rate F is twice the size of the average mortality rate per occupant. The corresponding value of R should also reflect the fact that there are two occupants on average. Consequently, the predicted value $\hat{R} = \hat{r}_1 + \hat{r}_2$ (i = 1 is the driver and i = 2 is the passenger), where $\hat{r}_i$ is the predicted probability of a fatality for an individual, defined in the previous section. The standardized inherent mortality rates for two male occupants for year 1995 automobiles are summarized by type of vehicle in Figure 4.1. The minimum, average and maximum risks of mortality for each type of vehicle are illustrated. Meanwhile, Figure 4.2 provides the corresponding scales for the raw (unadjusted) mortality data based on 1996-1997 FARS data. Comparing the two figures, the relative ranking among different types of vehicle are quite consistent, while the standardizing procedure significantly reduce the ranges of the risk of mortality.

One might be surprised by the implication from Figures 4.1 and 4.2 that large sports utility vehicles (SUVs) are not safer than middle size sedans and wagons. From Table 4.1 and Table 4.2, the average standardized and observed risks of mortality show that large SUVs are safer in two-vehicle and multiple-vehicle accidents. However, they are much less safe in one-vehicle accidents because, the probability of having an accident is higher. This point can be further illustrated by the information in Table 4.3. For two-vehicle accidents, large SUVs have the lowest mortality rate per occupant (.186) among all types of vehicle, which is less than half of the rate for middle size cars (.435). This advantage is partially offset by the higher accident rate for large SUVs compared to middle size cars. The impression that large SUVs are safer than other vehicles comes from observing that occupants in a large SUV are more likely to survive in a fatal accident than the occupants of other types of vehicle.

Another cost associated with reducing the risk of mortality and injury is buying more fuel because heavier vehicles are safer but have lower fuel efficiencies. Consequently, a hedonic model of fuel efficiency augments the standard hedonic model of the purchase price in our model. In this model, the cost of additional safety has a capital component and an operating component. In the latter case, the cost penalty corresponds to the reduced fuel efficiency when a heavier vehicle is purchased. The hedonic model of fuel efficiency has the same form as the hedonic model of the purchase price, and it can be written:

$$\log(fe\_city) = \alpha_0 + \sum_i \delta_i D_i + \sum_k \alpha_k \log(X_k) + e$$

where fe_city is the rated miles per gallon for city driving, X_k is a representative measured regressor, D_i is a dummy variable for vehicle type, δ_i , and α_k are the corresponding parameters and e is an unobserved residual.

4.2 The Data

The 1995 National Personal Transportation Survey (NPTS) is used to obtain information on each household's choice of automobiles. The 1995 NPTS was conducted by the Research Triangle Institute (RTI) under the sponsorship of the U.S. Department of Transportation (DOT). The survey covers 42,033 sampled households. A sub-data set of 4036 one-car households holding a 1990-1995 model year vehicle were merged with vehicle attribute data collected from industry and other sources for the same years. The vehicle price data were gathered from *NADA Official Used Car Guide*, and other attribute data were collected from *NADA Official Used Car Guide*, *Ward's Automotive Yearbook*, and *Consumer Reports*. The mortality rate is measured by the number of fatalities occurring in each make/model/year vehicle per 1000 vehicles sold. The number of fatalities is based on the U.S. Department of Transportation's Fatality Analysis Reporting System (FARS) for calendar year 1995-1997. Since the observed mortality rate is jointly determined by the inherent risk associated with the type of automobile and the driver's characteristics and behavior, driver's characteristics are also collected for each make, model and year to provide control variables using FARS 1995-1997 as the source.

In addition to the risk of mortality, a second safety measure, injury rate, is introduced. The injury rate by make and model of vehicle is published annually by the Highway Loss Data Institute. It is measured by the frequency of insurance claims filed under Personal Injury Protection coverages. The raw injury rates are adjusted by the same factors used to standardize raw mortality rates. The implicit assumption is that the "bad" driving characteristics that contribute to fatal accidents also affect injuries. The scatter plots of injury rate vs. mortality are presented in Appendix C. Appendices are available from the authors upon request.

The variables used in the hedonic price equation are summarized in Table 4.4, while Table 4.5 shows the descriptive statistics of selected vehicle attributes. The selection of vehicle attributes and driver's characteristics is similar to Dreyfus and Viscusi (1995) and Atkinson and Halvorsen (1990). It should be noted that the observed mean mortality rate is higher than the standardized mean and the observed standard deviation is also higher. The reason is that the standardized mortality is based on one average male driver and one average male passenger. Even though average values of the other regressors are used, the elimination of young drivers, for example, results in lower average mortality rates. The effect of standardizing drivers' characteristics to predict the inherent mortality rate has the effect, as expected, of reducing the variability of mortality among vehicles.

4.3 The Estimated Hedonic Models

Least square estimates of the hedonic price model and the fuel efficiency model are presented in Table 4.9. Model A is the hedonic equation of fuel efficiency, using the standardized mortality rate. Model B is the hedonic equation of capital cost, using the standardized mortality rate. In Model A and B, variables with small t ratios and perverse signs have been dropped.

The most important parameter for computing the VSL is the coefficient for the mortality rate, and the values in Model A and B have the right signs and are both significant. In other hedonic price models, fuel efficiency is included as a regressor in Model B, but it often has a large t ratio and a perverse negative sign (fuel efficiency is a positive attribute). Hence, some explanation is needed to explain why fuel efficiency is omitted in Model B. The implication of Model A is that fuel efficiency is a dependent variable, like the price, and is a function of the vehicle's characteristics. The model corresponds to a simplified reduced form for a system of two equations. If the predicted fuel efficiency from Model A is used as a regressor in Model B, the coefficient has a logical positive sign. The overall effect on the estimated VSL is small, however, if the direct effects of mortality on price and fuel efficiency are combined with the indirect effect on the price through fuel efficiency. This is not really surprising because the model presented in Table 4.6 is equivalent to a solved reduced form for a structural model which has fuel efficiency as a regressor in the hedonic price equation (the equation for fuel efficiency remains the same).

4.4 Estimates of VSL

The standard expression for determining the VSL from the hedonic price model for any make, model and year of vehicle, without the operating cost component, is:

$$VSL = \beta_m \left(P_{auto} / R \right) / \left(\sum_{t=1}^{L} \left(\frac{1}{1+i} \right)^t \right),$$

The expression represents the marginal change in the annualized capital cost for a reduction of one fatality. Adding the annual operating cost component, the full estimate of VSL for any make, model and year of vehicle is:

$$VSL = \beta_m \left(P_{auto} / R\right) / \left(\sum_{t=1}^{L} \left(\frac{1}{1+i} \right)^t \right) + P_{gas} (\alpha_m / fe\_city)(M / R),$$

where VSL is the average VSL for a household,

β_m is the coefficient for mortality in Model B,

P_{auto} is the purchase price of the vehicle,

R is the standardized mortality rate for a one-vehicle household,

i is the discount rate, set to 10 percent,

L is the expected vehicle life, set to 10 years,

P_{gas} is the average gasoline price in year 1995, set to $1.205 per gallon,

α_m is the regression coefficient for mortality in Model A,

fe_city is the city fuel efficiency in miles per gallon,

M is the average annual miles driven.

Determining the mortality rate R

The mortality rate for each one-vehicle household can be written as follows:

$$R = \sum_{i=1}^{n} m_i f_i$$

where i = 1, 2, …, n refers to individual i

m is the number of miles traveled as an occupant of the vehicle

f is the unconditional risk of fatality per mile.

The procedures used to determine f_i are exactly the ones described in the previous section for estimating the inherent mortality rates for different vehicles. The main difference is that the inherent mortality rates are based on the same set of characteristics for a driver and a passenger for every vehicle. In contrast, the estimated VSL for a single-vehicle household is based on the actual age composition of each household. To complete the determination of R, it is necessary to specify how many miles each member of a household rides in a vehicle (i.e. the mileage weights α).

There are two different types of variable in the NPTS data set that can be used to provide information about the mileage weights. The first is to compare average distances driven per vehicle for different types of household. It is interesting to note that the average distances driven for different types of vehicle are quite similar. Furthermore, the average distances per vehicle are similar for households with one vehicle or with more than one vehicle. In other words, the total distance driven by a household is roughly proportional to the number of vehicles owned. Nevertheless, there is an important difference in the distances driven per vehicle for different households. It is that households with seniors tend to drive less than other types of household.

The average annual mileages per vehicle are summarized in Table 4.7 for three different types of single-vehicle household. The assumption used to estimate the VSL is that Type 3 households with seniors drive 8.87/12.614 = 0.70 of the distance driven by Type 1 households. (Note that this lower distance driven by seniors partially offsets the lower survival rates in accidents for seniors). Type 2 households are given the same weight as Type 1 households because the difference between the average distances driven is relatively small.

The second type of information in the NPTS data set gives the distribution of ages of all members of a household in the sample. The results, summarized in Table 4.8, show that 24% of all household members are kids (0-15 years old), 65% are adults (16-64) and 11% are seniors (>65). Comparing these values with the corresponding age distribution of occupants of vehicles in accidents (i.e. information from the FARS data set) shows that only 11% of occupants are kids, 80% are adults, and 9% are seniors. (Note that this age distribution, identified as “FARS adjusted” in Table 4.8, is based on the estimated unconditional number of accidents and not just on the observed number of fatal accidents. This increases the relative importance of one-vehicle accidents with a driver only, for example, because the difference between the conditional and unconditional probabilities is greatest for this type of accident. Separate adjustments are made for each type of accident (one-vehicle, two-vehicle, and multiple-vehicle accidents) for the sub-samples used to estimate the hedonic models in the previous section. The results for each type of accident are then prorated to the full sample.) The main implication is that kids and seniors are underrepresented as occupants of vehicles in accidents and adults are overrepresented relative to the age distribution of household members in the NPTS data set. One reason for this difference is that over 60% of the vehicles involved in accidents have a driver only and no passengers. Hence, the mileage weights for passengers must be substantially lower than they are for drivers.

Estimates of the mileage weights for different ages in a household will be available later this year after a national survey of vehicle usage has been completed for the EPA. Given available information at this time, a relatively simple weighting scheme is adopted for kids, adults, and seniors.

Since the average annual mileages per vehicle are similar for households with one vehicle and more than one vehicle, each vehicle in the NPTS data set is assigned a driver (an adult or a senior) in a household. Hence, the number of assigned drivers in a household is less than or equal to the number of vehicles. Drivers are assigned a relative mileage weight of one. Additional household members who are not identified as drivers are given a relative mileage weight of .392 (corresponding to the observed proportion of vehicles in the FARS data set with passengers). Weighting all passengers, including kids, by 0.392 gives the proportions shown in Table 4.8 as "NPTS adjusted 1", and the adjusted proportion of kids is now much lower than the observed proportion.

A second adjustment to the NPTS proportions accounts for the lower number of miles driven by seniors. Weighting the proportion for seniors in "NPTS adjusted 1" by .703, and rescaling the three proportions to add to one, gives "NPTS adjusted 2". Comparing these proportions with "FARS adjusted" shows that the two sets of adjusted proportions for accidents (FARS adjusted) and households (NPTS adjusted 2) are very close to each other. In other words, converting all members of households to adult driver equivalents reconciles the major differences between the age distribution of the occupants of vehicles in the FARS data set and the observed age distribution of households in the NPTS data set.

Even though the adjusted age distributions for accidents (FARS adjusted) and for households (NPTS adjusted 2) are quite similar, the adjusted value for kids in households is still 10% larger than the adjusted value in accidents. Hence, the relative mileage weight for kids is reduced further to make the ratio of kids in FARS adjusted to NPTS adjusted 2 equal to the corresponding ratio for adults. The final weight is 0.345 = .392(.111/.124) x (.788/.801). For adults and seniors, the same weights used to adjust the NPTS data in Table 4.8 are adopted to determine the relative mileage weights for the household mortality rate, R. If the mileage weight for an adult driver is m_1, the relative mileage weights (m_i/M_1) used to determine R are:

adult drivers	1.00
other adults	0.39
senior drivers	0.70
other seniors	0.70 x 0.39 = 0.27
kids	0.35

Determining the average VSL by household type

In order to calculate the VSL, a simulation is conducted for each one-vehicle household based on the estimated mortality rates, R. (The average annual mileage driven by an adult driver is M = 13,989 from the NPTS data set.) To simplify the simulation, the differences associated with gender were not considered. The driver of each household is assumed to be a male with average driving characteristics (e.g. average alcohol involvement, etc.). Adult and senior passengers are assumed to fill the front seat first, while kids always sit in the back. Using these assumptions, the means of the estimated VSL for different types of household are presented in Table 4.9. The main implications of the values of VSL in Table 4.10 are 1) the operating cost component is relatively small compared to the capital cost component, and 2) the VSL of $4 million for Type 1 households, with no kids and no seniors, is over a third larger than the VSL for the other two types of household. One reason for this latter result is that incomes for Type 1 households are relatively high. Hence, adjusting for income differences is the next task.

Standardizing incomes for different types of household

Household income is reported in the NPTS data set, and the average values for different types of household are shown in Table 4.10 together with the average household sizes. Accounting for different sizes of household is the first step in standardizing income. To account for the shared benefits experiences by households with more than one member (e.g. lower housing costs per capita), household income is converted to income per adult equivalent. The standard weights adopted by the U.S. Bureau of the Census are used (see Table 4.10 for the values). Using this measure, the income per adult equivalent for one-vehicle households is over $33 thousand for households with no kids and no seniors (Type 1), but only $20 thousand for households with kids (Type 2). (Note that the difference would be even greater using income per capita.) The corresponding incomes for the full sample of all households are also given in Table 4.10. The main implications are that the income level for households with kids is higher for the sample of all households, but the income levels for Type 1 and 2 households in one-vehicle households are similar to the corresponding levels for all households.

Using income per adult equivalent as the measure of income for each type of household, it is straightforward to standardize income to the average value for all households ($28 thousand). Since VSL is almost certainly positively related to income, the differences in the values of VSL for different types of household in Table 4.9 are generally consistent with the differences in income in Table 4.10. Nevertheless, there is no unique way to adjust the value of VSL for a specified change of income.

Theory suggests that VSL is a function of wealth, which in turn is a function of income. If the income elasticity of the VSL is one, any percentage increase of income would give the same percentage increase of VSL. In contrast, Blomquist has argued that utility may include a non-wealth component. Since the VSL for an individual is mostly determined by the ratio of total utility to the marginal utility of consumption, the non-wealth component of utility will increase the VSL but reduce the size of the income effect. Blomquist's analysis suggests that an appropriate income elasticity may be 0.3 (1979).

The results in Table 4.11 are derived using the following formula to computed adjusted values of VSL:

$$VSL_i^* = VSL_i(\bar{I} / I_i)^\theta$$

Where i is the type of household

VSL_i is the value in Table 4.9

I_i is the income per adult in Table 4.10

$\bar{I}$ = \$28 thousand is the average income per adult for all households

VSL_i^* is the adjusted VSL

The adjusted values of VSL in Table 4.11 show the importance of the income elasticity. For both values of θ, the values of VSL are closer together compared to the unadjusted values. When θ = 1, the VSL for households with kids (Type 2) is larger than the values for the other types of household. The values of VSL for households with seniors (Type 3) are always the lowest. This is consistent with the life-cycle model of Viscusi that implies values of VSL decline with age.

Estimating the VSL for different age groups

The average VSL for adults (Type 1 households) and for seniors (Type 3 households) are estimated directly in Table 4.11, but the VSL for Type 2 households combines the values for kids and adults. The next step is to estimate a VSL for kids.

The basic economic and demographic characteristics for different types of household in the NPTS survey are listed in Table 4.10 for one-vehicle households and for the complete sample of all households. The average household size and the average number of adults in each type of household are calculated from the NPTS survey. Assuming the VSL of an adult in Type 1 and 2 households are the same, the VSL for kids can be derived from the values of VSL in Table 4.11 by decomposing the average VSL for a Type 2 household as follows:

$$\mathrm{VSL}_{\mathrm{adults}}\left(\frac{\#\text{ of adults of type 2}}{\text{household size of type 2}}\right) + \mathrm{VSL}_{\mathrm{kids}}\left(1 - \frac{\#\text{ of adults of type 2}}{\text{household size of type 2}}\right) = \mathrm{VSL}_{\mathrm{type\,2}}$$

The results in Table 4.12 summarize VSL values for kids, adults and seniors. They show that the VSL is very sensitive to the value of the income elasticity θ. When $\theta = 1$, the average VSL for kids is the largest among the three age groups, but when $\theta = .3$, it is the lowest. (When $\theta = .65$, the average VSL for kids is same as the VSL for adults.) The average VSL for seniors is lower than the VSL for adults for both values of θ. Even though seniors drive relatively safe vehicles and fewer miles, the additional safety is more than offset by higher fragility in accidents.

Adjusting for adult drivers' perception of risk

A final modification to the values of VSL is to account for the bias in drivers' perceptions about their driving ability and safety on the road (see Blomquist, Miller and Levy). Roughly 80 percent of drivers think that their driving skill is above average (i.e. their subjective probability of having a fatal accident is lower than the true probability by a factor of approximately .8/.5 = 1.6). The two modified VSL in Table 4.12 correspond to the two adjusted VSL x 1.634 following Blomquist et al. With this modification, VSL values of almost $8 million are obtained (for kids using modification 1).

The average ages of kids, adults and seniors in the NPTS data set are 7.8, 39.0 and 72.8 years, respectively. These representative ages are used in Figure 4.3 to plot the adjusted and modified VSL against age. Figure 4.3 provides additional evidence that the VSL for adults and seniors is consistent with the life-cycle model proposed by Moore and Viscusi (1988). It also shows how sensitive the VSL for kids are to assumptions about the income elasticity.

Section 5. Conclusions

Our analysis in the preceding sections, while encouraging for our forthcoming national survey of automobile usage, points out some important potential difficulties.

First, from the theoretical model of Section 2, it is apparent that we must collect very detailed data on usage by individuals, by automobile type, to estimate fraction of usage by age (child, adult, or senior) for multiple car families.

Second, since risk differs depending on seating position, these data must be collected as well.

Finally, considerable theoretical speculation exists that the value of a statistical life should differ by age. We find this hypothesis is consistent with our preliminary analysis in that the elderly have a lower VSL than adults.

However, our theoretical work is the first on children and suggests that the VSL of the young should be similar or slightly less than adults. To illustrate this point, consider the simplified intertemporal model of Moore and Viscusi (1988) who suggest that the VSL at age t is equal to the value of a life-year times the discounted present value of remaining life years, or $\overline{V}\left(1-e^{-\delta R(t)}\right)/\delta$, where $\overline{V}$ is the value of a life-year (assumed constant over time), δ is the discount rate, and R(t) is the remaining years of life at age t. For example, using average life expectancy for 1992 from the National Center for

Health Statistics, and an interest rate of 10%, the value of $\overline{V}$ is multiplied by 10 for an eight year old child, 9.8 for a 39 year old adult, and 7.1 for a seventy-three year old. Our best estimates of the VSL for adults (modified 2 in Table 4.12) of \$6.34 x 10^6 implies that $\overline{V}$ = \$646,939 in the Moore and Viscusi formula. Using this formula implies that a child's VSL would be only slightly more than an adult's (\$6.47 x 10^6 vs. \$6.34 x 10^6) and that the elderly value their lives somewhat less than adults (\$4.59 x 10^6 vs. \$6.34 x 10^6). The values of VSL for the quantity adjusted model are also shown in Table 4.12. Clearly our best estimates of the VSL for children and the elderly both fall below those implied by the predicted quantity adjusted value of life model. However, the latter model was developed for working adults. Note from our theoretical model, the VSL shown in Equation (8) is equal to the monetized value of utility plus the wage minus consumption. In an intertemporal model, this expression is the equivalent of $\overline{V}$ as defined above. Thus, since wage minus consumption is positive for adults, but negative for children and the elderly, it is unsurprising that the Moore and Viscusi model overpredicts the VSL of children and the elderly.

However, this interpretation depends on the assumption of an income elasticity of .3 for the VSL incorporated into the "Modified 2" values in Table 4.12. With new primary data, which will allow inclusion of multiple car, higher income families with children, we hope to be able to directly estimate values for similar income groups as well as estimate the income elasticity of the VSL. In addition, we hope to estimate a full life cycle model of the VSL in a family.

References:

Agee, M.D., and T.D. Crocker. (1996). "Parental Altruism and Child Lead Exposure: Inferences from the Demand for Chelation Therapy." The Journal of Human Resources 31:677-691.

Atkinson, Scott E and Robert Halvorsen (1990). “The Valuation of Risks to Life: Evidence from the Market for Automobiles.” The Review of Economics and Statistics 72 (1): 133-136.

Becker, G. S. (1974) “A Theory of Social Interactions,” Journal of Political Economy, 82, 1095-1117.

Becker, G. S. (1991) A Treatise on the Family, Harvard University Press, enlarged edn.

Bergstrom, T. C. (1996) “Economics in a Family Way,” Journal of Economic Literature, 34(4), 1903-1934.

Blomquist, Glenn C., (1979). “Value of Life Saving: Implications of Consumption Activity,” The Journal of Political Economy 87(3): 540-558.

Blomquist, Glenn C., David Levy and Ted R. Miller (1996) “Values of Risk Reduction Implied by Motorist Use of Protection Equipment: New Evidence from Different Populations,” Journal of Transport Economics and Policy 30 (January 1996): 55-66.

Chestnut, L. and W. Schulze. (1998). “Valuing the Long Term Health Risks From Wartime Toxic Exposures,” Proceedings of the First International Conference on Addressing Environmental Consequences of War.

Citro, Constance F. and Robert T. Michael. (1995). *Measuring Poverty: a New Approach.* National Academy Press.

Cropper, M.L. and A.M. Freeman III. (1991). "Environmental Health Effects." Measuring the Demand for Environmental Quality. J.B. Braden and C.D. Kolstad
(ed.) North-Holland. New York.

Dreyfus, Mark K. and W. Kip Viscusi (1995). "Rates of Time Preference and Consumer Valuations of Automobile Safety and Fuel Efficiency." Journal of Law and Economics, vol XXXVIII (April 1995): 79-105.

Greene, William H. (1997). *Econometric Analysis*. 3rd Edition. Prentice-Hall, Inc.

Harrison, D. and A.L. Nichols. (1990). Benefits of the 1989 Air Quality Management Plan for the South Coast Air Basin: A Reassessment. Prepared for California Council for Environmental and Economic Balance by National Economic Research Associates, Inc., Cambridge, Massachusetts.

Jenkins, Robin, Nicole Owens, and Lanelle Bembenek Wiggins (1999) "The Value of a Statistical Child's Life: The Case of Bicycle Helmets." USEPA Office of Economy and Environment.

Jones-Lee, M.W., M. Hammerton, and P.R. Philips. (1985). "The Value of Safety:
Results of a National Sample Survey." The Economic Journal 95(March):49-72.

Joyce, T.J., M. Grossman, and F. Goldman. (1989). "An Assessment of the Benefits of Air Pollution Control: The Case of Infant Health." Journal of Urban Economics 25:32-51.

McElroy, M. and M. Horney (1981) "Nash-bargained decisions: Toward a Generalization of the Theory of Demand," International Economic Review 22:333-349.

Miller, T.R. (1989). "Willingness to Pay Comes of Age: Will the System Survive?" Northwestern University Law Review 83:876-907.

Moore, M. J. and W. K. Viscusi (1988). "The Quantity-Adjusted Value of Life." Economic Inquiry 26:369-388.

NADA official used car guide (1995) Vol. 62. McLean, Va., National Automobile Dealers Used Car Guide Co.

Powell, James L. (1986). "Symmetrically Trimmed Least Square Estimation for Tobit Models." Econometrica 54 (6): 1435-1460.

Rosen, Sherwin (1974). "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." Journal of Political Economy 82 (1): 34-55.

Shepard, D.S. and R.J. Zeckhauser. (1982). "Life-Cycle Consumption and Willingness to Pay for Increased Survival." in M.W. Jones-Lee (ed.). The Value of Life and Safety. New York: North-Holland.

U.S. Bureau of the Census (1993). Current Population Reports, Series P60, No. 185, *Poverty in the United States: 1992*. U.S. Department of Commerce, Washington, D.C..

Ward's Automotive Yearbook (1990-1996). Vol. 52-58.

Table 4.1: The Standardized Risk of Mortality by Vehicle and Type of Accident

Vehicle Type	Total Risk	One-Car Accidents	Two-Car Accidents	Multiple-Car Accidents
small sedans & wagons	9.2	3.4	4.4	1.4
middle sedans & wagons	6.9	3.3	2.5	1.0
large sedans & wagons	6.5	3.5	2.1	0.8
luxury sedans & wagons	7.2	4.7	1.7	0.8
small & mid. specialties	9.5	5.6	3.0	1.0
luxury sports	25.3	21.8	2.6	0.9
small suv	17.1	12.0	3.6	1.6
large suv	9.4	7.1	1.6	0.7
van (minivan)	5.0	2.7	1.5	0.8
small pickup	12.4	7.7	3.5	1.2
large pickup	8.6	5.8	2.0	0.8

Table 4.2: The Observed Risk of Mortality by Vehicle and Type of Accident (Year 1996-1997 Average)

Vehicle Type	Total Risk	One-Car Accidents	Two-Car Accidents	Multiple-Car Accidents
small sedans & wagons	30.8	12.2	14.1	4.6
middle sedans & wagons	22.5	12.1	8.8	1.6
large sedans & wagons	17.7	5.4	11.2	1.2
luxury sedans & wagons	9.3	4.9	3.2	1.2
small & mid. specialties	33.8	20.5	10.3	3.0
luxury sports	26.2	23.6	2.6	0.0
small suv	53.4	26.6	25.6	1.2
large suv	21.1	16.2	3.5	1.4
van (minivan)	24.6	12.8	8.8	3.0
small pickup	26.1	17.6	7.2	1.4
large pickup	17.6	11.6	4.8	1.2

Table 4.3: The Observed Mortality Rates Per Occupant and Accident Rates per 1000 Vehicles for Fatal Two-vehicle Accidents (Average 1996-1997)

Vehicle Type	Mortality Rate	Accident Rate
small sedan & wagons	0.512	0.193
middle sedan & wagons	0.435	0.169
large sedan & wagons	0.370	0.159
luxury sedan & wagons	0.369	0.113
small & mid. specialties	0.429	0.171
luxury sports	0.329	0.109
small suv	0.430	0.189
large suv	0.186	0.193
van (minivan)	0.218	0.221
small pickup	0.368	0.188
large pickup	0.201	0.244

Table 4.4: Variable Definitions

Variable Name	Definition
Price	Vehicle price as of end-of-year 1995.
Value Retained	Original sales value retained, as of end-of-year 1995.
Mortality Rate, Observed	Number of fatalities occurring in that make/model/year vehicle per 1000 of that vehicle sold.
Mortality Rate, Standardized	Predicted number of fatalities in that make/model/year vehicle per 1000 of that vehicle sold with average 2 occupants.
Injury Rate	An Index based on the frequency of insurance claims. The lower, the safer.
CityFuel efficiency	Miles per gallon in city area.
CityFuel efficiency Predicted	Predicted Miles per gallon in city area.
Reliability Rating	A discrete variable coded from 1 to 5, 5 is the highest while 1 is the lowest.
Acceleration	The horsepower-to-weight ratio.
Traditional Styling	Length plus width divided by height.
ClassX	Discrete variables coded as 1 for the appropriate class. Class1 to class7 represent small, middle, large, luxury, SUV, van, and pick-up truck, respectively.
YearXX	Discrete variables coded as 1 for the vehicle model year.
Young Driver	Proportion of fatalities in this make/model/year vehicle in which the driver was younger than 25 years.
Older Driver	Proportion of fatalities in this make/model/year vehicle in which the driver was 65 or older.
Alcohol	Proportion of fatalities in this make/model/year vehicle in which the alcohol involvement was reported.
Gender of Driver	Proportion of fatalities in this make/model/year vehicle in which the driver was male.
Seat Belt	Proportion of fatalities in this make/model/year vehicle in which the driver was wearing a seat belt.
Previous Offenses	Proportion of fatalities in this make/model/year vehicle in which the driver had no previous offense.
Late Night	Proportion of fatalities in this make/model/year vehicle which occurred between 12:00am to 5:59am.
One-car Accident	Proportion of fatalities in this make/model/year vehicle in which only one vehicle was involved.
Ford, GM, Chrysler, Germany, Japan	Discrete variables coded as 1 for the manufacturer and 0 otherwise.
MB	Dummy variable coded as 1 for Mercedes Benz, 0 otherwise.

Table 4.5: Summary Statistics of Selected Variables

Variable	Mean	Standard Deviation
Price	15703.53	9371.57
Value Retained	0.7720	0.1753
Mortality Rate, Observed	0.1345	0.0994
Mortality Rate, Standardized	0.0939	0.0401
Injury Rate	73.72	42.12
City Fuel-efficiency	20.26	4.82
Reliability Rating	3.019	1.321
Acceleration	0.0475	0.0102
Traditional Styling	4.451	0.519

Table 4.6: Parameter Estimates for the Hedonic Equations

Variable	Model A Estimated Coefficient	t ratio	Model B Estimated Coefficient	t ratio
Dependent	**Fe_city**		**P_{auto}**	
Constant	2.5689	14.13	7.7174	25.45
Value Retained	0.0549	3.35	0.4594	11.10
Mortality Rate	0.0258	1.99	-0.0690	-3.53
Injury Rate	0.0330	4.01	-0.0161	-1.31
Reliability Rating	0.0170	5.05	0.0617	5.23
Acceleration	-0.2290	-8.04	0.6014	13.99
Traditional Styling	-0.2786	-5.21	0.6035	7.56
Class2	-0.1873	-16.56	0.2426	14.34
Class3	-0.2751	-14.69	0.3734	13.28
Class4	-0.2852	-19.29	0.6752	29.76
Class5	-0.6397	-37.47	0.8127	31.94
Class6	-0.4846	-24.84	0.6558	22.67
Class7	-0.4352	-27.49	0.3398	14.31
Year91			0.1137	6.31
Year92			0.2100	10.53
Year93			0.2977	13.16
Year94			0.3880	15.30
Year95			0.4474	16.14
Ford	0.0347	1.90	-0.0972	-3.58
GM	0.0334	1.94	-0.0879	-3.44
Chrysler	0.0196	1.12	-0.1148	-4.43
Germany	-0.0562	-2.84	0.1489	5.05
Japan	0.0470	2.73	-0.0430	-1.71
MB	-0.0078	-0.33	0.5237	14.89
R^2	0.7626		0.8996	

Table 4.7: Average annual mileage for single-vehicle households:

Household Category	1000 miles
Grand Mean	12.327
Type 1	12.614
Type 2	13.410
Type 3	8.870

Type 1: Household with no one retired and no kids.
Type 2: Household with kids.
Type 3: Household with a retired member and no kids.

Table 4.8: Age distributions of vehicle occupants and households:

	Kids	Adults	Seniors
FARS observed	.113	.785	.102
FARS adjusted[a]	.111	.801	.088
NPTS observed	.241	.646	.113
NPTS adjusted 1[b]	.119	.759	.122
NPTS adjusted 2[c]	.124	.788	.088

[a] to account for accidents with no fatalities.
[b] weighting passengers by .392 and drivers by 1.
[c] adjustment (1) plus weighting seniors by .703 to account for lower annual miles driven

Table 4.9: Estimated Values of a Statistical Life (VSL) for Different Types of Households

Household Category	Capital cost component ($millions)	Operating cost component ($millions)	VSL ($millions)
Grand Mean	3.09	0.38	3.47
Type 1	3.64	0.45	4.09
Type 2	2.63	0.33	2.97
Type 3	2.42	0.30	2.72

Type 1: Household with no one retired and no kids.
Type 2: Household with kids.
Type 3: Household with a retired member and no kids.

Table 4.10: Economic and Demographic Characteristics of Households

Household Category	Average Household Size	Average number of Adults	Average Number of Adult Equivalents*	Average Household Income ($1000)	Average Income Per Adult ($1000)
One-Vehicle Households					
Grand Mean	1.9	1.5	1.278	33.649	27.743
Type 1	1.4	1.4	1.103	36.378	33.436
Type 2	3.5	1.7	1.795	35.380	20.430
Type 3	1.6	1.6	1.174	26.934	23.025
All Households					
Grand Mean	2.7	1.9	1.516	41.080	28.007
Type 1	1.9	1.9	1.273	42.725	33.570
Type 2	4.0	2.0	1.964	45.269	23.607
Type 3	1.8	1.8	1.229	29.110	23.264

*The equivalence scale is based on the official weighted average poverty thresholds for 1992 (Data Source: Bureau of the Census (1993: Table A)), following the Table 3-1 of Citro and Michael (1995). The values of the equivalence scale are 1, 1.279, 1.566, 2.007, 2.323, 2.679, 3.023, 3.367 and 4.024 for family size 1, 2, 3, 4, 5, 6, 7, 8, and 9 or more, respectively.

Type 1: Household with no one retired and no kids.
Type 2: Household with kids.
Type 3: Household with a retired member and no kids.

Table 4.11: Values of VSL Adjusted for Income

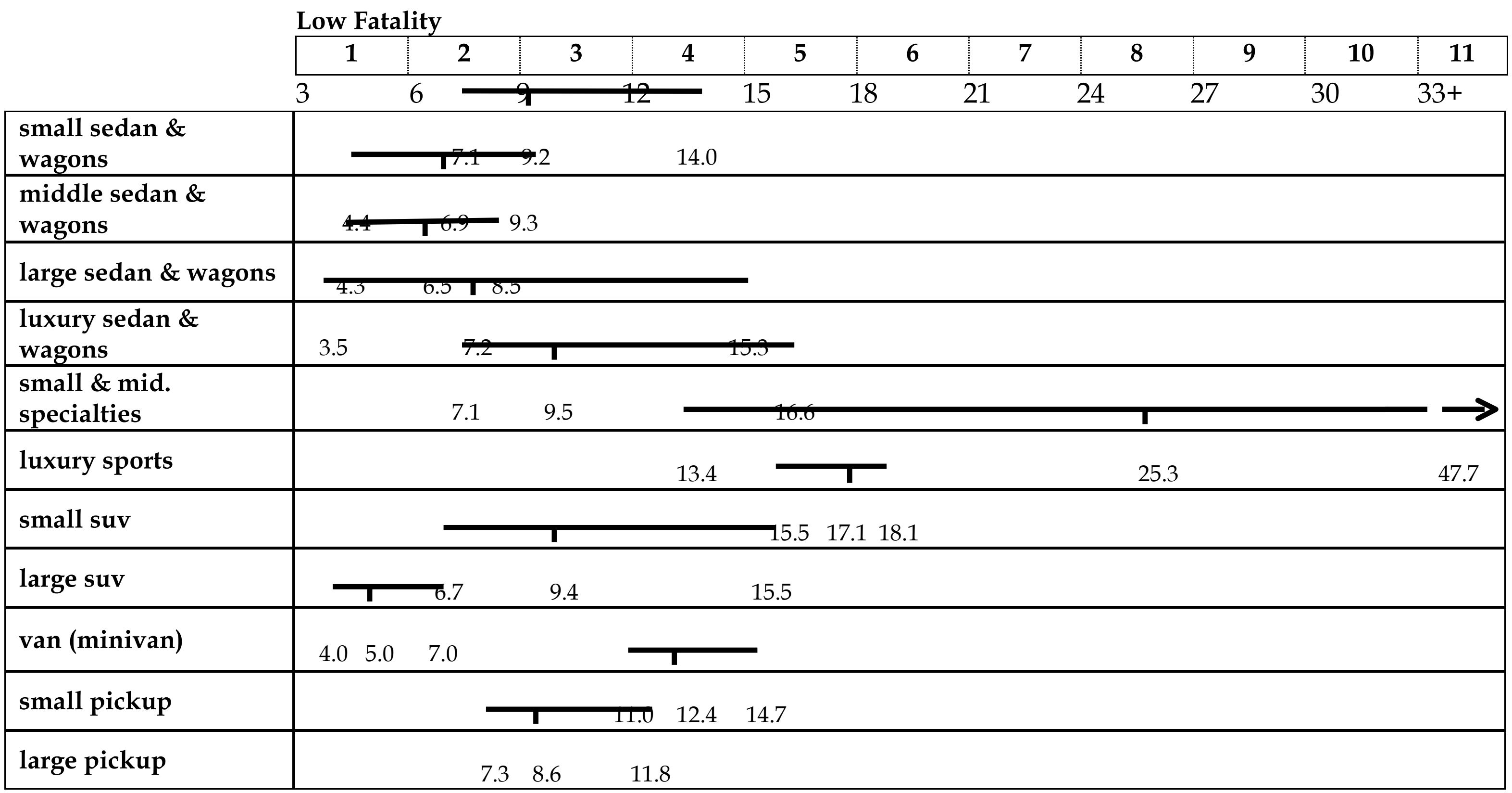
Household Category	VSL (Table 4.9) $ million	VSL* $\theta = 1.0$ $ million	VSL* $\theta = 0.3$ $ million
Grand Mean	3.47	3.51	3.48
Type 1	4.09	3.42	3.88
Type 2	2.97	4.07	3.26
Type 3	2.72	3.31	2.89

Type 1: Household with no one retired and no kids.
Type 2: Household with kids.
Type 3: Household with a retired member and no kids.

Table 4.12: Estimated Values of a Statistical Life (VSL) for Different Age Groups

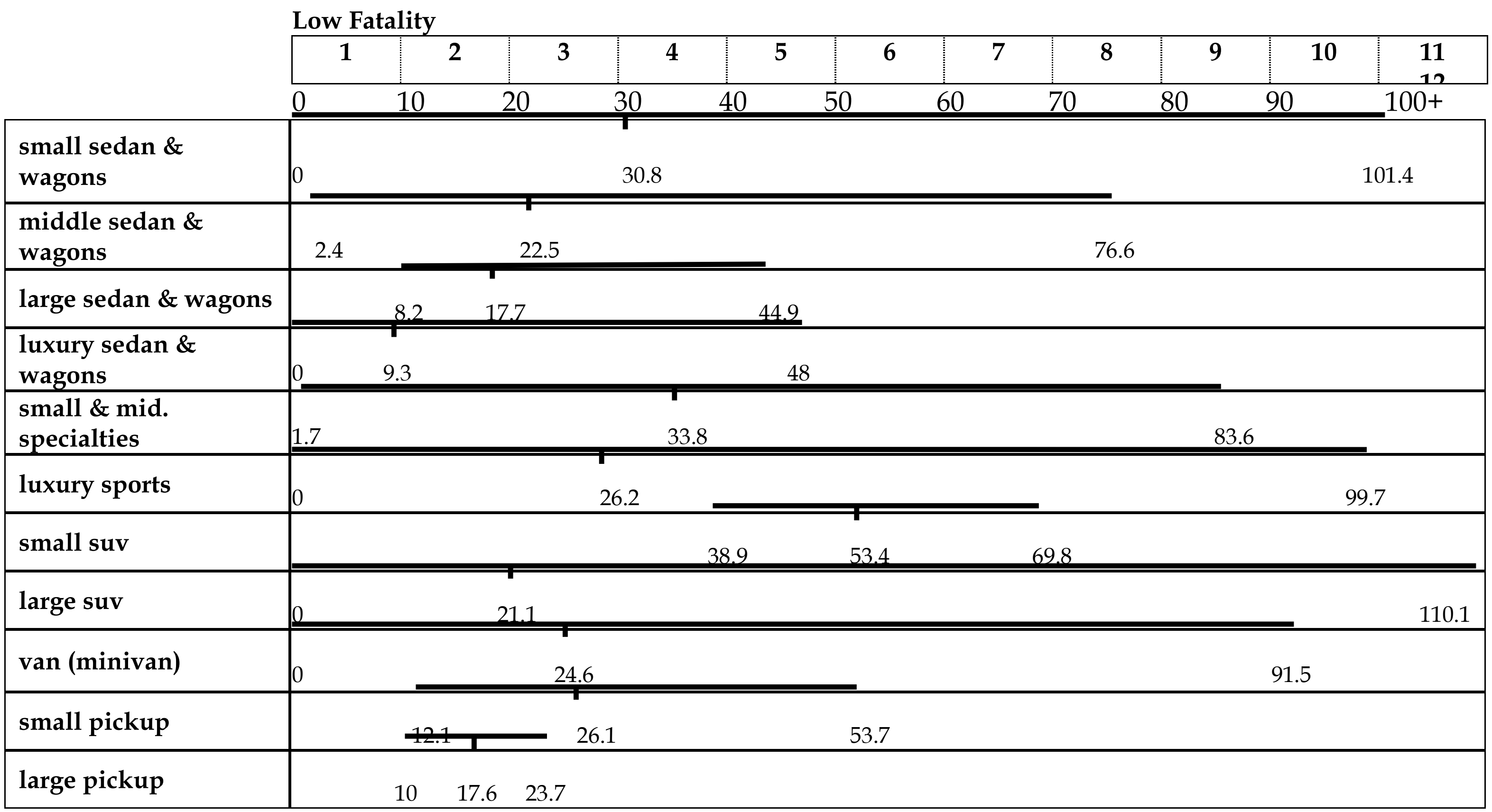
VSL ($Millions)	Kids	Adults	Seniors
Adjusted for Income			
Adjusted 1	4.73	3.42	3.31
Adjusted 2	2.62	3.88	2.89
Adjusted for Risk Perception			
Modified 1	7.74	5.60	5.42
Modified 2	4.28	6.34	4.72
Quantity Adjusted Model	6.47	6.34	4.59
Age	7.8	39	72.8

Figure 4.1: Standardized Scales for the Risk of Mortality



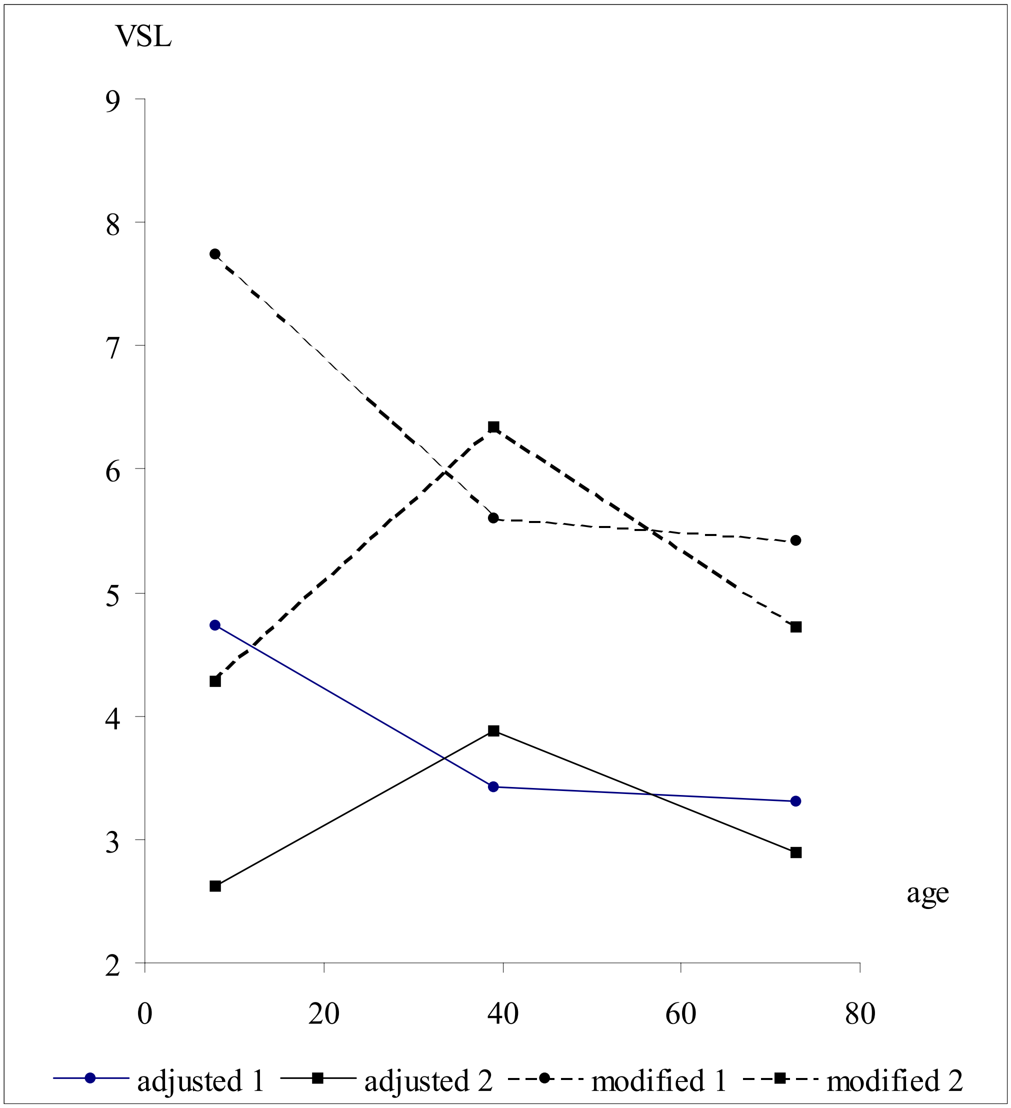
Note: The scale is based on predicted total fatalities per 100,000 vehicles (1995 model year) per 10,000 miles driven with 2 occupants.

Figure 4.2: Unadjusted Scales for the Risk of Mortality



Note: The scale is based on the observed total fatalities in year 1996-1997 per 100,000 vehicles (1995 model year) on road per 10,000 miles driven (average annual miles driven is 13989 miles).

Figure 4.3: The Value of a Statistical Life by Age



Definitions are provided in Tables 4.11 and 4.12.

Appendix A: Econometric Models Used to Estimate the Mortality Rate for Vehicles

The inherent mortality rate for an individual in a vehicle can be decomposed as follows:

$$r = \left[P\{V1\}(1 - P_1 *) + P\{V2\}(1 - P_2 *) + P\{Vm\}(1 - P_m *)\right]M ,$$

where r is the annual fatality rate per capita;

P{V1} is the probability of having a one-vehicle accident per 10000 miles;

P{V2} is the probability of having a two vehicle accident per 10000 miles;

P{Vm} is the probability of having a multiple (three or more) vehicle accident per 10000 miles;

P_1* is the probability of surviving in a one-vehicle accident;

P_2* is the probability of surviving in a two-vehicle accident;

P_3* is the probability of surviving in a multiple-vehicle (three or more) accident;

M is the average annual mileage traveled (13989 miles from the NPTS).

The likelihood function for the probability of survival can be specified as:

One-car accidents:

$$L = \prod_{k=1}^{K} \frac{\prod_{i=1}^{n_k} P_{ik}^{*\,Y_{ik}} (1 - P_{ik}^{*})^{1-Y_{ik}}}{(1 - \prod_{i=1}^{n_k} P_{ik}^{*})}$$

Two-car accidents:

$$L = \prod_{k=1}^{K} \frac{\prod_{j=1}^{2} \prod_{i=1}^{n_{jk}} P_{ijk}^{*\ Y_{ijk}} (1 - P_{ijk}^{*})^{1-Y_{ijk}}}{(1 - \prod_{j=1}^{2} \prod_{i=1}^{n_j} P_{ijk}^{*})}$$

Multiple-car accidents:

$$L = \prod_{k=1}^{K} \prod_{j=1}^{m_k} \prod_{i=1}^{n_{jk}} P_{ijk}^{\ Y_{ijk}} (1 - P_{ijk})^{1-Y_{ijk}}$$

where i = 1, …, n, individuals;

j = 1, …, m, vehicle;

k = 1, …, K, accidents;

Y_{ijk} = 1 if survived, else 0.

Survival rates P_1* and P_2* are specified as logit form and estimated by maximum likelihood in GAUSS using data from the FARS augmented with additional data about vehicle characteristics (step 1 in section 3). The explanatory variables are summarized in Table A4.1, and the estimated equations are shown in Tables A4.2 and A4.3, respectively. For survival rate in the one-vehicle accidents, the effect of using a restraint (seat belt or car seat) is very important and clearly positive, but the effect of an airbag was not significant. The number of occupants is significant, but without a clear explanation. The survival rate is relatively high in pickup trucks (Class 7). The very inexperienced driver, 16 years or younger, has a strong negative effect on the survival rate.

The equation for P_2* in Table A4.3 implies that the weight ratio is the most important explanatory variable. Being in a larger vehicle increases the chance of survival and visa-versa. The number of occupants is also important. The positive effect of using a restraint (seat belt or car seat) is substantially larger than the effect of airbags. In general, the effects of the class of vehicle are consistent with the effect of the weight ratio. Seating in a small vehicle (Class1) reduces the probability of survival, while hitting a small vehicle increases the probability of survival.

The equation for P{V1} and P{V2} are specified as censored regression models to allow for a point mass at zero (18% and 27% of the models having no recorded fatalities for one-vehicle and two- vehicle accidents, respectively). This specification worked much better than a linear probability model. The data used corresponds to observations of make, model and year augmented by average driving characteristics from the FARS. Since the observed probabilities of having a fatal accident per 1000 vehicles are very small, it was unnecessary to impose an explicit upper limit of one on the dependent variable. The equations were estimated in SAS.

In order to be consistent with the unconditional probability of survival, each fatal accident is scaled by the inverse of the probability of observing the accident, i.e. at least one fatality occurred. The scaling is very easy for one-vehicle accidents. But for two-vehicle accidents, we need to know the characteristics, e.g. weight, of both vehicles. Among the 25126 two-vehicle accidents that occurred in 1995-1997 involving at least one of the vehicles we studied, there are 8282 accidents having complete information for both vehicles' characteristics. Thus, only one-third of the accidents have complete information about both vehicles' characteristics. There are two possible solutions: one is to find out the complete information of the other vehicle, the other is to scale the accidents with unknown characteristics of the other vehicle by the same scalar used to scale accidents with both vehicles' characteristics known. If the pattern of hitting the other vehicle is the same for each make/model/year vehicle whether the characteristics of the other vehicle is known or not, then the second way is a reasonable approximation.

A goodness-of-fit test is used to test whether the pattern of accidents is the same or not. The probability having a two-vehicle accident is calculated by each make/model/year, but due to the limited number of observations, accidents for each make/model/year were aggregated to 23 types of vehicle. The overall χ^2 test is rejected, but when we only consider the first 21 types, the χ^2 test cannot be rejected. The remaining two types are small and large pick-up trucks. After comparing the distribution of the other vehicles hit by the 21 types, and by small and large pick-ups, it is found pick-up trucks tend to hit more than their share of old vehicles, whose characteristics are not collected by this study. Since old and new vehicles are similar in weight, and the age of vehicle isn't a significant factor determining the probability of survival, all accidents by make/model/year are inflated by the scalar derived from the subset with complete vehicle characteristics.

The remaining part of the mortality rate estimation is to estimate the probability for multiple vehicle accidents P{Vm} and P_m. Unlike two-vehicle accidents, the pattern of collision is very hard to identify in multiple-vehicle accidents. Some of the vehicles may have no direct impact on each other. Therefore, the model for P_m is more like the model for a one-vehicle accident, i.e. no information of the other vehicles is included. In addition, we assume that all multiple-vehicle accidents are observed. Since the total number of vehicle occupants involved in a multiple-vehicle accident could be quite large (at least 3), this is a reasonable approximation. Also, the fatalities in multiple-vehicle accidents are only 8.5% of the total fatalities that occurred in 1995-1997. The equation for the survival rate P_m is specified as a regular logit model and estimated by maximum likelihood in SAS. The equation for P{Vm} is specified as censored regression model to allow for a point mass at zero (24% of the models) and estimated by SAS.

Explanatory variables in the censored models for P_1, P_2, P_m that are not listed in Table A4.1 are described in Table A4.5. The basic differences are that small subdivisions of the classes of vehicles are made, for example, to identify sports cars from non-sports cars for one-vehicle accidents. In addition, variables such as styling ((length plus width/height) are included to provide more information about the type of vehicle.

12 of the total of 1261 vehicle types were dropped, because they had sales less than 500 vehicles, before estimating the censored regression of P{V1}, P{V2}, P{Vm}. With a very small number of vehicles on the road, even one fatal accident for that make/model/year will count as a big probability. The increase in the number of subclasses of vehicle for P{V1} was prompted by inspection of the raw data. The effects of variables such as alcohol and previous convictions are partly responsible for the high rates of accidents for some types of vehicles. For P{V1}, P{V2} and P{Vm}, the accident rate increases for young drivers, for older drivers and, surprisingly, for female drivers. Accidents are more likely to occur at highway speeds (Sp), and for all three types of accidents, powerful vehicles (Acceleration) are more likely to have accidents, especially for one-vehicle accidents. The use of alcohol and previous convictions increases P{V1}, P{V2} and P{Vm}. The overall

conclusion is that driving behavior does matter and affects the probabilities of having a fatal accident for different types of vehicle.

Table A4.1: Variable Definition for the Estimation of Probability of Survival

Variable Name	Definition
Restraint	Coded as 1 if the passenger used restraint, 0 otherwise.
Age0_5	Coded as 1 if the passenger age is ≤5, 0 otherwise.
Age15	Coded as 1 if the passenger age is ≥6 but ≤15, 0 otherwise.
Age21	Coded as 1 if the passenger age is ≥16 but ≤21, 0 otherwise.
Age24	Coded as 1 if the passenger age is ≥22 but ≤24, 0 otherwise.
Age_o	Coded as 1 if the passenger age is ≥65, 0 otherwise.
female	Coded as 1 if the passenger is female, 0 otherwise.
Occupants Number	logarithm of number of occupants.
ClassX	Discrete variables coded as 1 for the appropriate class. Class1 to class7 represent small, middle, large, luxury, SUV, van, and pick-up truck, respectively, class40, class41 represents luxury non-sports and luxury sports, respectively.
Weight	Weight of the vehicle (1000lb).
Weight Ratio	Weight ratio of the vehicle to the other vehicle in a two-vehicle accident.
Acceleration	Horsepower to weight ratio.
Vehicle Age	The age of the vehicle when the accident happened.
O_classX	The class code for the other vehicle.
Female Driver	Code as 1 if the driver is female.
Driver 16	Code as 1 if the driver is ≤ 16.
Young Driver	Coded as 1 if the driver is ≥16 but ≤24, 0 otherwise.
Older Driver	Coded as 1 if the driver is 65 or older.
Alcohol	Coded as 1 if the alcohol involvement is reported
Late Night	Code as 1 if the accident occurred between 12:00am to 5:59am.
No Previous Offenses	Code as 1 if the driver had no previous offenses.
Sp_limit	Speed limit (10 miles).
Seatfp	Coded as 1 for front seat non-driver passenger.
Seatb	Coded as 1 for back seat passenger.
airbag	Coded as 1 for airbag in that seat position.

Table A4.2: The Probability of Survival in a One-Vehicle Accident

Parameters	Estimates	t ratio	Prob.
Constant	1.485	5.067	0
Restraint	1.0943	25.028	0
Age0_5	0.2011	2.407	0.0161
Age15	0.6061	9.552	0
Age21	0.4501	8.547	0
Age24	0.3464	5.701	0
Age_o	-1.0999	-10.49	0
female	-0.2026	-5.948	0
Occupants Number	0.3961	5.999	0
Weight	-0.0737	-1.176	0.2395
Acceleration	-8.913	-2.539	0.0111
Vehicle Age	0.0025	0.154	0.8777
Class2	-0.0018	-0.022	0.9821
Class3	-0.0028	-0.016	0.9875
Class40	-0.1111	-0.739	0.4602
Class41	0.2465	0.924	0.3555
Class5	0.4951	3.601	0.0003
Class6	0.2715	1.992	0.0463
Class7	0.62	5.024	0
Sp_limit	-0.0627	-2.792	0.0052
airbag	-0.0054	-0.109	0.9132
Seatfp	-0.0612	-1.854	0.0637
Seatb	0.1249	2.535	0.0112
Driver 16	-0.4153	-4.046	0.0001
Young Driver	-0.2345	-3.42	0.0006
Older Driver	0.5054	3.625	0.0003
Female Driver	0.2059	3.248	0.0012
Alcohol	-0.0607	-0.965	0.3347
No Previous Offenses	-0.1916	-3.417	0.0006
Late Night	-0.1014	-1.672	0.0945

Table A4.3: The Probability of Survival in a Two-vehicle Accident

Parameters	Estimates	t ratio	Prob.
Constant	2.0436	6.544	0
Restraint	0.9234	18.053	0
Age0_5	0.0402	0.297	0.7663
Age15	0.4342	3.81	0.0001
Age21	0.4615	4.758	0
Age24	0.4571	4.14	0
Age_o	-1.5055	-13.25	0
female	-0.1744	-3.467	0.0005
Occupants Number	0.2572	4.65	0
Weight	0.1566	1.431	0.1523
Weight ratio	1.3538	6.626	0
Vehicle Age	-0.0018	-1.02	0.3077
Class2	0.0684	0.735	0.4624
Class3	0.0407	0.254	0.7993
Class40	-0.28	-1.773	0.0762
Class41	-0.4624	-1.136	0.2561
Class5	0.4542	2.837	0.0046
Class6	0.4554	3.184	0.0015
Class7	0.6685	5.06	0
O_class2	-0.1202	-1.167	0.2431
O_class3	-0.0789	-0.47	0.6387
O_class4	-0.289	-1.792	0.0732
O_class41	-1.7245	-5.119	0
O_class5	-0.3271	-2.064	0.039
O_class6	-0.3155	-2.067	0.0387
O_class7	-0.4214	-3.12	0.0018
Sp_limit	-0.447	12.714	0
airbag	0.1316	2.395	0.0166
Seatfp	-0.114	-1.306	0.1915
Seatb	0.2585	2.418	0.0156
Driver 16	-0.613	-4	0.0001
Young Driver	-0.0165	-0.177	0.8597
Older Driver	0.172	1.445	0.1484
Female Driver	-0.0423	-0.696	0.4862
Alcohol	-0.6062	-7.997	0
No Previous Offenses	-0.0599	-1.15	0.2501
Late Night	-0.6222	-6.02	0
Acceleration	-2.2826	-0.594	0.5523

Table A4.4: Logit Model for the Survival Rate in a Multiple-vehicle Accident

Parameters	Estimates	Wald χ^2	Prob.
Constant	0.0075	0.001	0.9774
Restraint	0.9570	437.001	0.0001
Age0_5	-0.0337	0.066	0.7978
Age15	0.2592	5.312	0.0212
Age21	0.2673	6.274	0.0123
Age24	0.4072	10.560	0.0012
Age_o	-1.5429	177.689	0.0001
female	-0.1232	4.486	0.0342
Occupants Number	0.5399	100.273	0.0001
Weight	0.3223	35.925	0.0001
Acceleration	7.7231	4.251	0.0392
Vehicle Age	-0.0072	0.216	0.6424
Class2	0.2409	11.420	0.0007
Class3	0.4135	11.050	0.0009
Class40	0.2486	3.611	0.0574
Class41	-0.0749	0.042	0.8380
Class5	0.7535	33.524	0.0001
Class6	0.6679	36.716	0.0001
Class7	0.7384	52.197	0.0001
Sp_limit	-0.2297	116.802	0.0001
airbag	0.1788	8.380	0.0038
Seatfp	-0.0683	1.050	0.3055
Seatb	0.2892	8.380	0.0038
Driver 16	-0.3106	2.337	0.1264
Young Driver	-0.0581	0.354	0.5518
Older Driver	0.2642	4.784	0.0287
Female Driver	0.0268	0.199	0.6557
Alcohol	-0.8750	108.279	0.0001
No Previous Offenses	0.0627	1.844	0.1745
Late Night	0.0192	0.045	0.8316

Table A4.5: Variable Definition for the Censored Regression

Variable Name	Definition
TypeXX	Coded as 1 for the appropriate type. Type1 to Type23 represent lower, upper small, small specialty, lower, upper middle, middle specialty, large, large specialty, lower, middle, upper luxury, luxury specialty, luxury sport, small, middle, large, luxury suv, small, middle, large, luxury van, small, large pickup, respectively.
Alcohol	Proportion of accidents in this make/model/year vehicle in which the alcohol involvement was reported.
No Previous Offenses	Proportion of accidents in this make/model/year vehicle in which the driver had no previous offense.
Late Night	Proportion of accidents in this make/model/year vehicle which occurre between 12:00am to 5:59am.
Driver 16	Proportion of accidents in this make/model/year vehicle in which the driver is 16 or younger.
Young Driver	Proportion of accidents in this make/model/year vehicle in which the driver is younger than 25 years, but older than 16..
Older Driver	Proportion of accidents in this make/model/year vehicle in which the driver is 65 or older.
Female Driver	Proportion of accidents in this make/model/year vehicle in which the driver was female.
Sp	Proportion of accidents at highway speed.
Acceleration	The horsepower-to-weight ratio.
Traditional Styling	Length plus width divided by height.
D_airbag	Coded as 1 for the driver-side airbag.
P_airbag	Coded as 1 for the passenger-side airbag.

Table A4.6: Censored Regression for the Probability of Having a One-Vehicle Accident

parameter	Estimate	std. Error	ChiSquare
constant	-0.2231	0.097	5.25
Alcohol	0.0925	0.019	22.69
No Previous Offenses	-0.1255	0.017	51.53
Late Night	-0.0056	0.022	0.07
Driver 16	0.1232	0.048	6.69
Young Driver	0.1814	0.022	69.42
Older Driver	0.1038	0.028	13.44
Female Driver	0.1018	0.019	28.28
Sp	0.1247	0.017	53.40
Acceleration	3.9825	0.654	37.08
Traditional Styling	0.0280	0.025	1.24
Weight	-0.0059	0.018	0.11
D_airbag	-0.0360	0.012	8.74
P_airbag	-0.0159	0.013	1.42
Type2	-0.0702	0.023	9.11
Type3	-0.0604	0.029	4.21
Type4	-0.0965	0.027	12.94
Type5	-0.0805	0.028	8.17
Type6	-0.0632	0.031	4.23
Type7	-0.0969	0.038	6.44
Type8	-0.0996	0.054	3.40
Type9	-0.0985	0.034	8.45
Type10	-0.1095	0.033	10.70
Type11	-0.1558	0.041	14.47
Type12	-0.0797	0.043	3.40
Type13	0.0725	0.038	3.64
Type14	0.1634	0.044	13.98
Type15	0.1194	0.045	7.07
Type16	0.0569	0.060	0.91
Type17	0.1454	0.057	6.58
Type18	-0.0232	0.040	0.33
Type19	-0.0491	0.054	0.82
Type20	0.0110	0.062	0.03
Type21	-0.0366	0.065	0.31
Type22	0.1359	0.032	18.32
Type23	0.0932	0.052	3.23
Sigma	0.1526	0.003	

Table A4.7: Censored Regression for the Probability of Having a Two Vehicle Accident

parameter	Estimate	std. Error	ChiSquare
constant	0.0197	0.061	0.11
Alcohol	0.0063	0.023	0.08
No Previous Offenses	-0.1561	0.015	109.11
Late Night	0.0448	0.024	3.54
Driver 16	0.1449	0.051	7.94
Young Driver	0.0980	0.018	30.26
Older Driver	0.1098	0.019	32.47
Female Driver	0.1102	0.014	61.16
Sp	0.1067	0.014	58.72
Acceleration	0.0858	0.428	0.04
Traditional Styling	0.0074	0.015	0.23
Weight	0.0261	0.010	6.22
D_airbag	-0.0119	0.007	2.88
P_airbag	-0.0062	0.008	0.61
Type2	-0.0258	0.013	3.84
Type3	-0.0377	0.017	4.75
Type4	-0.0630	0.015	16.87
Type5	-0.0476	0.016	8.36
Type6	-0.0549	0.019	8.73
Type7	-0.0577	0.022	6.68
Type8	-0.0887	0.031	8.28
Type9	-0.1069	0.020	28.49
Type10	-0.0972	0.021	22.06
Type11	-0.1146	0.026	19.86
Type12	-0.0909	0.027	11.46
Type13	-0.0939	0.026	13.39
Type14	-0.0438	0.026	2.75
Type15	-0.0274	0.026	1.09
Type16	-0.0196	0.035	0.31
Type17	-0.0679	0.034	3.98
Type18	-0.0469	0.023	4.03
Type19	-0.0627	0.031	3.98
Type20	-0.0193	0.036	0.29
Type21	-0.0590	0.036	2.64
Type22	0.0338	0.018	3.35
Type23	0.0218	0.030	0.52
Sigma	0.0851	0.002	

Table A4.8: Censored Regression for the Probability of Having a Multiple (three or more) Vehicle Accident

parameter	Estimate	std. Error	ChiSquare
constant	-0.0067	0.013	0.25
Alcohol	0.0051	0.005	1.10
No Previous Offenses	-0.0192	0.002	68.94
Late Night	0.0118	0.005	6.58
Driver 16	0.0291	0.008	11.76
Young Driver	0.0130	0.003	17.97
Older Driver	0.0180	0.004	26.31
Female Driver	0.0213	0.002	93.38
Sp	0.0214	0.002	103.77
Acceleration	0.1539	0.092	2.79
Traditional Styling	0.0012	0.003	0.13
Weight	0.0038	0.002	2.58
D_airbag	0.0006	0.002	0.12
P_airbag	0.0004	0.002	0.06
Type2	-0.0100	0.003	10.86
Type3	-0.0133	0.004	11.30
Type4	-0.0128	0.004	13.05
Type5	-0.0078	0.004	4.30
Type6	-0.0124	0.004	9.00
Type7	-0.0099	0.005	3.80
Type8	-0.0139	0.007	3.77
Type9	-0.0164	0.005	12.73
Type10	-0.0163	0.004	13.13
Type11	-0.0203	0.006	12.85
Type12	-0.0020	0.006	0.12
Type13	-0.0173	0.005	10.03
Type14	-0.0005	0.006	0.01
Type15	-0.0006	0.006	0.01
Type16	-0.0035	0.008	0.20
Type17	-0.0136	0.008	3.13
Type18	-0.0086	0.005	2.65
Type19	-0.0049	0.007	0.49
Type20	-0.0036	0.008	0.20
Type21	-0.0031	0.008	0.16
Type22	0.0030	0.004	0.50
Type23	-0.0018	0.007	0.07
Sigma	0.0200	0.000	

Appendix B: Additional Issues on the Censored Model for the Probability of having Vehicle Accidents

The probability of having vehicle accidents is estimated by a censored regression model to allow for a probability mass at zero. However, the consistency and asymptotic normal distribution of the estimators are quite sensitive to the error distribution. There are two problems: First, likelihood-based estimators are inconsistent when the assumed parametric form of the likelihood function is incorrect. Second, when the heteroskedasticity of the error terms occurs, the parameter estimates are also inconsistent (Powell).

When estimating the probability of having an accident, several versions of the censored model were tested to deal with non-normality and heteroskedasticity of the error terms. Those methods can be summarized into three categories: (1) multiplicative heteroskedasticity of the error terms, (2) the error terms contain two parts: a normally distributed residual and a measurement error, and (3) as an alternative to maximum likelihood estimation, the symmetrically trimmed (censored) least square estimation (STLS) method proposed by Powell is used.

Method 1. Multiplicative Heteroskedasticity

The heteroskedastic censored model is defined as:

$p_i = \beta' x_i + u_i$ if RHS>0, (B1)

$p_i = 0$ otherwise, (B2)

where $u_i \sim N(0, \sigma_i^2)$, and $\sigma_i^2 = e^{\alpha' w_i}$.

A test of heteroskedasticity is to test $\alpha' = 0$, except for the intercept. Vector β' and α' contain coefficients corresponding to explanatory variables. The explanatory variables x and w may or may not be identical. Heteroskedastic censored models for one, two and multiple vehicle accidents with identical x and w are estimated, respectively. A likelihood-ratio test of heteroskedasticity is conducted and the homoskedastic censored model is rejected, indicating the existence of heteroskedasticity.

Method 2. Error Terms Containing Measurement Error

In the graph of the residual vs. the number of vehicles sold, the error terms have an inverse relation to the number of vehicles sold. One can view the chance of a vehicle having an accident as a Bernoulli trial. Given the number of vehicles sold n_i (assumed equivalent to the number of vehicles on the road), the difference of the observed probability of having a vehicle accident and the true probability is approximately normal distributed with zero mean and $\frac{p_i(1-p_i)}{n_i}$ variance. Further more, since the probability is very small, the variance can be approximated as p_i / n_i. The censored model in equations B1 and B2 can be specified with the error term written as follows:

$u_i = e_i + v_i$,

where $e_i \sim N(0, \sigma^2)$, and $v_i \sim N(0, \frac{p_i}{n_i})$,

v_i is the measurement error, p_i is the probability of having an accident, and n_i is the number of vehicles sold.

Assume cov[e_i, v_i] = 0, then $u_i \sim N(0, \sigma^2 + \frac{p_i}{n_i})$, and the log-likelihood function can be written as:

$$\log L = \sum_0 \log[1 - \Phi(\frac{\beta' x_i}{\sigma_i})] + \sum_1 \log[\frac{1}{\sigma_i}\phi(\frac{p_i - \beta' x_i}{\sigma_i})\,],$$

where σ_i is specified as

$$\sigma_i^{\,2} = \sigma^2 + \frac{\beta' x_i}{n_i}, \qquad \text{if } \beta' \mathrm{x_i} > 0,$$
$$\sigma_i^{\,2} = \sigma^2, \qquad \text{if } \beta' \mathrm{x_i} \le 0,$$

which is an approximation of $\sigma_i^{\,2} = \sigma^2 + \frac{p_i}{n_i}$.

Again the value of the log-likelihood function is significantly larger than that of the homoskedastic censored model even though the degrees of freedom are the same for the two models. Both method 1 and 2 are estimated using GAUSS.

Method 3. Symmetrically Trimmed Least Squares Estimation for Censored Model (STLS)

The STLS method is based on censoring of the upper tail of the distribution of the dependent variable so that symmetry is restored. The resulting semiparametric estimator is shown to be consistent and asymptotically normally distributed, given the assumption of symmetrically and independently distributed error terms.

First, the "true" underlying regression equation is

$p_i^* = \beta' x_i + u_i$.

In the censored regression model, only the values of x_i and $p_i = \max\{0, p_i^*\}$ are observed, which induces asymmetry in the distribution of the error terms. The error terms are of the form $e_i = \max\{u_i, -\beta' x_i\}$, the "symmetric censoring" would replace e_i with $\min\{u_i, \beta' x_i\}$ whenever $\beta' x_i > 0$, and the observations with $\beta' x_i < 0$ are deleted. Therefore, under STLS method, the upper tail of the distribution of the dependent variable is replaced with min $\{p_i, 2\beta' x_i\}$.

The symmetrically censored least squares estimator $\hat{\beta}_I$ is obtained by minimizing $S_I(\beta)$, defined as follows:

$$S_I(\beta) = \sum_{i=1}^{I} (p_i - \max\{\frac{1}{2} p_i, \beta' x_i\})^2 + \sum_{i=1}^{I} \mathrm{I}(p_i > 2\beta' x_i) \cdot [(\frac{1}{2} p_i)^2 - (\max\{0, \beta' x_i\})^2],$$

where $I(p_i > 2\beta' x_i)$ is an indicator function, which takes the value of 1 if true, and 0 if false. Minimizing $S_I(\beta)$ is conducted by the GAUSS optimization module. The asymptotic variance of $\hat{\beta}_I$ is

$$\frac{1}{I} C_I^{-1} D_I C_I^{-1}$$

where

$$\hat{C}_I = \frac{1}{I} \sum_{i=1}^{I} I(-\hat{\beta}_I' x_i < \hat{u}_i < \hat{\beta}_I' x_i) \cdot x_i x'_i, \qquad \text{and}$$

$$\hat{D}_I = \frac{1}{I} \sum_{i=1}^{I} I(\hat{\beta}_I' x_i > 0) \cdot \min\{\hat{u}_i^2, (\hat{\beta}_I' x_i)^2\} \cdot x_i x'_i$$

where $\hat{u}_i \equiv p_i - \hat{\beta}_I' x_i$, I is the total number of observations.

All three methods give the same signs for the estimated coefficients as the homoskedastic censored regression model, but the sizes of the coefficients vary from model to model. For method 1 and 2, the tests for the existence of heteroskedasticity, using a likelihood ratio test, reject homoskedasticity. Since method 3 does not use maximum likelihood, no direct test can be done for heteroskedasticity. The advantage of method 3 is that it automatically deals with the unknown form of heteroskedasticity using semiparametric estimation.

Although the three methods of dealing with heteroskedasticity and non-normality seem to be legitimate alternative specifications to the standard censored regression model, the lack of robustness of the estimated coefficients presents a practical problem of deciding which alternative is best. In general, different approaches to the same type of statistical problem should give similar results. Since this was not the case, the standard model was used to estimate the probabilities of having accidents of different types for the hedonic models. The issue of alternative specifications will be left for further research. It is quite possible that heteroskedasticity and non-normality are not the most important misspecifications. An implicit assumption from using the FARS data is that the average driver's characteristics for each type of vehicle derived from fatal accidents is representative of all drivers. This is a big assumption, and it would be interesting to look at drivers' characteristics in non-fatal accidents as well as fatal accidents. With data on all types of accidents, it would be possible to specify a multinomial model to estimate the probabilities of different types of accidents in the same model. Issues of heteroskedasticity and non-normality could still be considered, but they are not likely to be the most important way to improve the standard model in this application.

Appendix C: Scatter Plots of Injury Rate vs. Mortality Rate

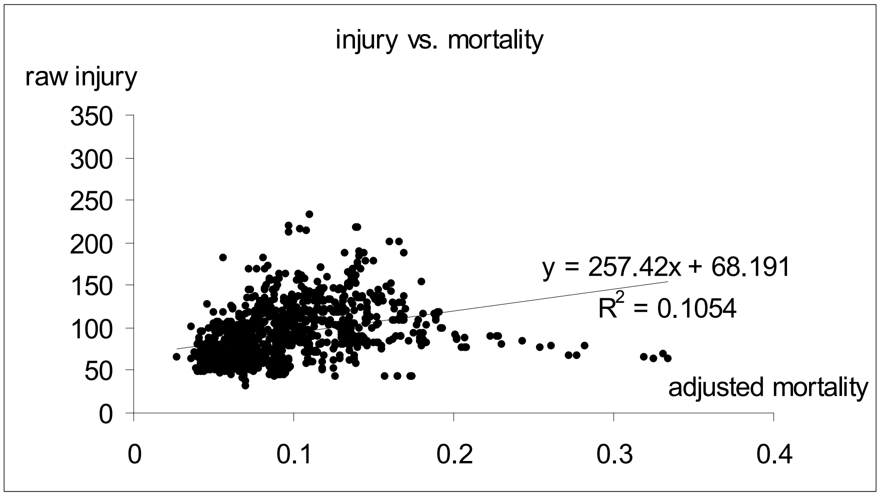


Figure A4.1: Scatter Plot of Raw Injury vs. Adjusted Mortality Rate

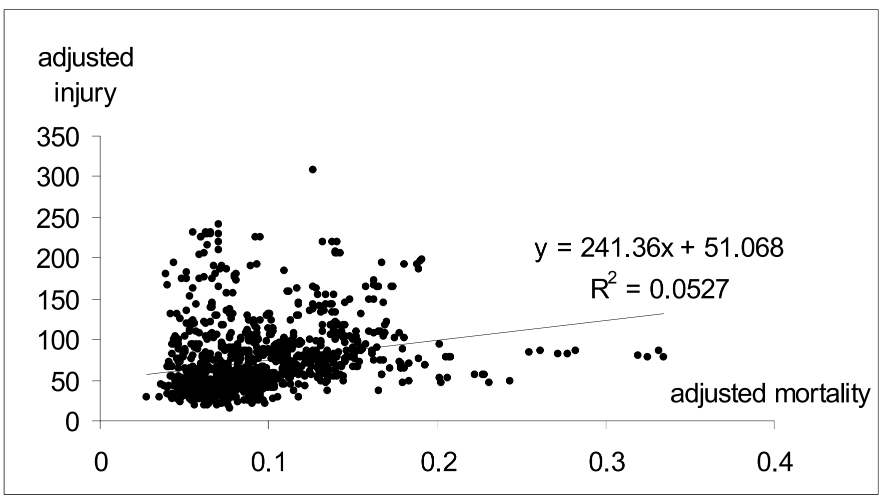


Figure A4.2: Scatter Plot of Adjusted Injury vs. Adjusted Mortality Rate

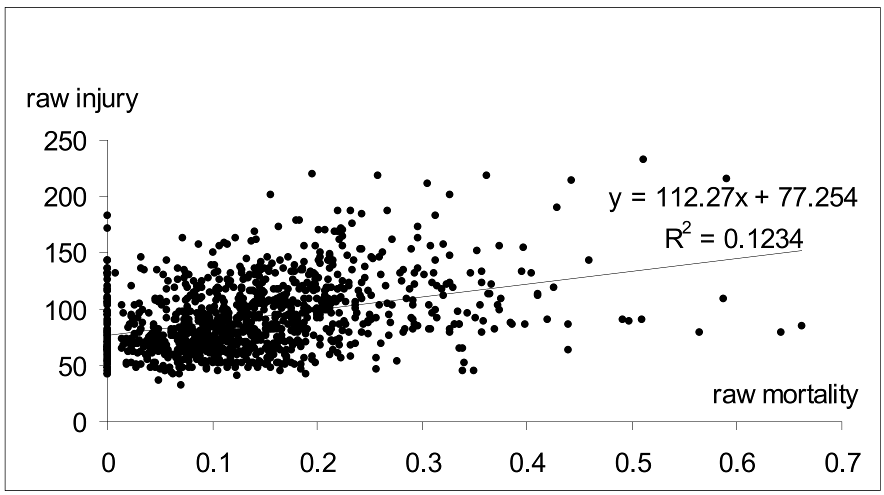


Figure A4.3: Scatter Plot of Raw Injury vs. Raw Mortality Rate