Jean L. Freeman, William W. Hunter, Gary G. Koch, and John H. Lacey. (1975). Shoulder Harness Usage in the Population of Drivers at Risk in North Carolina. Chapel Hill, NC: University of North Carolina Highway Safety Research Center.

ATTENTION

The enclosed report is a reprint of the original technical report which has recently gone out of print. Its content does not differ in any way from the original report. The format differs slightly due to time restrictions in the reprinting process.

We hope that this report will fulfill your interests. We appreciate your continued concern in highway safety.

Shoulder Harness Usage in the Population of Drivers at Risk in North Carolina

Jean L. Freeman
William W. Hunter
Gary G. Koch
John H. Lacey

University of North Carolina
Highway Safety Research Center
Chapel Hill, North Carolina

April 1975

The UNC Highway Safety Research Center was created by an act of the 1965 North Carolina General Assembly. A three-point mandate issued by the Governor authorized HSRC to 1) evaluate the state's highway safety programs, 2) conduct research, and 3) instruct and train other working professionals in highway safety.

This project was supported by
the Governor's Highway Safety Program
and a grant from the Institute of
General Medical Sciences.

ABSTRACT

Shoulder harness usage rates based on observations of vehicles in North Carolina in the summer of 1971 are reported. Tabulations of usage rates based on 19,338 observations are presented on a variety of demographic and environmental factors. A model of usage patterns is developed using a method of analysis of categorical data by linear models. The clusters derived with the highest predicted usage rates (33.3%) were foreign car drivers from out-of-state on Interstates, mature and older white male out-of-state drivers of foreign cars on rural non-interstate roads and young white male out-of-state foreign car drivers on rural four-lane divided highways. Values of variables which most consistently positively influenced usage rates were rural, foreign car, road size, and males. The overall usage rate is 4.8 percent with the males' rate being 5.4 percent and females' rate 3.2 percent. The influence on usage rate of 3-point versus 4-point shoulder harness systems in U.S. cars is examined with no real difference demonstrated. The effect of drivers' usage patterns on passenger versus behavior is discussed and a usage rate of 51.1 percent for passengers whose drivers were wearing shoulder harnesses is reported.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

Appreciation is expressed to Dr. Donald W. Reinfurt of the UNC Highway Safety Research Center for statistical consultation, to Ms. Diane S. Pozefsky for some special computer programming required in the analysis, and to the many HSRC staff members who assisted in the data collection.

I. REVIEW OF THE LITERATURE

There are numerous studies in the literature which show the general efficacy of lap and shoulder belts in injury reduction. Although governmental agencies, auto manufacturers, and research organizations have published such findings widely, the use of vehicle occupant restraints remains distressingly low. The findings of the current study reflect the same trend.

This report deals exclusively with shoulder harness usage. This restraint was designed to prevent occupants from striking interior surfaces of vehicles, and has been available on U.S. vehicles since 1968. Since the device is relatively new, and since usage rates have remained very low, few studies have been made which compare shoulder harness users to non-users, either in terms of injury effects or usage patterns. Anderson (1971) found an occupant usage rate of eight percent in vehicles equipped with shoulder harnesses. In 1972, the U.S. Department of Transportation estimated that four percent of drivers in cars equipped with shoulder belts used them. A study conducted by the Insurance Institute for Highway Safety reported a combined lap and shoulder belt usage of eight percent in 1973 model vehicles equipped with the buzzer-light system, as compared to a combined lap and shoulder belt usage of 44 percent in 1974 model vehicles equipped with the starter-interlock system.

Additional studies comparing injuries of shoulder harness users and non-users have been done in Europe, where the device has been more readily accepted, and in Australia, where complusory belt usage laws have been enacted in some areas. Most of the restraints used in these countries are the 3-point type which necessitates the use of a shoulder belt whenever the lap belt is worn. This literature review will include both domestic and foreign studies. However, one must exercise care when comparing U.S. and foreign studies, since many differences exist in the data. Foreign usage rates are higher than rates in the U.S.; furthermore there are great differences in vehicle design and roadway design.

Most foreign studies have been concerned with injury reduction or injury benefits. Injury benefit studies depend on the number of drivers wearing restraints. Since foreign usage rates are much higher than U.S. rates, any prediction for U.S. accident savings based on results of studies done in other countries could be misleading.

Bohlin (1967) analyzed 28,000 accident cases occurring over a 12-month period in 1965-1966 in Sweden involving Volvo vehicles, almost all of which were equipped with a 3-point lap and shoulder belt system. Injuries were classified as fatal, non-fatal, or none for 28,780 drivers and 8,731 front seat passengers. The 3-point belts were used by approximately 24 percent of the drivers and 30 percent of the front seat passengers. Usage was found to be directly related to estimated speed (probably due to higher belt usage on longer trips).

When occupants were questioned about how often they use belts, the responses were as follows: "Always" -- 19 percent; "Long trips only" -- 60 percent; "Seldom or never" -- 18 percent; "No answer" -- 3 percent. During the study, three times as many 20-year-old drivers were involved in crashes as 45-year-old drivers, although exposure to risk was not considered. Belt usage tended to increase with age, with 16 percent of the 20-year-old drivers wearing the restraint system, as compared with 30 percent of the 45-year-old drivers. Bohlin concluded that the use of the belt system tended to reduce or prevent injuries of all kinds and substantially reduced the frequency and severity of head and upper torso injuries.

Bohlin (1973) conducted a follow-up of his earlier Swedish study over a 12-month period completed in 1972. A total of 1,505 accidents involving Volvo vehicles with greater than $400 worth of damage were investigated. The study included 1,505 drivers, 503 front seat passengers, and 432 rear seat occupants. Injuries were classified according to the Abbreviated Injury Scale and vehicle deformation was codified by the Vehicle Damage Index. All of the vehicles had the 3-point restraint system for outboard front seat occupants, and 60 percent of the outboard rear seats had the 3-point system.

The overall usage rate for drivers and front seat passengers was 39.2 percent. Usage in urban crashes was 33.5 percent, while in rural crashes the usage rate was 43 percent. Only 7.8 percent of the rear seat occupants who had the restraint system available used the belts. Again, usage was found to increase with age; 30-35 percent of the 20-year-old occupants were wearing belts compared to 50-55 percent of the 50 to 55-year-old occupants. The users preferred a retractor belt over a non-retractor. Usage rates were approximately 30 percent higher for front seat occupants in cars with retractor belts than in cars not so equipped. Injury severity was again shown to be lower for belted occupants.

Vulcan (1973) attempted to examine the effect of a compulsory seat belt use law in Victoria, Australia. In 1967, the use of a 3-point system was 10 percent for accident-involved drivers in Victoria. Seat belts for outboard front seating positions became compulsory equipment on January 1, 1970, and for all seating positions after January 1, 1971, with few exceptions. The outboard positions were required to have the combined lap and shoulder 3-point system, and middle seats were required to have lap belts available. On December 22, 1970. compulsory use of installed belt systems by all occupants was mandated. A one-month grace period was granted to become familiar with the law. After January 23, 1971, it was a punishable offense for failure to wear seat belts when available, with penalty set at $20. A survey indicated that public attitude was positive, with over 75 percent approving of the law.

By 1971, 76 percent of all front seat positions were equipped with belts, with 90 percent of the belts being the 3-point type. A survey made in the latter part of 1971 to evaluate the effect of the legislation indicated that usage rates increased from approximately 25 percent before the legislation to 64-75 percent after the legislation for vehicles fitted with the belts. However, only 13.5 percent were found to be wearing the 3-point belt system correctly (i.e., the majority of those using the belt system either had it too loosely adjusted or improperly located across the body). It was apparent that usage rates were substantially increased by the legislation.

The Road Traffic Board of South Australia (1972) reviewed seat belt usage in South Australia from 1964 through 1971. A law passed in South Australia in 1963 required anchorages and belts to be provided for drivers and front seat passengers. Station wagons, panel vans, and service vehicles were initially exempted. Motor vehicles were officially required to be fitted by January 1, 1967.

Belt use by all occupants, including those in vehicles not equipped with belts, increased from 8.2 percent in 1964 to 22.6 percent in 1971 with most of the increase coming in the last year when compulsory belt use was imminent. Over the seven year period the percent of vehicles equipped with belts increased from 15.1 to 66.9 percent. Comparison of usage rates with availability, however, indicates that even though the availability of belts increased dramatically, many chose not to use them.

Usage by all occupants with belts available was 60 percent in 1964. This rate decreased and leveled off to 26 percent in 1968, 1969, and

1970, and increased to 36 percent in 1971. Trends were the same for drivers and other passengers.

It is possible that the novelty appeal of the restraint system declined after its introduction. More likely, persons who had belts installed at their own initiative before the mandatory installation were more inclined to use their belts. The imminence of the compulsory belt use law apparently resulted in an upswing of actual belt use. South Australia passed a compulsory usage law on November 29, 1972, after reviewing the State of Victoria's experience.

Andreassend (1972) originated a continuing comprehensive study of the effects of the mandatory usage legislation in Victoria during 1971. Vehicles on the road were observed to determine the usage rates of drivers and front seat passengers during times of peak accident occurrence. A sample of motorists was interviewed to obtain not only demographic information but also travel and attitude data related to usage.

Andreassend sampled in the Melbourne area to determine what percent of vehicles were fitted with 3-point belts, how many drivers were wearing the fitted belts, and the percentage of all drivers wearing belts. Interview teams also checked to see whether the belts were being worn correctly.

A total of 63,587 cars were observed in metropolitan Melbourne and the overall driver usage rate was 75 percent. Overall driver usage was 64 percent in the rural area, based on 30,065 observed cars. Driver usage in both areas was approximately eight percent higher than passenger usage. The primary results are shown below.

	Percent of Cars Equipped With Belt Systems (1)	Usage Rate Where Available (Drivers) (2)	Overall Usage Rate in Total Observed Population (1) x (2)
Urban	73%	75%	55%
Rural	65%	64%	42%

Although women tended to have a higher wearing rate than men, their vehicles were less likely to be fitted with belts; therefore, the overall usage among women was lower than men. An interaction between driver and passenger seemed to be in effect. When a passenger was present, it was more likely to find both driver and passenger wearing belts or neither wearing belts than to find one wearing belts and not the other.

Overall use rates were similar for the interview results (76 percent for drivers, 77 percent for front passengers). Both women drivers and passengers tended to have a higher rate than their male counterparts. No significant relationship was found between age and usage (although usage was lowest for those under 20 years of age), perhaps because of small sample sizes. As far as trip length was concerned, the highest usage rate (82 percent) occurred on trips of 11-15 miles, while the lowest rate (57 percent) occurred on trips of 0-5 miles. Only 14 percent were wearing their belts correctly at the time of the interview. Thirty-four percent of the drivers stated that they wore their belts because of the safety provided, while 22 percent said they did so because of the compulsory legislation. Older age groups (40-49, 50-59, and greater than 60) tended to disapprove of the legislation.

Anderson (1971) attempted to estimate the percentage of drivers using an available shoulder belt according to a number of variables. Field observation of 1,707 drivers moving in traffic were collected across North Carolina, and an overall utilization rate of 8.26 percent was determined. Other results included the following:

1. Male drivers (9.51 percent) used the shoulder harness more than female drivers (4.82 percent).

2. Drivers of small foreign vehicles (19.86 percent) used the shoulder belt more than drivers of U.S. manufactured vehicles (5.96 percent).

3. Young drivers (11.15 percent) were observed to be using the shoulder harness more than either mature (7.16 percent) or older drivers (5.32 percent).

4. Drivers of out-of-state vehicles (12.19 percent) had a higher utliization rate than their in-state counterparts (7.19 percent), perhaps as a result of trip length.

Most of the studies in the literature, especially the European studies, point out the injury reducing potential of the lap and shoulder harness restraints. Occupant usage in these foreign countries is generally much higher than in the U.S. It is clear that in order to realize the tremendous potential of seat belts and shoulder harnesses for injury reduction, new ways must be found to ensure their use. Innovative approaches are particularly needed now in light of recent legislative measures rescinding the seat belt interlock requirement that was introduced in the 1974 model year autos. The observations for the current study were made before the interlock devices were introduced and thus may be meaningful in terms of future use patterns when cars are no longer equipped with the seat belt interlock systems.

II. METHODOLOGY

In February, 1971, Anderson reported on shoulder belt usage among drivers of cars observed in June, 1970, in a multi-county area of North Carolina. His analysis revealed an overall usage rate of 8.26 percent. Earlier North Carolina seat belt use studies by Campbell (1967), Campbell, et al. (1968) and Council (1969) revealed that patterns of use change with time. It was decided that a replication of Anderson's study observing patterns of use in North Carolina would be useful.

Shoulder harnesses became mandatory equipment in vehicles manufactured in the U.S. in 1968; vehicles included in Anderson's study were manufactured after January 1, 1969. Only 18 months had elapsed from that date to the time of Anderson's study; therefore, most of the vehicles which were equipped with shoulder harnesses were relatively new cars. A follow-up study would contain vehicles a year older, and consequently we could expect a different mix of drivers with regard to age and socio-economic status. Other studies have indicated that seat belt usage depends on both of these factors.

We wanted to collect data on many of the variables used by Anderson so that comparisons could be made. Moreover, we wanted to obtain as representative a sample of the summer driving population as possible. The summertime was chosen for logistic reasons -- availability of personnel and extended daylight hours.

It was decided to stratify the sample according to the following five factors: 1) month; 2) day of week; 3) hour of day; 4) type of road; 5) geographic area.

June, July, and August of 1971 were chosen as the three months for data collection. Days of week were grouped to form the following periods: 6 p.m. Friday to 2 a.m. Monday (weekend); 2 a.m. Monday to 10 a.m. Wednesday (beginning part of the workweek); and 10 a.m. Wednesday to 6 p.m. Friday (latter part of workweek). Hour of day was also grouped to form three periods: 7 a.m. - 1 p.m. (morning traffic), 1 - 6 p.m. (afternoon traffic); and 7 p.m. - 12 (evening traffic).

No shoulder harness data were collected during the 7 p.m. to midnight period, because we found that daylight was necessary to observe accurately shoulder harness usage--particularly in rural areas. The feasibility of using an infrared light source as in a sniper scope was tested. However, changes in levels of the ambient light conditions created by vehicle headlights, coupled with the narrowed field of vision of the scope, combined to render this method impractical.

Flexibility was built into the data collection process by using the broader time categories of section of week and part of day rather than identifying a particular hour of a specific day of the week. This allowed increased travel time between sampling sites if needed, and thus broader coverage of geographic areas. Within both the 7 a.m. - 1 p.m. and 1 p.m. - 6 p.m. time intervals, two one-and-one-half hour sampling periods were identified. These were 8 - 9:30 a.m. (work traffic) and 11:30 a.m. - 1 p.m., and 1 - 2:30 p.m. and 4 - 5:30 p.m. (work traffic), respectively.

The three geographical areas delineated were the eastern, the piedmont, and the western parts of the state. These were in turn subdivided by Highway Patrol troop and district for assignment. The wide choice of districts within each region served to simplify scheduling.

The actual type of sampling location within month, time period and geographical region was determined by road type. Road type was identified in terms of the following: 1) urban or rural; 2) number of lanes; and 3) presence of median. Thus, urban and rural could be applied across 2-lane, 4-lane, 4-lane divided, and Interstate roadways. If within the area assigned for sampling the particular road type described was unavailable (in some areas there were no Interstate roads), the observers were told to select a site whose road type was nearest to the prescribed road type yet still in the area specified.

Within the categories described above, observation teams were assigned to provide a balanced representation of each of the sampling variables. Thus, a typical assignment might be to observe an urban, two-lane roadway in Troop D, District 3, between 8 and 9:30 on Tuesday in June.

To be included in the sample a vehicle almost certainly had to be equipped with shoulder harnesses. All automobiles manufactured for sale in the United States were mandatorily equipped with shoulder harnesses for both outboard front seat passengers after January 1, 1968, and with factory installed head restraints after January 1, 1969. Therefore, almost all cars equipped with head restraints also have shoulder harnesses; cars that were optionally equipped with head restraints and not shoulder harnesses prior to January 1, 1968, are the most likely exceptions. Observers were told to make observations only on cars equipped with factory-installed head restraints. Considering the previous availability of head restraints as optional equipment, Anderson estimated that a maximum of one percent of all vehicles having head restraints would not be equipped with shoulder harnesses at the time of his study. We assumed that the mix of cars one year later would include even fewer of these exceptions.

Two-man observation teams (one observer and one recorder) were instructed to screen vehicles passing their position for eligibility for sampling (i.e., presence of factory-installed head restraints). Once it was determined that a vehicle was eligible, the following information was collected:

1. Shoulder harness use of driver.

2. Approximate age, race and sex of driver.

3. Presence or absence of right front passenger.

4. Shoulder harness use of right front passenger

5. License number, if North Carolina licensed vehicle.

6. Make of passenger car (General Motors, Ford, Foreign, Other U.S., unknown).

Observers were instructed to record data on as many vehicles as could be accurately observed within the sampling period. If the vehicle was a driver training vehicle they were instructed to also record that information. Provision was made for the observers to record a value representing "unknown" on any of the variables taken for each vehicle, and they were instructed to code "unknown" for any item they were not sure of. The observers were also instructed to try to ascertain or record the variables in the order they appeared on the form from left to right (see Appendix F). Thus, a hierarchy of importance was established among the variables with age, race, sex and shoulder harness use of driver being the most important. The result is that there are not equal numbers of valid observations for each variable. However, the values given may be more accurate than otherwise.

Observers sampled for a full one-and-one-half hour period. On all urban roads except Interstates and on two-lane rural roads, they stationed themselves beside the road to collect data. On all Interstates and four-lane rural roads they travelled in passenger vehicles at a speed approximately 10 miles per hour less than the prevailing traffic speed and collected data on occupants of vehicles passing them travelling in the same direction. A total of 19,338 observations were made.

The forms used for data collection were designed to allow recording of up to 20 separate vehicles. This form is included as Appendix F.

III. RESULTS OF ANALYSIS

In the methodology section, the sample design and data collection activities were described. For every vehicle that was considered eligible for inclusion in the sample, information on the variables shown in Table 1 was gathered. Preliminary screening tests using Pearson chi-square tests of association between each variable and driver usage were performed (see Appendix A). The following seven variables were found to be significantly associated with use of the shoulder belt: race, sex, age, vehicle registration, vehicle make, road type, and location. The cell frequencies (i.e., number of vehicles in each factor combination of race, sex, age, license, location, road type, car make) are presented in Table 2. The observed number of those drivers in each factor combination wearing a shoulder harness is given in Table 3. Finally, the usage rate or proportion of drivers wearing a shoulder harness for each cell is given in Table 4. The usage rate or proportion is determined by taking the ratio, on a cell by cell basis, of the numbers of drivers observed wearing harnesses, as noted in Table 3, over the total number of vehicles as noted in Table 2.

We proceeded to analyze this table of proportions using a method developed by Grizzle, Starmer, and Koch (1969) which is based on fitting linear regression models to functions of the cell proportions by the method of weighted least squares. A detailed mathematical description of the methodology can be found in Appendix B, while a discussion of how it was applied in this analysis appears in Appendix C. In this section we will present only the results of the analysis.

A linear regression model was fitted to the proportions which divided the drivers into nine homogeneous clusters with corresponding predicted usage rates of 8 percent, 2.5 percent, 4.1 percent, 5.4 percent, 7.1 percent, 9.7 percent, 15.8 percent, 27.2 percent and 33.3 percent (see Table 6). In the highest prediction rate (33.3

Table 1. A summary of the variables which were tested for association with the use of shoulder harnesses.

Driver Variables		Vehicle Variables		Environment Variables	
Variable	Values	Variable	Values	Variable	Values
Age	Young Mature (MA) Older (O)	Registration	N.C. Other (O)	Location	Rural (R) Urban (U)
Race	Blacks,Indian and Others (B)[1] White (W)	Vehicle make	U.S. Foreign (F)	Road type	Two-lane (2L) Four-lane (4NL) (Undivided) Four lane (4DL) (Divided) Four lane (4IL) (Interstate)
Sex	Male (M) Female (F)				
Driver Harness Use	Yes No			Pavement Condition	Wet Dry
Passenger	Present Absent			Month	March June July August September
Passenger Harness Use	Yes No			Day of Week	Sunday to Saturday
				Time of Day	One Hour Intervals From ____ to ____
				Speed	Posted Speed
				Observer's Position	Stationary En Route

[1] *In the remainder of the text, this group will be referred to as Blacks.*

Table 2. Observed number of vehicles.

Vehicle Type	Vehicle Registration	Race-Sex	Age	Location Site X Road Type Combination: Urban 2L	Urban 4NL	Urban 4DL	Urban 4IL	Rural 2L	Rural 4NL	Rural 4DL	Rural 4IL	Total
U.S.	N.C.	W-M	Y	293	319	280	15	412	233	692	234	2478
U.S.	N.C.	W-M	MA	527	686	370	38	602	499	1090	453	4265
U.S.	N.C.	W-M	O	138	109	81	4	142	113	220	99	906
U.S.	N.C.	W-F	Y	223	208	156	8	228	132	331	108	1394
U.S.	N.C.	W-F	MA	261	279	180	3	288	175	382	104	1672
U.S.	N.C.	W-F	O	59	36	29	2	43	34	75	25	303
U.S.	N.C.	B-M	Y	100	104	31	3	59	38	86	34	455
U.S.	N.C.	B-M	MA	57	52	22	0	21	19	50	17	238
U.S.	N.C.	B-M	O	7	2	1	0	3	0	4	0	17
U.S.	N.C.	B-F	Y	57	42	19	0	15	12	31	8	184
U.S.	N.C.	B-F	MA	33	25	7	0	14	7	31	7	124
U.S.	N.C.	B-F	O	2	2	0	0	0	3	1	0	8
U.S.	Other	W-M	Y	26	115	36	2	72	57	190	126	624
U.S.	Other	W-M	MA	65	167	70	15	168	122	390	229	1226
U.S.	Other	W-M	O	13	53	29	0	44	32	139	106	416
U.S.	Other	W-F	Y	17	21	11	3	22	23	58	39	194
U.S.	Other	W-F	MA	18	30	24	3	35	18	105	59	292
U.S.	Other	W-F	O	4	8	6	0	10	5	18	16	67
U.S.	Other	B-M	Y	5	8	1	1	3	4	45	32	99
U.S.	Other	B-M	MA	7	4	4	0	5	2	24	22	68
U.S.	Other	B-M	O	0	0	0	0	1	2	1	2	6
U.S.	Other	B-F	Y	4	5	1	0	1	2	5	8	26
U.S.	Other	B-F	MA	1	0	0	0	1	1	11	3	17
U.S.	Other	B-F	O	0	0	0	0	1	0	0	1	2
Foreign	N.C.	W-M	Y	80	103	70	1	101	74	134	55	618
Foreign	N.C.	W-M	MA	53	76	39	3	48	41	94	27	381
Foreign	N.C.	W-M	O	14	6	12	0	6	7	7	5	57
Foreign	N.C.	W-F	Y	31	72	42	0	65	42	77	23	352
Foreign	N.C.	W-F	MA	10	18	18	0	22	16	29	5	118
Foreign	N.C.	W-F	O	2	1	2	0	3	3	4	0	15
Foreign	N.C.	B-M	Y	17	14	2	1	2	10	22	3	71
Foreign	N.C.	B-M	MA	3	8	1	0	1	0	3	1	17
Foreign	N.C.	B-M	O	1	0	0	0	0	0	0	0	1
Foreign	N.C.	B-F	Y	2	7	3	0	1	2	9	2	26
Foreign	N.C.	B-F	MA	6	1	1	0	2	2	0	0	12
Foreign	N.C.	B-F	O	0	0	0	0	0	0	0	0	0
Foreign	Other	W-M	Y	9	26	13	0	14	15	42	34	153
Foreign	Other	W-M	MA	3	13	10	0	9	8	25	13	81
Foreign	Other	W-M	O	0	2	5	0	6	4	4	8	29
Foreign	Other	W-F	Y	4	7	4	0	6	4	23	12	60
Foreign	Other	W-F	MA	0	1	1	0	0	2	4	2	10
Foreign	Other	W-F	O	0	0	0	0	1	2	0	0	3
Foreign	Other	B-M	Y	4	3	0	0	1	4	7	0	19
Foreign	Other	B-M	MA	1	0	0	1	1	0	0	1	4
Foreign	Other	B-M	O	0	0	0	0	0	0	1	0	1
Foreign	Other	B-F	Y	1	0	0	0	1	0	1	1	4
Foreign	Other	B-F	MA	0	0	0	0	0	0	0	0	0
Foreign	Other	B-F	O	0	0	0	0	0	0	0	0	0
Total				2158	2633	1581	103	2480	1769	4465	1924	17113

Table 3. Observed number of drivers wearing a shoulder belt.

Vehicle Type	Vehicle Registration	Race-Sex	Age	Location Site X Road Type Combination: Urban 2L	Urban 4NL	Urban 4DL	Urban 4IL	Rural 2L	Rural 4NL	Rural 4DL	Rural 4IL	Total
U.S.	N.C.	W-M	Y	7	8	9	1	16	8	44	19	112
U.S.	N.C.	W-M	MA	12	20	10	6	27	22	45	30	172
U.S.	N.C.	W-M	O	1	2	4	1	4	3	5	6	26
U.S.	N.C.	W-F	Y	4	7	2	0	6	2	9	8	38
U.S.	N.C.	W-F	MA	3	3	1	0	2	3	3	6	21
U.S.	N.C.	W-F	O	2	0	0	0	1	0	0	1	4
U.S.	N.C.	B-M	Y	2	2	0	0	3	1	1	0	9
U.S.	N.C.	B-M	MA	0	1	0	0	0	0	1	0	2
U.S	N.C.	B-M	O	0	0	0	0	0	0	0	0	0
U.S.	N.C.	B-F	Y	1	1	0	0	1	0	0	0	3
U.S.	N.C.	B-F	MA	0	0	0	0	0	0	0	0	0
U.S.	N.C.	B-F	O	0	0	0	0	0	0	0	0	0
U.S.	Other	W-M	Y	1	5	4	0	5	3	13	14	45
U.S.	Other	W-M	MA	0	10	3	5	14	9	29	14	84
U.S.	Other	W-M	O	0	0	1	0	0	1	5	3	10
U.S.	Other	W-F	Y	1	2	1	1	1	1	5	4	16
U.S.	Other	W-F	MA	0	1	1	0	2	1	6	2	13
U.S.	Other	W-F	O	0	0	0	0	1	0	0	0	1
U.S.	Other	B-M	Y	0	0	0	0	0	0	1	7	8
U.S.	Other	B-M	MA	0	0	0	0	0	1	0	0	1
U.S.	Other	B-M	O	0	0	0	0	0	0	0	0	0
U.S.	Other	B-F	Y	0	0	0	0	0	0	0	1	1
U.S.	Other	B-F	MA	0	0	0	0	0	0	0	0	0
U.S.	Other	B-F	O	0	0	0	0	1	0	0	0	1
Foreign	N.C.	W-M	Y	2	9	5	0	10	8	13	12	59
Foreign	N.C.	W-M	MA	3	8	6	2	5	6	11	9	50
Foreign	N.C.	W-M	O	1	0	3	0	0	0	2	0	6
Foreign	N.C.	W-F	Y	3	4	3	0	5	4	10	2	31
Foreign	N.C.	W-F	MA	1	2	1	0	3	2	4	0	13
Foreign	N.C.	W-F	O	0	0	0	0	1	0	0	0	1
Foreign	N.C.	B-M	Y	0	0	0	0	0	0	2	0	2
Foreign	N.C.	B-M	MA	0	0	0	0	0	0	1	0	1
Foreign	N.C.	B-M	O	0	0	0	0	0	0	0	0	0
Foreign	N.C.	B-F	Y	0	1	0	0	0	1	0	0	2
Foreign	N.C.	B-F	MA	0	0	0	0	0	0	0	0	0
Foreign	N.C.	B-F	O	0	0	0	0	0	0	0	0	0
Foreign	Other	W-M	Y	0	6	2	0	2	2	17	10	39
Foreign	Other	W-M	MA	1	1	2	0	3	3	8	6	24
Foreign	Other	W-M	O	0	2	0	0	3	0	0	3	8
Foreign	Other	W-F	Y	1	0	0	0	0	3	1	5	10
Foreign	Other	W-F	MA	0	0	0	0	0	0	0	0	0
Foreign	Other	W-F	O	0	0	0	0	0	0	0	0	0
Foreign	Other	B-M	Y	1	1	0	0	0	0	0	0	2
Foreign	Other	B-M	MA	0	0	0	1	1	0	0	1	3
Foreign	Other	B-M	O	0	0	0	0	0	0	0	0	0
Foreign	Other	B-F	Y	0	0	0	0	1	0	0	0	1
Foreign	Other	B-F	MA	0	0	0	0	0	0	0	0	0
Foreign	Other	B-F	O	0	0	0	0	0	0	0	0	0
	Total			47	96	58	17	118	84	236	163	819

Table 4. Observed shoulder harness usage rates.

Vehicle Type	Vehicle Registration	Race-Sex	Age	Location Site X Road Type Combination: Urban 2L	Urban 4NL	Urban 4DL	Urban 4IL	Rural 2L	Rural 4NL	Rural 4DL	Rural 4IL	Total
U.S.	N.C.	W-M	Y	.024	.025	.032	.067	.039	.034	.064	.081	.045
U.S.	N.C.	W-M	MA	.023	.029	.027	.158	.045	.044	.041	.066	.040
U.S.	N.C.	W-M	O	.007	.018	.049	.250	.028	.027	.023	.061	.029
U.S.	N.C.	W-F	Y	.018	.034	.013	.000	.026	.015	.027	.074	.027
U.S.	N.C.	W-F	MA	.011	.011	.006	.000	.007	.017	.008	.058	.013
U.S.	N.C.	W-F	O	.034	.000	.000	.000	.023	.000	.000	.040	.013
U.S.	N.C.	B-M	Y	.020	.019	.000	.000	.051	.026	.012	.000	.020
U.S.	N.C.	B-M	MA	.000	.019	.000		.000	.000	.020	.000	.008
U.S.	N.C.	B-M	O	.000	.000	.000		.000		.000		.000
U.S.	N.C.	B-F	Y	.018	.024	.000		.067	.000	.000	.000	.016
U.S.	N.C.	B-F	MA	.000	.000	.000		.000	.000	.000	.000	.000
U.S.	N.C.	B-F	O	.000	.000				.000	.000		.000
U.S.	Other	W-M	Y	.038	.043	.111	.000	.069	.053	.068	.111	.072
U.S.	Other	W-M	MA	.000	.060	.043	.333	.083	.074	.074	.061	.069
U.S.	Other	W-M	O	.000	.000	.034		.000	.031	.036	.028	.024
U.S.	Other	W-F	Y	.059	.095	.091	.333	.045	.043	.086	.103	.082
U.S.	Other	W-F	MA	.000	.033	.042	.000	.057	.056	.057	.034	.045
U.S.	Other	W-F	O	.000	.000	.000		.100	.000	.000	.000	.015
U.S.	Other	B-M	Y	.000	.000	.000	.000	.000	.000	.022	.219	.081
U.S.	Other	B-M	MA	.000	.000	.000		.000	.500	.000	.000	.015
U.S.	Other	B-M	O					.000	.000	.000	.000	.000
U.S.	Other	B-F	Y	.000	.000	.000		.000	.000	.000	.125	.038
U.S.	Other	B-F	MA	.000				.000	.000	.000	.000	.000
U.S.	Other	B-F	O					1.000			.000	.500
Foreign	N.C.	W-M	Y	.025	.087	.071	.000	.099	.108	.097	.218	.095
Foreign	N.C.	W-M	MA	.057	.105	.154	.667	.104	.146	.117	.333	.131
Foreign	N.C.	W-M	O	.071	.000	.250		.000	.000	.286	.000	.105
Foreign	N.C.	W-F	Y	.097	.056	.071		.077	.095	.130	.087	.088
Foreign	N.C.	W-F	MA	.100	.111	.056		.136	.125	.138	.000	.110
Foreign	N.C.	W-F	O	.000	.000	.000		.333	.000	.000		.067
Foreign	N.C.	B-M	Y	.000	.000	.000	.000	.000	.000	.091	.000	.028
Foreign	N.C.	B-M	MA	.000	.000	.000		.000		.333	.000	.059
Foreign	N.C.	B-M	O	.000								.000
Foreign	N.C.	B-F	Y	.000	.143	.000		.000	.500	.000	.000	.077
Foreign	N.C.	B-F	MA	.000	.000	.000		.000	.000			.000
Foreign	N.C.	B-F	O									
Foreign	Other	W-M	Y	.000	.231	.154		.143	.133	.405	.294	.255
Foreign	Other	W-M	MA	.333	.077	.200		.333	.375	.320	.462	.296
Foreign	Other	W-M	O		1.000	.000		.500	.000	.000	.375	.276
Foreign	Other	W-F	Y	.250	.000	.000		.000	.750	.043	.417	.167
Foreign	Other	W-F	MA		.000	.000			.000	.000	.000	.000
Foreign	Other	W-F	O					.000	.000			.000
Foreign	Other	B-M	Y	.250	.333			.000	.000	.000		.105
Foreign	Other	B-M	MA	.000			1.000	1.000			1.000	.750
Foreign	Other	B-M	O							.000		.000
Foreign	Other	B-F	Y	.000				1.000		.000	.000	.250
Foreign	Other	B-F	MA									
Foreign	Other	B-F	O									
Total				.022	.036	.037	.165	.048	.047	.053	.085	.048

Table 5. Predicted shoulder harness usage rate.

Vehicle Type	Vehicle Registration	Race-Sex	Age	Urban 2L	Urban 4NL	Urban 4DL	Urban 4IL	Rural 2L	Rural 4NL	Rural 4DL	Rural 4IL
				Location Site X Road Type Combination							
U.S.	N.C.	W-M	Y	.025	.025	.025	.071	.041	.041	.071	.071
U.S.	N.C.	W-M	MA	.025	.025	.025	.071	.041	.041	.041	.071
U.S.	N.C.	W-M	O	.008	.008	.025	.071	.025	.025	.025	.071
U.S.	N.C.	W-F	Y	.025	.025	.025	.071	.025	.025	.025	.071
U.S.	N.C.	W-F	MA	.008	.008	.008	.054	.008	.008	.008	.054
U.S.	N.C.	W-F	O	.008	.008	.008	.054	.008	.008	.008	.054
U.S.	N.C.	B-M	Y	.025	.025	.025	.025	.025	.025	.025	.025
U.S.	N.C.	B-M	MA	.008	.008	.008	.008	.008	.008	.008	.008
U.S.	N.C.	B-M	O	.008	.008	.008	.008	.008	.008	.008	.008
U.S.	N.C.	B-F	Y	.025	.025	.025	.025	.025	.025	.025	.025
U.S.	N.C.	B-F	MA	.008	.008	.008	.008	.008	.008	.008	.008
U.S.	N.C.	B-F	O	.008	.008	.008	.008	.008	.008	.008	.008
U.S.	Other	W-M	Y	.041	.041	.071	.097	.041	.041	.071	.097
U.S.	Other	W-M	MA	.041	.041	.041	.071	.071	.071	.071	.071
U.S.	Other	W-M	O	.008	.008	.025	.025	.025	.025	.025	.025
U.S.	Other	W-F	Y	.041	.041	.071	.097	.041	.041	.071	.097
U.S.	Other	W-F	MA	.041	.041	.041	.071	.071	.071	.071	.071
U.S.	Other	W-F	O	.008	.008	.025	.025	.025	.025	.025	.025
U.S.	Other	B-M	Y	.025	.025	.025	.097	.025	.025	.025	.097
U.S.	Other	B-M	MA	.025	.025	.025	.025	.025	.025	.025	.025
U.S.	Other	B-M	O	.025	.025	.025	.025	.025	.025	.025	.025
U.S.	Other	B-F	Y	.025	.025	.025	.097	.025	.025	.025	.097
U.S.	Other	B-F	MA	.025	.025	.025	.025	.025	.025	.025	.025
U.S.	Other	B-F	O	.025	.025	.025	.025	.025	.025	.025	.025
Foreign	N.C.	W-M	Y	.041	.097	.097	.272	.097	.097	.097	.272
Foreign	N.C.	W-M	MA	.041	.097	.158	.272	.097	.097	.158	.272
Foreign	N.C.	W-M	O	.041	.097	.158	.272	.097	.097	.158	.272
Foreign	N.C.	W-F	Y	.097	.097	.097	.097	.097	.097	.097	.097
Foreign	N.C.	W-F	MA	.097	.097	.097	.097	.097	.097	.097	.097
Foreign	N.C.	W-F	O	.097	.097	.097	.097	.097	.097	.097	.097
Foreign	N.C.	B-M	Y	.041	.041	.041	.041	.041	.041	.041	.041
Foreign	N.C.	B-M	MA	.041	.041	.041	.041	.041	.041	.041	.041
Foreign	N.C.	B-M	O	.041	.041	.041	.041	.041	.041	.041	.041
Foreign	N.C.	B-F	Y	.041	.041	.041	.041	.041	.041	.041	041
Foreign	N.C.	B-F	MA	.041	.041	.041	.041	.041	.041	.041	.041
Foreign	N.C.	B-F	O	.041	.041	.041	.041	.041	.041	.041	.041
Foreign	Other	W-M	Y	.158	.158	.158	.333	.158	.158	.333	.333
Foreign	Other	W-M	MA	.158	.158	.158	.333	.333	.333	.333	.333
Foreign	Other	W-M	O	.158	.158	.158	.333	.333	.333	.333	.333
Foreign	Other	W-F	Y	.097	.097	.097	.333	.097	.097	.097	.333
Foreign	Other	W-F	MA	.097	.097	.097	.333	.097	.097	.097	.333
Foreign	Other	W-F	O	.097	.097	.097	.333	.097	.097	.097	.333
Foreign	Other	B-M	Y	.158	.158	.158	.333	.158	.158	.158	.333
Foreign	Other	B-M	MA	.158	.158	.158	.333	.158	.158	.158	.333
Foreign	Other	B-M	O	.158	.158	.158	.333	.158	.158	.158	.333
Foreign	Other	B-F	Y	.158	.158	.158	.333	.158	.158	.158	.333
Foreign	Other	B-F	MA	.158	.158	.158	.333	.158	.158	.158	.333
Foreign	Other	B-F	O	.158	.158	.158	.333	.158	.158	.158	.333

Table 6. Final clusters of all drivers.

Cluster	Car Make	Registration	Road Type	Race-Sex	Age	Location	Predicted Usage Rate
I	U.S.	N.C. + 0	2L + 4NL	W	0	U	
	U.S.	N.C.	2L + 4NL + 4DL	W-F	MA + 0	U + R	
	U.S.	N.C.	All	B	MA + 0	U + R	0.8%
II	U.S.	N.C.	2L + 4NL + 4DL	W-M	Y + MA	U	
	U.S.	N.C. + 0	2L + 4NL	W-M	0	R	
	U.S.	N.C. + 0	4DL	W-M	0	U + R	
	U.S.	N.C.	2L + 4NL + 4DL	W-F	Y	U + R	
	U.S.	N.C.	All	B	Y	U + R	
	U.S.	0	All	W-F	0	R	
	U.S.	0	4DL + 4IL	W	0	U + R	
	U.S.	0	2L + 4NL + 4DL	B	Y	U + R	
	U.S.	0	All	B	MA + 0	U + R	2.5%
III	U.S.	N.C.	2L + 4NL	W-M	Y + MA	R	
	U.S.	N.C.	4DL	W-M	MA	R	
	U.S.	0	2L + 4NL	W	Y	U + R	
	U.S.	0	2L + 4NL + 4DL	W	MA	U	
	F	N.C.	2L	W-M	All	U	
	F	N.C.	All	B	All	U + R	4.1%
IV	U.S.	N.C.	4IL	W-F	MA + 0	U + R	5.4%
V	U.S.	N.C.	4DL	W-M	Y	R	
	U.S.	N.C.	4IL	W-M	All	U + R	
	U.S.	N.C.	4IL	W-M	Y	U + R	
	U.S.	0	4L + 4NL + 4DL	W	MA	R	
	U.S.	0	4DL	W	Y	U + R	
	U.S.	0	4IL	W	MA	U + R	7.1%
VI	U.S.	0	4IL	All	Y	U + R	
	F	N.C.	2L	W-M	All	R	
	F	N.C.	2L	W-F	All	U + R	
	F	N.C.	4NL	W	All	U + R	
	F	N.C.	4DL	W-M	Y	U + R	
	F	N.C.	4DL + 4IL	W-F	All	U + R	
	F	0	2L + 4NL + 4DL	W-F	All	U + R	9.7%
VII	F	N.C.	4DL	W-M	MA + 0	U + R	
	F	0	2L + 4NL + 4DL	W-M	All	U	
	F	0	2L + 4NL	W-M	Y	R	
	F	0	2L + 4NL + 4DL	B	All	U + R	15.8%
VIII	F	N.C.	4IL	W-M	All	U + R	27.2%
IX	F	0	4DL	W-M	Y	R	
	F	0	2L + 4NL + 4DL	W-M	MA + 0	R	
	F	0	4IL	All	All	U + R	33.3%

percent group are: (1) foreign car drivers with out-of-state registration on four-lane Interstates; (2) mature and older white male foreign car drivers with out-of-state registration on two-lane, four-lane divided, and four-lane undivided highways in rural areas; and (3) younger white male foreign car drivers with out-of-state registration on four-lane divided highways in rural areas. The lowest predicted usage rate (0.8 percent) group contains (1) older white drivers of U.S. cars on two-lane and four-lane undivided highways in urban areas; (2) mature and older white female drivers of U.S. cars with N.C. registration on two-lane, four-lane undivided, and four-lane divided highways; and (3)mature and older black drivers of U.S. cars with N.C. registration. For a description of the drivers in the intermediate clusters refer to Table 6.

Referring to this clustering scheme we examined each of the six variables (car make, road type, location, registration, race-sex, and age) separately while holding the other five constant. When differences existed, foreign car drivers had higher predicted usage rates than U.S. car drivers, drivers in rural areas had higher predicted rates than those in urban areas, and drivers on four-lane Interstates had higher predicted rates than those on four-lane divided highways, who in turn had higher predicted rates than those on two-lane and four-lane undivided highways.

When differences existed between the vehicle registration, drivers of cars with out-of-state registration had higher predicted rates than drivers of N.C. cars except for older white drivers of U.S. cars on four-lane Interstates. In this case, those with out-of-state registration had a rate of 2.5 percent while females and males with N.C. car licenses had predicted rates of 5.4 percent and 7.1 percent respectively.

With two exceptions, predicted rates for younger drivers were greater than or equal to rates for mature drivers which were greater than those for older drivers. The first exception concerns white drivers of U.S. cars with out-of-state registration on two-lane and four-lane undivided rural highways. Younger drivers had a predicted rate of 4.1 percent while the predicted rate for mature drivers was 7.1 percent. The second exception is for white male drivers of foreign cars on two-lane and four-lane undivided rural highways. Younger drivers in this category had a predicted rate of 15.8 percent and mature and older drivers had a predicted rate of 33.3 percent.

White males had higher predicted rates than white females when there was a significant difference, except for foreign car drivers with N.C. registration on two-lane urban highways. In this case, the males had a predicted rate of 4.1 percent while the females had a predicted rate of 9.7 percent. White drivers had predicted rates greater than or equal to their black counterparts, with two exceptions. Older white U.S. car drivers with out-of-state registration on two-lane and four-lane undivided urban highways had a predicted rate of 0.8 percent and the corresponding predicted rate for blacks was 2.5 percent. Also, white female drivers of foreign cars with out-of-state registration on two-lane, four-lane undivided and four-lane divided highways had a predicted rate of 9.7 percent while that for blacks was 15.8 percent.

The previous discussion summarizes the results obtained with the methodology described in Grizzle, Starmer, & Koch (1969). In particular, this technique enables one to identify very specific subgroups of the population that wear shoulder harnesses to a different extent than the population as a whole. These groups have been identified in terms of the clusters defined in this discussion.

One benefit of this approach is that the initial steps taken to identify significant variables to incorporate into the model enables one to look at the trends within each variable considered. It is perhaps meaningful to discuss the implications of some of these intermediate results in greater detail. Perusal of the tables in Appendix A reveals some interesting trends which can be related to Anderson's 1971 study.

One of the most apparent contrasts is between usage rates for U.S. and foreign make cars (Table A2). The drivers of foreign cars wear their shoulder harnesses more than three times as much as drivers of U.S. model cars. This finding is consistent with that of Anderson and generally holds when other variables of interest (age, race, sex, urbanization and road type) are held constant.

One possible explanation of this phenomenon is that the shoulder harness configuration in foreign cars is more convenient (almost universally of the three-point design), unseparable from the lap belt (if you wear one you must wear the other) and comfortable (in many foreign cars there is a pillar between the front and rear seat which facilitates fastening the shoulder harness anchorage at a lower point than on most U.S. cars and thus eliminates the irritation of the shoulder belt rubbing the neck).

This explanation could be tested by contrasting usage of shoulder harnesses in U.S. cars with the three and four-point system. This, in fact, was done and is described in detail in Appendix D. The three point system was worn a little more frequently than the four point (3.2 percent compared to 2.4 percent) but not enough to be statistically significant. However, this does not fully test the postulated explanation, because the U.S. three-point systems of that era only meet the convenience criterion of the explanation. They generally can be separated from the lap belt and are anchored near the roof line.

Another striking feature is the difference between cars of North Carolina and out-of-state registration (Table A1). The out-of-state registration cars have twice the usage rate as the North Carolina cars, which agrees with Anderson's findings. This also generally holds true across the other variables of interest and may be explained if one assumes the out-of-state cars are more likely to be on trips of greater length -- an extension of Bohlin's interview findings that persons more often stated that they wore shoulder harnesses on long trips than shorter ones. Another possible interpretation is that drivers of out-of-state registration cars were likely to be of higher socio-economic level than the corresponding in-state driver group (in that they could afford to be on the longer trips) and thus might also be more aware of the benefit of wearing the restraint device.

Usage rates rise steadily as the road type gets wider, from 3.0 percent for two-lane to 3.8 percent for four-lane to 4.7 percent for four-lane divided to 8.4 percent for Interstate (see Table A3). These road widths may also be construed to reflect different design characteristics, with Interstate roads obviously constituting the highest type of design due to horizontal and vertical alignment standards, median width, access control, etc. These better-designed roads may be associated with longer trips or rurality (see Table A7), with both factors being related to increased belt usage.

Demographic variables are also considered. In this regard, whites tend to use their shoulder belts more often than blacks (see Table A4), and males tend to use theirs more often than do females (see Table A5). Also, shoulder belt usage tends to decrease with an increase in age (see Table A6). These trends hold when controlling for other variables (see Tables A18-A20) and are consistent with Anderson's earlier study.

The inverse relationship of age to usage rate is consistent with Anderson's findings but opposite to the Swedish and Australian findings. Factors possibly relevant to usage rate may be behaving differently in each country. Otherwise, the apparent differences in usage associated with demographic variables like age, race, and sex are difficult to explain except possibly in terms of the extent to which these groups have different perceptions of the potential injury reduction benefits provided by shoulder belts.

Table A34 cross-tabulates passenger usage and driver usage. We find that 76 percent of drivers in cars where the passenger is using the shoulder harness are also wearing it while 59 percent of passengers wore the belts when the driver was wearing belts (see also Table A15). From these data, we do not know who buckled up first. However, it is interesting to note how high the usage rates for one becomes if the other is also wearing the belt system. This confirms that the "follow the leader" effect reported by Council (1974) for seat belts alone also extends to shoulder harnesses. Table A29 also confirms such an interaction for both in-state and out-of-state vehicles.

In summary, the study has illustrated the distressingly low shoulder harness usage rates (4.8 percent overall) in the North Carolina driving population in the absence of interlock devices or mandatory belt usage legislation. Perhaps, the only heartening point that can be drawn is that the younger drivers have higher usage rates than mature or older drivers. Hopefully, as they age their higher usage rate will continue and the new young drivers will continue the trend, thus raising the overall usage rate. Additionally, the "follow the leader" effect may imply that as each person is convinced to start wearing the restraint system, an additional effect may be felt by their drivers and passengers.

REFERENCES

Anderson, T.E. Shoulder belt utilization. Chapel Hill: University of North Carolina Highway Safety Research Center, 1971.

Andreassend, D.C. The effects of compulsory seat belt wearing legislation in Victoria. (Symposium pre-print) Canberra, Australia: National Road Safety Symposium, 1972.

Bohlin, N.I. A statistical analysis of 28,000 accident cases with emphasis on occupant restraint value. 11th Stapp Car Crash Conference. New York: Society of Automotive Engineers, 1967, pp. 299-308.

Bohlin, N.I., Norm, H., & Andersson, A. A statistical traffic accident analysis with reference to occupant restraint value and crashworthiness of the Volvo Experimental Safety Car (VESC). Gothenburg, Sweden: AB Volvo, 1973.

Campbell, B.J. Seat belts and injury reduction in 1967 North Carolina automobile accidents. Chapel Hill: University of North Carolina Highway Safety Research Center, 1968.

Campbell, B.J. Driver injury in automobile accidents involving certain car models. Journal of Safety Research, 1970, 2 (4), 207-228.

Campbell, B.J., Waller, P.F., & Council, F.M. Seat belts: A pilot study of their use under normal driving conditions. Chapel Hill: University of North Carolina Highway Safety Research Center, 1967.

Cochran, W.G. Some methods of strengthening the common χ^2 test. Biometrics, 1954, 10, 417-451.

Commonwealth Bureau of Census and Statistics. Report on a statistical investigation into the effectiveness of seat belts in motor vehicle accidents. Canberra, Australia: Author, 1973.

Council, F.M. Seat belts: A follow-up study of their use under normal driving conditions. Chapel Hill: University of North Carolina Highway Safety Research Center, 1969.

Council, F.M., & Hunter, W.W. Seat belt usage and benefits in North Carolina accidents. Chapel Hill: University of North Carolina Highway Safety Research Center, 1974.

Grizzle, J.E., Starmer, C.F., & Koch, G.G. Analysis of categorical data by linear models. Biometrics, 1969, 25 (3), 489-504.

Koch, G.G., & Reinfurt, D.W. An analysis of the relationship between driver injury and vehicle age for automobiles involved in North Carolina accidents during 1966-1970. Chapel Hill: University of North Carolina Highway Safety Research Center, 1973.

Mantel, N., & Haenszel, W. Statistical aspects of the analysis of data from retrospective studies of disease. Journal of the National Cancer Institute, 1959, 22, 719-748.

Neyman, J. Contribution to the theory of the χ^2 test. Proceedings of Berkeley Symposium on Mathematical Statistics and Probability. Berkeley: University of California Press, 1949.

Road Traffic Board of South Australia. A review of seat belt usage in South Australia: 1964-1971. Victoria, Australia: Author, 1972.

Robertson, L.S. Interlock. Washington, D.C.: Insurance Institute for Highway Safety, 1974.

Vulcan, A.P. Compulsory seat belt wearing in Australia. Paper presented at the Seventeenth Stapp Car Crash Conference, Oklahoma City, November 1973.

APPENDIX A

Preliminary Investigation of Shoulder Harness Usage As Related to Sample Design Variables and Driver Vehicle Variables

As stated in the methodology section of the paper, there were initially eight factors considered in the sampling design (road type, location, observer's position, time, pavement condition, day of week, posted speed limit, and month), as well as seven driver/vehicle variables that were collected once a vehicle was considered eligible for inclusion in the sample (race of driver, sex of driver, age of driver, passenger status, passenger usage, vehicle registration, and make of car). In addition, it was noted whether the driver was using his shoulder harness or not. The purpose of the study is to evaluate the extent to which certain driver related variables (age, race, sex), highway related variables (road type, location) and vehicle-situational variables (make of car, vehicle license, passenger status, passenger usage) have an effect on shoulder harness usage. In this regard, we assume the sample is a representative one with respect to the design variables (observer's position, time, pavement condition, day of week, posted speed limit, and month) which will eventually be omitted from consideration after the presentation of certain preliminary descriptive tabulations. Pearson chi square ($\chi^2{}_p$) tests of association were performed between each of the variables under study and shoulder harness usage of the driver to determine their relative importance. On the basis of these and subsequent tests where we controlled for the dominant variables, we were able to identify the combinations of variables to include in further analyses.

Initial Pearson chi square tests, whose results appear in Tables A1 through A15, reveal that car make and registration are the most significant factors (χ^2_p = 302.18, df = 1 and χ^2_p = 110.72, df = 1 respectively). Significant relationships (α = .05) also exist with respect to road type (χ^2_p = 118.14, df = 3), location (χ^2_p = 75.36, df = 1), race (χ^2_p = 16.52, df = 1), sex (χ^2_p = 37.87, df = 1), and age (χ^2_p = 24.49, df = 2). However, the variable passenger status appeared to have no effect at all on shoulder harness usage (χ^2_p = 37.87, df = 1). Considering the variables in the sample design and their effect on usage, pavement condition (χ^2_p = 3.69, df = 1) and month (χ^2_p = 5.40, df = 2) were not significant, but were, nevertheless, large enough to warrant further investigation. Otherwise, the other variables are clearly significant (α = .05): sampler (or observer's position (χ^2_p = 42.00, df = 1), time (χ^2_p = 45.72, df = 11), day (χ^2_p = 26.43, df = 6), and speed limit (χ^2_p = 117.26, df = 8).

One of the driver-vehicle variables (passenger usage) was considered as a co-dependent variable along with driver usage rather than as an independent variable like the others. All the chi square tests involving this variable are highly significant, which can be seen by referring to Tables A15, A29, A31, and A33. It is analyzed separately in Appendix E.

Tables A16 through A29 give the Pearson chi square statistics computed after controlling for registration. As these tables indicate, statistics for car make and age were significant for both N.C. and out-of-state vehicles while those for race and sex were significant for N.C. vehicles only and those for passenger status for out-of-state vehicles only. In addition, statistics corresponding to the sample design variables location, road type, and day were significant for both N.C. and out-of-state vehicles. Those for sampler location, time, and speed limit were significant for N.C. vehicles and that for month was significant for out-of-state vehicles.

Tables A30 and A31 display the quotients of Pearson chi square statistics divided by their corresponding degrees of freedom (df) after adjustment is made for registration and car make. Significant statistics (α = .05) are designated by an asterisk. Tables A32 and A33 contain the quotients of Pearson chi square statistics divided by their corresponding degrees of freedom (df) after controlling for the three factors of license, car make, and road type. Again asterisks indicate the subpopulations in which significant relationships exist between the variables and shoulder harness usage.

After examining the results of this sequence of tables, we were able to draw the following conclusions. Registration, car make, and road type are the dominant factors affecting shoulder harness usage. After controlling for these, the other four variables of interest -- location, race, age, and sex -- were still significant for some of the factors combinations. Passenger status was significant for only one subpopulation -- out-of-state vehicles on four-lane Interstates. This was regarded as an artifact, and since the chi square statistics across most of the other fifteen subpopulations were very nonsignificant, this variable was not considered in further analyses. In Appendix C we investigated in more detail the pattern of significant relationships revealed in the tables based on combinations of these variables by an analysis of contingency table data developed by Grizzle, Starmer, and Koch (1969).

In addition, while the tables indicate that the sample design variables of observer's position, time, day, month, and speed still tend to have a significant effect on shoulder harness usage, this effect is considerably muted once we control for car make, vehicle license and road type. We have reason to believe that this diminishing trend would continue if we controlled for all the variables of interest (car make, vehicle license, road type, race, sex, age, and location). Thus, we feel the results based on the linear model analysis in Appendix B are robust at least with respect to the conclusions for N.C. vehicles, independent of whether the sample is representative or not.

Table A1

Usage Rates by Registration

Registration	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
In state	15835	580	0.037
Out of state	3503	270	0.077
χ^2_P (df=1) = 110.72			
Combined	19338	850	0.044

Table A2

Usage Rates by Car Make

Car Make	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
U.S.	15636	578	0.037
Foreign	2100	258	0.123
χ^2_P (df=1) = 302.18			
Combined	17736	836	0.047
Missing	1602	14	0.009

Table A3

Usage Rates by Road Type

Road Type	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
2-Lane	5609	166	0.030
4-Lane	5049	190	0.038
4-Div. Lane	6465	306	0.047
Interstate	2205	185	0.084
χ^2_P (df=3) = 118.14			
Combined	19328	847	0.044
Missing	10	3	0.300

Table A4

Usage Rates by Race

Race	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
White	16158	804	0.050
Black	1446	37	0.026
χ^2_P (df=1) = 16.52			
Combined	17604	841	0.048
Missing	1734	9	0.005

Table A5

Usage Rates by Sex

Sex	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
Male	12577	680	0.054
Female	5027	161	0.032
χ^2_P (df=1) = 37.87			
Combined	17604	841	.048
Missing	1734	9	.005

Table A6

Usage Rates by Age

Age	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
Young	6912	390	0.056
Mature	8698	391	0.045
Old	1876	58	0.031
χ^2_P (df=2) = 24.49			
Combined	17486	839	.048
Missing	1852	11	.006

Table A7

Usage Rates by Location

Location	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
Urban	8006	229	0.029
Rural	11302	618	0.055
χ^2_P (df=1) = 75.36			
Combined	19308	847	0.044
Missing	30	3	0.100

Table A8

Usage Rates by Passenger Status

Passenger Status	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
Not Present	11297	490	0.043
Present	7915	353	0.045
χ^2_P (df=1) = 0.14			
Combined	19212	843	0.044
Missing	126	7	0.056

Table A9

Usage Rates by Observer Position

Observer Position	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
Stationary	17969	743	0.041
Enroute	1350	107	0.079
χ^2_P (df) = 42.00			
Combined	19319	850	0.044
Missing	19	0	0.000

Table A10

Usage Rates by Time

Time	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
7am-8am	70	3	0.043
8am-9am	1859	85	0.046
9am-10am	1227	62	0.051
10am-11am	310	24	0.077
11am-noon	1863	86	0.046
noon-1pm	2964	110	0.037
1pm-2pm	2725	105	0.039
2pm-3pm	1772	72	0.041
3pm-4pm	317	33	0.104
4pm-5pm	2772	147	0.053
5pm-6pm	1616	82	0.051
6pm-7pm	2	0	0.000
χ^2_P (df=11) = 45.72			
Combined	17497	809	0.046
Missing	1841	41	0.022

Table A11

Usage Rates by Pavement Condition

Pavement	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
Dry	17431	783	0.045
Wet	1907	67	0.035
χ^2_P (df=1) = 3.69			
Combined	19338	850	0.044
Missing	---	---	---

Table A12

Usage Rates by Day

Day	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
Sunday	730	30	0.041
Monday	5485	278	0.051
Tuesday	5155	213	0.041
Wednesday	5523	243	0.044
Thursday	691	37	0.054
Friday	405	15	0.037
Saturday	1322	27	0.020
χ^2_P (df=6) = 26.43			
Combined	19311	843	0.044
Missing	27	7	0.259

Table A13

Usage Rates by Month

Month	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
June	6374	303	0.048
July	7415	295	0.040
August	5449	249	0.046
χ^2_P (df=2) = 5.40			
Combined	19238	847	0.044
Missing	100	3	0.030

Table A14

Usage Rates by Posted Speed Limit

Posted Limit	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
20	1923	42	0.022
25	371	0	0.000
30	18	0	0.000
35	2102	57	0.027
45	3665	140	0.038
50	129	6	0.047
55	4096	221	0.054
60	2523	150	0.059
65	2120	153	0.072
χ^2_P (df=8) = 117.26			
Combined	16947	769	0.045
Missing	2391	11	0.005

Table A15

Usage Rates by Passenger Use

Passenger Use	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
No Passenger	11297	490	0.043
Passenger Not Wearing	7629	135	0.018
Passenger Wearing	259	197	0.761
χ^2_P (df=2) = 3371.47			
Combined	19185	822	0.043
Missing	153	28	0.183

Table A16

Usage Rates by Car Make and Registration Combinations

In State				Out of State			
Car Make	Number Vehicles Observed	Number Driver Users	Observed Usage Rate	Car Make	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
U.S.	12535	397	0.032	U.S.	3101	181	0.058
Foreign	1731	171	0.099	Foreign	369	87	0.236
χ^2_P = 177.46 (df=1)				χ^2_P = 143.14 (df=1)			
Combined	14266	568	0.040	Combined	3470	268	0.077
Missing	1569	12	0.008	Missing	33	2	0.061

Table A17

Usage Rates by Road Type and Registration Combinations

In State				Out of State			
Road Type	Number Vehicles Observed	Number Driver Users	Observed Usage Rate	Road Type	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
2-Lane	5006	127	0.025	2-Lane	603	39	0.065
4-Lane	4261	137	0.032	4-Lane	788	53	0.067
4-Div. Lane	5124	207	0.040	4-Div. Lane	1341	99	0.074
Interstate	1437	107	0.074	Interstate	768	78	0.102
χ^2_P = 80.98 (df=3)				χ^2_P = 9.06 (df=3)			
Combined	15828	578	0.037	Combined	3500	269	0.077
Missing	7	2	0.286	Missing	3	1	0.333

Table A18

Usage Rates by Race and Registration Combinations

In State				Out of State			
Race	Number Vehicles Observed	Number Driver Users	Observed Usage Rate	Race	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
White	12961	552	0.043	White	3197	252	0.079
Black	1195	19	0.016	Black	251	18	0.072
χ^2_P = 19.45 (df=1)				χ^2_P = 0.08 (df=1)			
Combined	14156	571	0.040	Combined	3448	270	0.078
Missing	1679	9	0.005	Missing	55	0	0.000

Table A19

Usage Rates by Sex and Registration Combinations

In State				Out of State			
Sex	Number Vehicles Observed	Number Driver Users	Observed Usage Rate	Sex	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
Male	9811	454	0.046	Male	2766	226	0.082
Female	4345	117	0.027	Female	682	44	0.065
$\chi^2_P = 28.62$ (df=1)				$\chi^2_P = 2.01$ (df=1)			
Combined	14156	571	0.040	Combined	3448	270	0.078
Missing	1679	9	0.005	Missing	55	0	0.000

Table A20

Usage Rates by Age and Registration Combinations

In State				Out of State			
Age	Number Vehicles Observed	Number Driver Users	Observed Usage Rate	Age	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
Young	5718	266	0.047	Young	1194	124	0.104
Mature	6984	265	0.038	Mature	1714	126	0.074
Old	1347	38	0.028	Old	529	20	0.038
$\chi^2_P = 11.74$ (df=2)				$\chi^2_P = 23.29$ (df=2)			
Combined	14049	569	0.041	Combined	3437	270	0.079
Missing	1786	11	0.006	Missing	66	0	0.000

Table A21

Usage Rates by Location and Registration Combinations

In State				Out of State			
Location	Number Vehicles Observed	Number Driver Users	Observed Usage Rate	Location	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
Urban	7097	173	0.024	Urban	909	56	0.062
Rural	8710	404	0.046	Rural	2592	214	0.083
$\chi^2_P = 53.23$ (df=1)				$\chi^2_P = 3.86$ (df=1)			
Combined	15807	577	0.037	Combined	3501	270	0.077
Missing	28	3	0.107	Missing	2	0	0.000

Table A22

Usage Rates by Passenger Status and Registration Combinations

In State				Out of State			
Passenger Status	Number Vehicles Observed	Number Driver Users	Observed Usage Rate	Passenger Status	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
Not Present	10242	381	0.037	Not Present	1055	109	0.103
Present	5481	194	0.035	Present	2434	159	0.065
	χ^2_P = 0.28 (df=1)				χ^2_P = 14.45 (df=1)		
Combined	15723	575	0.037	Combined	3489	268	0.077
Missing	112	5	0.045	Missing	14	2	0.143

Table A23

Usage Rates by Sampler and Registration Combinations

In State				Out of State			
Sampler	Number Vehicles Observed	Number Driver Users	Observed Usage Rate	Sampler	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
Stationary	14908	513	0.034	Stationary	3061	230	0.075
Enroute	910	67	0.074	Enroute	440	40	0.091
	χ^2_P (df=1) = 36.24				χ^2_P (df=1) = 1.13		
Combined	15818	580	0.037	Combined	3501	270	0.077
Missing	17	0	0.000	Missing	2	0	0.000

Table A24

Usage Rates by Time and Registration Combinations

In State				Out of State			
Time	Number Vehicles Observed	Number Driver Users	Observed Usage Rate	Time	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
7 AM-8 AM	59	3	0.051	7 AM-8 AM	11	0	0.000
8 AM-9 AM	1572	69	0.044	8 AM-9 AM	287	16	0.056
9 AM-10 AM	970	45	0.046	9 AM-10 AM	257	17	0.066
10 AM-11 AM	234	16	0.068	10 AM-11 AM	76	8	0.105
11 AM- noon	1440	58	0.040	11 AM- noon	423	28	0.066
noon-1 PM	2278	67	0.029	noon-1 PM	686	43	0.063
1 PM-2 PM	2170	64	0.029	1 PM-2 PM	555	41	0.074
2 PM-3 PM	1443	50	0.035	2 PM-3 PM	329	22	0.067
3 PM-4 PM	199	18	0.090	3 PM-4 PM	118	15	0.127
4 PM-5 PM	2339	96	0.041	4 PM-5 PM	433	51	0.118
5 PM-6 PM	1434	66	0.046	5 PM-6 PM	182	16	0.088
6 PM-7 PM	1	0	0.000	6 PM-7 PM	1	0	0.000
	χ^2_P (df=11) = 35.85				χ^2_P (df=11) = 22.05		
Combined	14139	552	0.039	Combined	3358	257	0.007
Missing	1696	28	0.017	Missing	145	13	0.090

Table A25

Usage Rates by Pavement Condition and Registration Combinations

In State				Out of State			
Pavement Condition	Number Vehicles Observed	Number Driver Users	Observed Usage Rate	Pavement Condition	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
Dry	14200	530	0.037	Dry	3231	253	0.078
Wet	1635	50	0.031	Wet	272	17	0.062
χ^2_p (df=1) = 1.70				χ^2_p (df=1) = 0.67			
Combined	15835	580	0.037	Combined	3503	270	0.077
Missing	-	-		Missing	-	-	

Table A26

Usage Rates by Day and Registration Combinations

In State				Out of State			
Day	Number Vehicles Observed	Number Driver Users	Observed Usage Rate	Day	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
Sunday	597	17	0.028	Sunday	133	13	0.098
Monday	4556	190	0.042	Monday	929	88	0.095
Tuesday	4130	141	0.034	Tuesday	1025	72	0.070
Wednesday	4450	165	0.037	Wednesday	1073	78	0.073
Thursday	509	31	0.061	Thursday	182	6	0.033
Friday	377	11	0.029	Friday	28	4	0.143
Saturday	1192	19	0.016	Saturday	130	8	0.062
χ^2_p (df=6) = 28.95				χ^2_p (df=6) = 12.98			
Combined	15811	574	0.036	Combined	3500	269	0.077
Missing	24	6	0.250	Missing	3	1	0.333

Table A27

Usage Rates by Month and Registration Combinations

Month	Number Vehicles Observed	Number Driver Users	Observed Usage Rate	Month	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
June	5259	209	0.040	June	1115	94	0.084
July	5868	197	0.034	July	1547	98	0.063
August	4622	172	0.037	August	827	77	0.093
χ^2_p (df=2) = 3.04				χ^2_p (df=2) = 7.90			
Combined	15749	578	0.037	Combined	3489	269	0.077
Missing	86	2	0.023	Missing	14	1	0.071

Table A28

Usage Rates by Speed and Registration Combinations

In State				Out of State			
Speed	Number Vehicles Observed	Number Driver Users	Observed Usage Rate	Speed	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
20	1741	33	0.019	20	182	9	0.049
25	370	0	0.000	25	1	0	0.000
30	18	0	0.000	30	-	-	-
35	1737	40	0.023	35	365	17	0.047
45	3316	110	0.033	45	349	30	0.086
50	98	5	0.051	50	31	1	0.032
55	3217	153	0.048	55	879	68	0.077
60	1843	84	0.046	60	680	66	0.097
65	1431	96	0.067	65	689	57	0.083
χ^2_P (df=8) = 90.25				χ^2_P (df=7) = 12.01			
Combined	13771	521	0.038	Combined	3176	248	0.078
Missing	3064	59	0.029	Missing	327	22	0.067

Table A29

Usage Rates by Passenger Use and Registration Combinations

In State				Out of State			
Passenger Use	Number Vehicles Observed	Number Driver Users	Observed Usage Rate	Passenger Use	Number Vehicles Observed	Number Driver Users	Observed Usage Rate
No Passenger	10242	381	0.037	No Passenger	1055	109	0.103
Passenger Not Wear	5345	89	0.017	Passenger Not Wear	2284	46	0.020
Passenger Wearing	119	92	0.773	Passenger Wearing	140	105	0.750
χ^2_P (df=2) = 1932.38				χ^2_P (df=2) = 1034.10			
Combined	15706	562	0.036	Combined	3479	260	0.075
Missing	129	18	0.140	Missing	24	10	0.417

Table A30

Registration	Car Make	Road Type $\chi^2/(df=3)$	Race $\chi^2/(df=1)$	Sex $\chi^2/(df=1)$	Age $\chi^2/(df=2)$	Location $\chi^2/(df=1)$	Passenger $\chi^2/(df=1)$
In State	U.S.	*17.83	*11.87	*33.05	2.45	*23.14	3.42
In State	Foreign	5.99	* 4.88	0.57	1.79	* 8.96	0.00
Out of State	U.S.	1.74	0.21	0.18	*7.49	2.86	*12.48
Out of State	Foreign	2.47	0.01	* 4.20	.79	* 4.28	.12

Table A31

Registration	Car Make	Observer Location $\chi^2/(df=1)$	Time $\chi^2/(df=11)$	Pavement $\chi^2/(df=1)$	Day $\chi^2/(df=6)$	Month $\chi^2/(df=2)$	Speed $\chi^2/(df=8)$	Pass. Use $\chi^2/(df=2)$
In State	U.S.	*42.73	4.02	0.21	1.64	0.12	5.89[1]	*818.46
In State	Foreign	1.54	0.63[1]	0.02	2.01	3.64	*3.76[1]	* 91.62
Out of State	U.S.	0.01	*1.64	2.24	1.35	4.30	1.40[1]	*391.82
Out of State	Foreign	* 4.67	1.05[1]	0.00	1.48	0.02	1.31[2]	* 68.80

[1] $\chi^2/(df=n-1)$
[2] $\chi^2/(df=n-2)$
[3] $\chi^2/(df=n-3)$
⋮
[7] $\chi^2/(df=n-7)$

Table A32

Registration	Car Make	Road Type	Race $\chi^2/(df=1)$	Sex $\chi^2/(df=1)$	Age $\chi^2/(df=2)$	Location $\chi^2/(df=1)$	Passenger $\chi^2/(df=1)$
In State	U.S.	2-Lane	0.44	* 5.87	.55	*8.18	0.00
In State	U.S.	4-Lane	0.93	* 4.65	.65	0.73	1.10
In State	U.S.	4-Div. Lane	*5.50	*21.83	2.70	*4.81	2.52
In State	U.S.	Interstate	*4.25	0.36	.04	1.88	2.50
In State	Foreign	2-Lane	1.86	0.64	.20	*4.23	.01
In State	Foreign	4-Lane	0.62	0.01	1.30	1.07	.02
In State	Foreign	4-Div. Lane	0.27	0.12	1.59	0.43	1.33
In State	Foreign	Interstate	0.72	3.13	2.22	0.44	.21
Out of State	U.S.	2-Lane	0.01	0.01	.44	*5.94	.26
Out of State	U.S.	4-Lane	0.01	0.01	2.42	0.59	.03
Out of State	U.S.	4-Div. Lane	3.31	0.03	1.73	0.01	.90
Out of State	U.S.	Interstate	1.17	0.35	*6.82	*8.62	*28.20
Out of State	Foreign	2-Lane	0.26	0.04	2.76	0.87	.01
Out of State	Foreign	4-Lane	0.01	0.05	.15	0.01	.07
Out of State	Foreign	4-Div. Lane	1.42	* 7.20	1.53	1.56	.05
Out of State	Foreign	Interstate	0.28	0.01	.63	0.09	.40

Table A33

Residence	Car Make	Road Type	Observer Location $\chi^2/(df=1)$	Time $\chi^2/(df=11)$	Pavement $\chi^2/(df=1)$	Day $\chi^2/(df=6)$	Month $\chi^2/(df=2)$	Speed $\chi^2/(df=8)$	Pass. Use $\chi^2/(df=2)$
In State	U.S.	2-Lane	*37.72	1.27[1]	0.02	*3.78	0.55	*2.50[3]	*238.89
In State	U.S.	4-Lane	1.21	0.67[2]	0.01	*2.20	0.14[2]	1.07[2]	*184.85
In State	U.S.	4-Div. Lane	0.45	2.60[1]	0.10	1.05	0.57	1.66[2]	*276.18
In State	U.S.	Interstate	1.44	1.78[2]	0.00	*3.35	2.79	2.31[4]	* 99.77
In State	Foreign	2-Lane	0.81	1.08[3]	2.31	*2.81	2.21	2.20[3]	* 45.42
In State	Foreign	4-Lane	0.00	1.59[3]	0.06	0.50	*3.24[2]	0.82[3]	* 39.24
In State	Foreign	4-Div. Lane	0.09	0.85[1]	1.79	1.12	*3.60	2.16[3]	* 18.17
In State	Foreign	Interstate	0.00	1.55[2]	0.65	1.05[1]	1.16[2]	0.58[6]	* 3.06
Out of State	U.S.	2-Lane	3.39	1.82[3]	0.08	1.96	2.64	1.35[3]	* 72.29
Out of State	U.S.	4-Lane	1.40	1.19[3]	0.12	0.57	0.11[2]	0.81[3]	*129.31
Out of State	U.S.	4-Div. Lane	0.05	0.84[1]	3.39	*5.07	1.73	0.49[2]	*136.83
Out of State	U.S.	Interstate	3.80	*2.96[1]	0.05	*7.16	*6.50	*5.53[4]	* 82.31
Out of State	Foreign	2-Lane	0.33	1.71[3]	0.17	2.17[1]	*3.07	0.71[5]	* 15.95
Out of State	Foreign	4-Lane	1.83	1.22[4]	0.21	1.52[1]	1.07	0.33[3]	* 13.75
Out of State	Foreign	4-Div. Lane	0.01	1.29[1]	0.11	1.07[2]	0.29	0.27[2]	* 26.66
Out of State	Foreign	Interstate	0.17	1.26[2]	0.28	1.69	0.88	0.44[7]	* 14.08

[1] $\chi^2/(df=n-1)$
[2] $\chi^2/(df=n-2)$
[3] $\chi^2/(df=n-3)$
:
[7] $\chi^2/(df=n-7)$

Table A34

Usage Rates by Driver Use and Passenger Use

	Passenger Use	
Driver Use	Yes	No
Yes	197	135
No	62	7494

χ^2_P (df=1) = 410.78

APPENDIX B

Discussion of General Methodology for Regression Analysis of Proportions

Introduction

This study is concerned with analyzing profiles of shoulder harness usage rates in North Carolina. A general methodology for analyzing such proportions with respect to certin types of effects and their interactions has been discussed by Grizzle, Starmer, and Koch (1969). It is based on fitting linear regression models to the proportions by the method of weighted least squares in a manner which properly accounts for the inherent variability in these quantities. Using this method, it is possible to investigate the relationship between shoulder harness usage and certain factors related to driver characteristics, vehicle make, and road type. Such questions are pursued in the same spirit as that used in analysis of variance and stepwise regression as applied to quantitative data.

The test statistics derived from this type of analysis correspond identically to the minimum modified chi square statistics of Neyman (1949). The basic assumption required is that the sample size is sufficiently large; in particular, each proportion should be based on at least 10 observations and preferably 25 or more. The only other requirement is that there be either a sampling or observational basis for arguing that the inherent variability in the respective proportions can be characterized by binomial probability distributions. A discussion of the mathematical details pertaining to this methodology is given in the following section.

Statistical Methodology

Let $j = 1, 2, \ldots, r$ index a set of categories which correspond to responses associated with a dependent variable of interest. Similarly, let $i = 1, 2, \ldots, s$ index a set of categories which correspond to distinct subpopulations as defined in terms of pertinent independent variables. If samples of size n_i where $i = 1, 2, \ldots, s$ are independently selected from the respective subpopulations, then the resulting data can be summarized in an s x r contingency table as shown in Table B1

Table B1

Subpopulation	1	2	...	r	Totals
1	n_{11}	n_{12}	...	n_{1r}	n_1
2	n_{21}	n_{22}	...	n_{2r}	n_2
...	...	...	...	...	...
s	n_{s1}	n_{s2}	...	n_{sr}	n_s

where n_{ij} denotes the frequency of response category j in the sample from the i-th subpopulation. It then follows that the proportion $p_{ij} = (n_{ij}/n_i)$ is an estimate of the probability that an observation from the i-th subpopulation falls in the j-th response category. In those cases when a particular response category or combination of categories is of interest, the j-subscript can be deleted without loss of generality. Thus, $P_1, P_2, \ldots, P_s$ represent estimated proportions for the response event of interest in the $i = 1, 2, \ldots, s$ subpopulations.

Under the assumption that the inherent variability in the proportions $P_1, P_2, \ldots, P_s$ can be characterized by independent binomial probability distributions, it follows that elements of their corresponding variance-covariance matrix can be estimated as

$$\text{Var}\{P_i\} \hat{=} V_i = \frac{P_i(1-P_i)}{n_i}, \quad \text{Cov}\{P_i, P_{i'}\} = 0 \text{ for } i \neq i'$$

where "$\hat{=}$" means "is estimated by." Thus, in matrix notation, this formulation may be summarized by writing

$$\underset{s\times 1}{\underset{\sim}{P}} = \begin{bmatrix} P_1 \\ P_2 \\ \cdots \\ P_s \end{bmatrix}, \quad \underset{\sim}{\text{Var}}\{\underset{\sim}{P}\} \hat{=} \underset{s\times s}{\underset{\sim}{V}} = \begin{bmatrix} V_1 & 0 & \cdots & 0 \\ 0 & V_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & V_s \end{bmatrix}.$$

The relationship between variation among the proportions $P_1, P_2, \ldots, P_s$ and certain aspects of the nature of the subpopulations can be investigated by fitting linear regression models to the vector $\underset{\sim}{P}$ by the method of weighted least squares. This aspect of the methodology may be characterized by writing

$$\underset{(s\times 1)}{\underset{\sim}{P}} \hat{=} \underset{(s\times t)}{\underset{\sim}{X}}\ \underset{(t\times 1)}{\underset{\sim}{b}}$$

where $\underset{\sim}{X}$ is the design (or independent variable) matrix of full rank t and $\underset{\sim}{b}$ is the tx1 vector of estimated parameters (or effects). The estimators $\underset{\sim}{b}$ are determined as

$$\underset{\sim}{b} = (\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{P}$$

where $\underset{\sim}{b}$ minimizes the quadratic function

$$\chi^2_W = (\underset{\sim}{P}-\underset{\sim}{X}\underset{\sim}{b})'\underset{\sim}{V}^{-1}(\underset{\sim}{P}-\underset{\sim}{X}\underset{\sim}{b}).$$

The variance-covariance matrix of $\underset{\sim}{b}$ is consistently estimated by

$$\underset{\sim}{V}_{\underset{\sim}{b}} = (\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1}.$$

Justification for a linear regression model is provided by the residual sum of squares χ^2_W. If the model fits, χ^2_W is distributed approximately as χ^2 with df = g-t.

When an appropriate model has been determined, statistical tests of significance involving $\underset{\sim}{b}$ may be performed by analogous standard multiple regression procedures. Linear hypotheses are formulated as H_0: $\underset{\sim}{C}\underset{\sim}{b} \hat{=} \underset{\sim}{0}$, where $\underset{\sim}{C}$ is a known (dxt) coefficient matrix, and tested using the statistic

$$\chi^2_{WC} = \underset{\sim}{b}'\underset{\sim}{C}'[\underset{\sim}{C}(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1}\underset{\sim}{C}']^{-1}\underset{\sim}{C}\underset{\sim}{b}$$

which is approximately distributed as χ^2 with df = d when the hypothesis H_0 is true.

Successive uses of the goodness of fit test and the significance tests specified by the $\underset{\sim}{C}$ matrices represent ways of partitioning the model components into specific sources of variance. In this context, the $\underset{\sim}{C}$ matrices reflect the amount the residual sum of squares χ^2_W would increase if one simplified (or reduced) the model by substituting in the conditions described by H_0: $\underset{\sim}{C}\underset{\sim}{b} = \underset{\sim}{0}$. This partitioning of total variance into specific sources represents a statistically valid analysis of variance for proportions.

Finally, predicted values corresponding to any specific model can be calculated from

$$\underset{\sim}{\hat{P}} = \underset{\sim}{X}\,\underset{\sim}{b} = X(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{P}$$

and corresponding estimates of variance can be obtained from the diagonal elements of

$$\underset{\sim}{V} = \underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-1}\underset{\sim}{X}'.$$

This type of linear model analysis can be undertaken by using a computer program written and used extensively in the Biostatistics Department at the Universitv of North Carolina. The program was used in this study of shoulder harness usage rates. In cases where cell sizes $n_{ij} = 0$, it was necessary to replace them by 1/2 in order to prevent $\underset{\sim}{V}$ from being singular.

General Remarks

The advantage of performing this type of analysis on proportions is that it ultimately leads to predicted values for the original proportions which for most purposes are different only if the corresponding observed values are significantly different. These predicted values not only have the advantage of characterizing essentially all the important features of the variations in the original data, but also represent better estimates than the original proportions since they are based on the data for the entire sample as opposed to its component parts. Finally, they are descriptively advantageous in the sense that they make trends more apparent and permit a clearer interpretation of the effects of the respective independent variables comprising $\underset{\sim}{X}$ on the vector $\underset{\sim}{P}$.

In addition to the Q statistic described in the previous section for assessing the goodness of fit of a linear regression model, it is also useful to determine a statistic which reflects the amount of variation explained by the model. On an absolute basis this is given by the difference between the total variation statistic and the goodness of fit statistic. Thus, an appropriate relative measure of percent explained variation is the ratio of this difference with respect to the corresponding total variation.

APPENDIX C

Application of Regression Analysis Methodology for Proportions to Shoulder Harness Data

Introduction

As described in the body of the paper, the purpose of this study is to investigate patterns of shoulder harness usage in the driving population of North Carolina. In order to evaluate the effects of certain driver, road, and vehicle characteristics on usage rates, a sample survey was conducted and data collected by two-man observation teams throughout the state. Details concerning the sampling scheme and data collection process are described in the paper. Once a vehicle was determined to be eligible for inclusion in the sample (i.e., presence of factory installed head restraints) information on the following variables was gathered: 1) race of driver (white, black); 2) sex of driver (male, female); 3) age of driver (younger -- 16 to 35; mature -- 36 to 55; and older -- 56+); 4) passenger status (right front passenger or not); 5) passenger usage (right front passenger using shoulder harness or not); 6) registration of car (N.C. or other) 7) make of car (U.S. or foreign); and 8) driver usage (driver using shoulder harness or not). In addition, data relating to factors considered in the sampling design were recorded: 1) road type; 2) month; 3) location; 4) observer's position (moving or stationary); 5) time; 6) pavement condition; 7) day of week; and 8) posted speed limit. The objective of the statistical analysis is to assess the extent to which the above variables have an effect on shoulder harness usage.

Preliminary chi square tests of association were computed between each of the variables and shoulder harness usage. On the basis of these tests the following seven driver, road, and vehicle factors were determined to be of interest for further study (see Appendix A for details): 1) race; 2) sex; 3) age; 4) license; 5) car make; 6) road type; and 7) location. A multidimensional contingency table was then generated where the rows corresponded to the 384 different factor combinations of the seven variables, and the columns refer to the two levels of shoulder harness usage (used, not used).

A method for analyzing contingency table data has been developed by Grizzle, Starmer, and Koch (1969; hereafter abbreviated as GSK), which is based on fitting linear regression models to functions of the cell proportions by the method of weighted least squares. A detailed mathematical discussion of the methodology can be found in Appendix B. We shall present here a discussion of how the method was applied in this particular study.

Referring to the notation and terminology of Appendix B, the functions of interest $\underset{\sim}{F}$ in this case are the proportions corresponding to shoulder harness usage. We then fit a series of linear regression models to the functions by the method of weighted least squares. This can be represented in matrix notation as

$$\underset{g\times 1}{\underset{\sim}{F}} = \underset{g\times t}{\underset{\sim}{X}}\ \underset{t\times 1}{\underset{\sim}{b}}$$

where $\underset{\sim}{X}$ is a matrix of known coefficients that indicates the ways in which the estimated effect parameters in $\underset{\sim}{b}$ explain the variation of the values of $\underset{\sim}{F}$. Thus the role of driver, vehicle, and road factors in shoulder harness usage can be evaluated by constructing $\underset{\sim}{X}$ so that the elements of $\underset{\sim}{b}$ correspond to such effects. To evaluate the suitability of each model, a goodness of fit statistic, the residual sum of squares, was computed. If the model fits, it is distributed approximately as χ^2 with df = g - t. Once an appropriate model (i.e., one with as few parameters as possible that characterizes all the essential features of the variation in the data) was determined, statistical tests of significance involving the $\underset{\sim}{b}$ were performed. They are formulated as $H_0: \underset{\sim}{C}\hat{\underset{\sim}{b}} \hat{=} \hat{0}$ where $\underset{\sim}{C}$ is a known (d x t) coefficient matrix and are tested using a statistic which is distributed as χ^2 with df = d when the hypothesis is true. Predicted values corresponding to any specific model $\underset{\sim}{X}$ can be calculated from

$$\hat{\underset{\sim}{P}} = \underset{\sim}{X}\underset{\sim}{b}.$$

Owing to the small sample sizes in certain cells, some road type, race-sex, age, and location categories were combined after chi square tests for homogeneity were performed to justify the collapsing. The 384 x 2 table was then reduced to a 78 x 2 table. For ease of analysis this table was broken up into sections or modules of data which were examined separately then recombined for a final model fit. The preliminary chi square analyses discussed in Appendix A indicate that car make was the most significant factor, (χ^2_p = 302.18, df = 1) followed by vehicle registration (χ^2_p = 110.72, df = 1) and road type (χ^2_p = 118.14, df = 3). Further examination of the other factors within car make, type of license, and road type indicated the following breakdown is appropriate:

1. U.S. cars with N.C. registration
   a) white males on two-lane and four-lane undivided highways
   b) white males on four-lane divided highways
   c) white males on four-lane Interstates
   d) white females
   e) blacks
2. U.S. cars with out-of-state registration
3. Foreign cars with N.C. registration
   a) white males on two-lane highways
   b) white males on four-lane divided highways
   c) white males on four-lane undivided highways
   d) white males on four-lane Interstates
   e) white females
   f) blacks
4. Foreign cars with out-of-state registration

Results from the analyses of the above groups were applied to the overall analysis of U.S. and foreign car drivers, which in turn led to a final analysis of all drivers considered together. At each stage of the analysis, we used the GSK method to develop linear models that would define homogeneous clusters of drivers that would account for most of the variation in the data. As we combined subpopulations of drivers to form four larger groups (North Carolina registered U.S. car drivers; out-of-state registered U.S. car drivers; North Carolina registered foreign car drivers; out-of-state registered foreign car drivers), and then two groups (all U.S. car drivers; all foreign car drivers), and finally one group (all drivers), special attention was applied to similarities among the clusters in defining overall models. In the following sections, we describe the models and corresponding clusters developed at each stage, concluding with the final model.

Drivers of U.S. Cars with N.C. Registration

White males on two-lane and four-lane undivided highways.

Since total variation (χ^2_W = 14.13, df = 5) is significant, differences among the observed usage rates may be said to exist and further consideration of how to characterize these differences is appropriate. The observed usage rates suggest that younger and mature drivers have a higher usage rate than older drivers, and that within each age group rural driver rates are consistently higher than those for urban drivers. Our preliminary model includes an overall mean, two age effects, a location effect and two interaction (age x location) effects. Statistical tests of significance on the age effects (b_2 and b_3) and the location effect (b_4) do not give a clear indication of significance or nonsignificance ($.05 < p < .25$), but do indicate the equality of the two age effects; the two interaction effects, however, are definitely nonsignificant ($p \geq .25$). The final model, therefore, includes an overall mean, one age effect (younger and mature vs. older) and a location effect. The fit is good with a residual chi square of .62, df = 3, and both parameters are significant, $\alpha = .05$. Under this model, predicted usage rates are the highest for younger and mature rural drivers (4.1 percent), followed by older rural drivers (2.8 percent), younger and mature urban drivers (2.6 percent), and finally older urban drivers (1.2 percent).

<u>White males on four-lane divided highways</u>.

Since the total variation ($\chi^2_W = 12.41$) is statistically significant with respect to a χ^2 distribution with 5 degrees of freedom, differences among the usage rates may be said to exist. Following the same preliminary analysis as for white males on two-lane and four-lane undivided highways, we begin with a model that has two age effects, a location effect and two interaction effects. Both the age effects and the location effect are non-significant ($\chi^2_{WC} = .81$,

TABLE C1

Road Type	Race-Sex	Age	Location	Number Observed	Number Users	Observed Usage Rate	Est. s.e.	Predicted Usage Rate	Est. s.e.
2L + 4NL	W-M	Y	U	612	15	.025	.006	.026	.004
2L + 4NL	W-M	Y	R	645	24	.037	.007	.041	.004
2L + 4NL	W-M	MA	U	1213	32	.026	.005	.026	.004
2L + 4NL	W-M	MA	R	1101	49	.045	.006	.041	.004
2L + 4NL	W-M	O	U	247	3	.012	.007	.012	.006
2L + 4NL	W-M	O	R	255	7	.027	.010	.028	.007

Preliminary Model

$$\underset{\sim}{X} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}, \quad \underset{\sim}{b} = \begin{bmatrix} .012 \\ .012 \\ .014 \\ .015 \\ -.003 \\ -.003 \end{bmatrix}$$

Statistical Tests	D.F.	X^2_W
$b_2, b_3 \hat{=} 0$	1	3.02
$b_4 \hat{=} 0$	1	1.53
$b_2 - b_3 \hat{=} 0$	1	.06
$b_5, b_6 \hat{=} 0$	2	.19
$b_5 - b_6 \hat{=} 0$	1	.19
Total Variation	5	14.13

Final Model

$$\underset{\sim}{X} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad \underset{\sim}{b} = \begin{bmatrix} .012 \\ .014 \\ .016 \end{bmatrix}$$

Statistical Tests	D.F.	X^2_W
$b_2 \hat{=} 0$	1	4.51
$b_3 \hat{=} 0$	1	8.35
Model	2	13.51
Residual	3	.62

df = 2 and χ^2_{WC} = 1.04, df = 1, respectively). Although the test for the interaction is not clearly significant, it was large enough to motivate further investigation. Our final model includes an overall mean, a location effect for younger and mature drivers, and an additional effect to account for the exceptionally high usage proportion for young rural drivers. The residual chi square is non-significant, χ^2_W = 1.23 and df = 3, indicating that the model fits. The effect for younger rural drivers is significant, but not the location effect for younger and mature drivers. However, the location effect is large enough to warrant its inclusion in the model in order to maintain consistency with analagous results for two-lane and four-lane undivided highways. Predicted usage rates are lowest (2.8 percent) for older drivers and younger and mature urban drivers. The second highest rate is for mature rural drivers (4.1 percent), and the highest rate corresponds to the younger rural drivers (6.4 percent).

White males on four-lane Interstates.

Since total variation in the preliminary model is non-significant, differences among the three age groups do not appear to exist. Thus, our final involves only an overall mean parameter, yielding a predicted usage rate of 7.5 percent.

TABLE C2

Road Type	Race-Sex	Age	Location	Number Observed	Number Users	Observed Usage Rate	Est. s.e.	Predicted Usage Rate	Est. s.e.
4DL	W-M	Y	U	280	9	.032	.011	.028	.005
4DL	W-M	Y	R	692	44	.064	.009	.064	.009
4DL	W-M	MA	U	370	10	.027	.008	.028	.005
4DL	W-M	MA	R	1090	45	.041	.006	.041	.006
4DL	W-M	O	U	81	4	.049	.024	.028	.005
4DL	W-M	O	R	220	5	.023	.010	.028	.005

Preliminary Model

$$\underset{\sim}{X} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}, \quad \underset{\sim}{b} = \begin{bmatrix} .049 \\ -.017 \\ -.022 \\ -.027 \\ .058 \\ .041 \end{bmatrix}$$

Statistical Tests	D.F.	X^2_W
$b_2, b_3 \hat{=} 0$	2	.81
$b_4 \hat{=} 0$	1	1.04
$b_2 - b_3 \hat{=} 0$	1	.14
$b_5, b_6 \hat{=} 0$	2	3.91
$b_5 - b_6 \hat{=} 0$	1	.97
Total Variation	5	12.41

Final Model

$$\underset{\sim}{X} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \underset{\sim}{b} = \begin{bmatrix} .028 \\ .013 \\ .022 \end{bmatrix}$$

Statistical Tests	D.F.	X^2_W
$b_2 \hat{=} 0$	1	2.61
$b_3 \hat{=} 0$	1	4.06
Model	2	11.18
Residual	3	1.23

TABLE C3

Road Type	Race-Sex	Age	Location	Number Observed	Number Users	Observed Usage Rate	Est. s.e.	Predicted Usage Rate	Est. s.e.
4IL	W-M	Y	U + R	249	20	.080	.017	.075	.009
4IL	W-M	MA	U + R	491	36	.073	.012	.075	.009
4IL	W-M	O	U + R	103	7	.068	.025	.075	.009

Preliminary Model

$$\underset{\sim}{X} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \quad \underset{\sim}{b} = \begin{bmatrix} .068 \\ .012 \\ .005 \end{bmatrix}$$

Statistical Tests	D.F.	X^2_W
Total Variation	2	.19

Final Model

$$\underset{\sim}{X} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \underset{\sim}{b} = [.075]$$

Statistical Tests	D.F.	X^2_W
Model	1	.19

White males.

The individual models for each of the road types were combined to form a preliminary overall model for white male drivers of U.S. cars with North Carolina registration. Statistical tests were performed to detect any structural differences or similarities among the three groups. The results of these tests motivated the final model, which fits, with a residual chi square of 2.92 and 12 degrees of freedom, and has significant effect parameters. This model divides the white males driving U.S. cars with N.C. registration into four clusters (see Table C4).

White females.

For the white females, analyses on the three road type groups separately did not indicate any significant total variation among the estimated proportions. Because of the relatively small sample sizes compared to the white male drivers, we combined all the white female drivers and observed that across the four road types the younger drivers had higher usage rates than their mature and older counterparts. Moreover, usage rates on four lane Interstates were considerably higher than those for the other highway types. Thus, we fit a model with two means -- one for four-lane Interstates and one for the other road types -- as well as an age effect (younger vs. mature and older). Statistical tests confirm the significance of the age effect as well as the difference between the two means.

TABLE C4

Final Clusters of White Male Drivers of U.S. Cars with a N.C. License

Cluster	Road Type	Age	Location	Predicted Usage Rate
I	2L + 4NL	O	U	1.2%
II	2L + 4NL + 4DL	Y	U	
	2L + 4NL + 4DL	MA	U	
	2L + 4NL	O	R	
	4DL	O	U + R	2.7%
III	2L + 4NL	Y	R	
	2L + 4NL + 4DL	MA	R	4.1%
IV	4DL	Y	R	
	4IL	All	U + R	6.9%

TABLE C5

Road Type	Race-Sex	Age	Location	Number Observed	Number Users	Observed Usage Rate	Est. s.e.	Predicted Usage Rate	Est. s.e.
2L + 4NL	W-M	Y	U	612	15	.025	.006	.027	.002
2L + 4NL	W-M	Y	R	645	24	.037	.007	.041	.003
2L + 4NL	W-M	MA	U	1213	32	.026	.005	.027	.002
2L + 4NL	W-M	MA	R	1101	49	.045	.006	.041	.003
2L + 4NL	W-M	O	U	247	3	.012	.007	.012	.005
2L + 4NL	W-M	O	R	255	7	.027	.010	.027	.002
4DL	W-M	Y	U	280	9	.032	.011	.027	.002
4DL	W-M	Y	R	692	44	.064	.009	.069	.006
4DL	W-M	MA	U	370	10	.027	.008	.027	.002
4DL	W-M	MA	R	1090	45	.041	.006	.041	.003
4DL	W-M	O	U	81	4	.049	.024	.027	.002
4DL	W-M	O	R	220	5	.023	.010	.027	.002
4IL	W-M	Y	U + R	249	20	.080	.017	.069	.006
4IL	W-M	MA	U + R	491	36	.073	.012	.069	.006
4IL	W-M	O	U + R	103	7	.068	.025	.069	.006

Preliminary Model

$$\underset{\sim}{X} = \begin{bmatrix} 1 & 1 & 0 & & & & \\ 1 & 1 & 1 & & & & \\ 1 & 1 & 0 & & & & \\ 1 & 1 & 1 & & & & \\ 1 & 0 & 0 & & & & \\ 1 & 0 & 1 & & & & \\ & & & 1 & 0 & 0 & \\ & & & 1 & 1 & 1 & \\ & & & 1 & 0 & 0 & \\ & & & 1 & 1 & 0 & \\ & & & 1 & 0 & 0 & \\ & & & 1 & 0 & 0 & \\ & & & & & & 1 \\ & & & & & & 1 \\ & & & & & & 1 \end{bmatrix}, \quad \underset{\sim}{b} = \begin{bmatrix} .012 \\ .014 \\ .016 \\ .028 \\ .013 \\ .022 \\ .075 \end{bmatrix}$$

Statistical Tests	D.F.	X^2_W
$b_1 + b_2 - b_4 \hat{=} 0$	1	0.15
$b_3 - b_5 \hat{=} 0$	1	0.07
$b_2 - b_3 \hat{=} 0$	1	0.05
$b_4 + b_5 + b_6 - b_7 \hat{=} 0$	1	0.71
$b_1 + b_2 - b_4,\ b_3 - b_5,\ b_2 - b_3,\ b_4 + b_5 + b_6 - b_7$	4	0.87
$b_2 \hat{=} 0$	1	4.51
$b_3 \hat{=} 0$	1	8.35
$b_5 \hat{=} 0$	1	2.61
$b_6 \hat{=} 0$	1	4.06
Model	6	50.56
Residual	8	2.05

Final Model

$$
\underset{\sim}{X} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \\ 1 & 2 & 1 \\ 1 & 2 & 1 \end{bmatrix} , \quad \underset{\sim}{b} = \begin{bmatrix} .012 \\ .015 \\ .028 \end{bmatrix}
$$

Statistical Tests	D.F.	X^2_W
$b_2 \mathrel{\hat{=}} 0$	1	17.22
$b_3 \mathrel{\hat{=}} 0$	1	14.36
Model	2	49.69
Residual	12	2.92

TABLE C6

Final Clusters for White Female Drivers of U.S. Cars with a N.C. License

Cluster	Road Type	Age	Location	Predicted Usage Rate
I	2L + 4NL + 4DL	MA + O	U + R	.9%
II	2L + 4NL + 4DL	Y	U + R	2.2%
III	4IL	MA + O	U + R	5.3%
IV	2L + 4NL + 4DL	Y	U + R	6.7%

TABLE C7

Road Type	Race-Sex	Age	Location	Number Observed	Number Users	Observed Usage Rate	Est. s.e.	Predicted Usage Rate	Est. s.e.
2L + 4NL	W-F	Y	U	431	11	.026	.008	.022	.004
2L + 4NL	W-F	Y	R	360	8	.022	.008	.022	.004
2L + 4NL	W-F	MA + O	U	635	8	.013	.004	.009	.002
2L + 4NL	W-F	MA + O	R	540	6	.011	.005	.009	.002
4DL	W-F	Y	U	156	2	.013	.009	.022	.004
4DL	W-F	Y	R	331	9	.027	.009	.022	.004
4DL	W-F	MA + O	U	209	1	.005	.005	.009	.002
4DL	W-F	MA + O	R	457	3	.007	.004	.009	.002
4IL	W-F	Y	U + R	116	8	.069	.024	.067	.015
4IL	W-F	MA + O	U + R	134	7	.052	.019	.053	.015

Final Model

$$\underset{\sim}{X} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad \underset{\sim}{b} = \begin{bmatrix} .009 \\ .014 \\ .053 \end{bmatrix}$$

Statistical Tests	D.F.	X^2_W
$b_2 \hat{=} 0$	1	8.82
$b_1 - b_3 \hat{=} 0$	1	8.88
Model	2	18.76
Residual	7	3.65

Moreover, the residual chi square is nonsignificant (χ^2_W = 3.65, df = 7) indicating that the model fits and confirming the existence of the trend we noted in the observed proportions. This has the effect of grouping the white females into four clusters.

Black drivers.

For black drivers the total variation statistic (χ^2_N = 4.40, df = 1) indicates there is a significant difference between the usage rates for younger vs. mature and older drivers. We therefore use the two parameter model, which simply re-identifies the usage rates with themselves, as our final model to preserve this difference.

TABLE C8

Road Type	Race-Sex	Age	Location	Number Observed	Number Users	Observed Usage Rate	Est. s.e.	Predicted Usage Rate	Est. s.e.
ALL	B	Y	U + R	639	12	.019	.005	.019	.005
ALL	B	MA + O	U + R	387	2	.005	.004	.005	.004

Final Model

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \quad \underset{\sim}{b} = \begin{bmatrix} .019 \\ .005 \end{bmatrix}$$

Statistical Tests	D.F.	χ^2_W
$b_2 \hat{=} 0$	1	4.40

Overall model for drivers of U.S. cars with N.C. registration.

The final model is derived by combining the overall models for white female, white male, and black drivers, rearranging their respective rows to correspond to the order of the subpopulations given in Table C10. It contains an overall mean and two effects, both of which are significant. The model fits, χ^2_N = 9.57 and df = 24, and accounts for approximately 94 percent of the variation in the data.

Five clusters of the subpopulations defined by combinations of road type, race, sex, age, and location are identified with corresponding predicted usage rates of 0.8 percent, 2.4 percent, 4.1 percent, 5.3 percent, and 6.9 percent.

TABLE C9

Final Clusters of All Drivers of U.S. Cars with a N.C. License

Cluster	Road Type	Race-Sex	Age	Location	Predicted Usage Rate
I	2L + 4NL	W-M	O	U	
	2L + 4NL + 4DL	W-F	MA + O	U + R	
	All	B	MA + O	U + R	.8%
II	2L + 4NL	W-M	Y + MA	U	
	2L + 4NL	W-M	O	R	
	2L + 4NL	W-F	Y	U + R	
	4DL	W-M	Y + MA	U	
	4DL	W-M	O	U + R	
	4DL	W-F	Y	U + R	
	All	B	Y	U + R	2.4%
III	2L + 4NL	W-M	Y + MA	R	
	4DL	W-M	MA	R	4.1%
IV	4IL	W-F	MA + O	U + R	5.3%
V	4DL	W-M	Y	R	
	4IL	W-M	All	U + R	
	4IL	W-F	Y	U + R	6.9%

TABLE C10

Road Type	Race-Sex	Age	Location	Number Observed	Number Users	Observed Usage Rate	Est. s.e.	Predicted Usage Rate	Est. s.e.
2L + 4NL	W-M	Y	U	612	15	.025	.006	.024	.002
2L + 4NL	W-M	Y	R	645	24	.037	.007	.041	.003
2L + 4NL	W-M	MA	U	1213	32	.026	.005	.024	.002
2L + 4NL	W-M	MA	R	1101	49	.045	.006	.041	.003
2L + 4NL	W-M	O	U	247	3	.012	.007	.008	.002
2L + 4NL	W-M	O	R	255	7	.027	.010	.024	.002
2L + 4NL	W-F	Y	U	431	11	.026	.008	.024	.002
2L + 4NL	W-F	Y	R	360	8	.022	.008	.024	.002
2L + 4NL	W-F	MA + O	U	.635	8	.013	.004	.008	.002
2L + 4NL	W-F	MA + O	R	540	6	.011	.005	.008	.002
4DL	W-M	Y	U	280	9	.032	.011	.024	.002
4DL	W-M	Y	R	692	44	.064	.009	.069	.006
4DL	W-M	MA	U	370	10	.027	.008	.024	.002
4DL	W-M	MA	R	1090	45	.041	.006	.041	.003
4DL	W-M	O	U	81	4	.049	.024	.024	.002
4DL	W-M	O	R	220	5	.023	.010	.024	.002
4DL	W-F	Y	U	156	2	.013	.009	.024	.002
4DL	W-F	Y	R	331	9	.027	.009	.024	.002
4DL	W-F	MA + O	U	209	1	.005	.005	.008	.002
4DL	W-F	MA + O	R	457	3	.007	.004	.008	.002
4IL	W-M	Y	U + R	249	20	.080	.017	.069	.006
4IL	W-M	MA	U + R	491	36	.073	.012	.069	.006
4IL	W-M	O	U + R	103	7	.068	.025	.069	.006
4IL	W-F	Y	U + R	116	8	.069	.024	.069	.006
4IL	W-F	MA + O	U + R	134	7	.052	.019	.053	.006
All	N-W	Y	U + R	639	12	.019	.005	.024	.002
All	N-W	MA + O	U + R	387	2	.005	.004	.008	.002

$$
\underset{\sim}{X} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 2 & 1 \\ 1 & 2 & 1 \\ 1 & 2 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \underset{\sim}{b} = \begin{bmatrix} .008 \\ .017 \\ .028 \end{bmatrix}
$$

Final Model

Source of Variation	D.F.	χ^2_W
$b_2 = 0$	1	78.33
$b_3 = 0$	1	18.54
Model	2	145.37
Residual	24	9.57
Total	26	154.94

Drivers of U.S. Cars with Out-of-State Registration

Since total variation is significant with respect to a chi square distribution with 17 degrees of freedom, significant differences exist among the subpopulations of U.S. car drivers with an out-of-state registration. Proceeding in a similar manner as for U.S. drivers with N.C. licenses, we did preliminary analyses on subset nodules of the data to investigate the nature of these differences. These analyses led to the final model which has an overall mean and two incremental effects, smoothing the data so that the predicted usage rates for the driver subpopulations fall into four clusters with respective rates of 1.7 percent, 4.5 percent, 7.3 percent, and 12.2 percent. The model fits (residual χ^2_N = 6.67, df = 15) and accounts for approximately 90 percent of the variation.

Drivers of Foreign Cars with N.C. Registration

White males on two-lane highways.

Although total variation is nonsignificant, χ^2_W = 6.02, df = 3, it is apparent that on two-lane highways rural drivers of foreign cars with N.C. registration have higher usage rates than urban drivers. It was also observed in other driver populations discussed previously that when differences did exist between urban and rural drivers the latter had higher rates. At this stage we therefore

Table C11

Final Clusters of All Drivers of U.S. Cars with Other Licenses

Cluster	Road Type	Race-Sex	Age	Location	Predicted Usage Rate
I	All	W	O	U + R	
	2L+4NL+4DL	B	Y	U + R	
	All	B	MA + O	U + R	1.7%
II	2L+4NL	W	Y	U + R	
	2L+4NL+4DL	W	MA	U	4.5%
III	2L+4NL	W	MA	R	
	4DL	W	Y	U + R	
	4DL	W	MA	R	
	4IL	W	MA	U + R	7.3%
IV	4IL	W + B	Y	U + R	12.2%

TABLE C12

Road Type	Race-Sex	Age	Location	Number Observed	Number Users	Observed Usage Rate	Est. s.e.	Predicted Usage Rate	Est. s.e.
2L + 4NL	W	Y	U	179	9	.050	.016	.045	.004
2L + 4NL	W	Y	R	174	10	.057	.018	.045	.004
2L + 4NL	W	MA	U	280	11	.039	.012	.045	.004
2L + 4NL	W	MA	R	343	26	.076	.014	.073	.006
2L + 4NL	W	O	U	78	0	.000	.009	.017	.005
2L + 4NL	W	O	R	91	2	.022	.015	.017	.005
4DL	W	Y	U	47	5	.106	.045	.073	.006
4DL	W	Y	R	248	18	.073	.016	.073	.006
4DL	W	MA	U	94	4	.043	.021	.045	.004
4DL	W	MA	R	495	35	.071	.011	.073	.006
4DL	W	O	U	35	1	.029	.028	.017	.005
4DL	W	O	R	157	5	.032	.014	.017	.005
4IL	W	Y	U + R	171	19	.111	.024	.122	.022
4IL	W	MA	U + R	306	21	.069	.014	.073	.006
4IL	W	O	U + R	122	3	.025	.014	.017	.005
2L + 4NL + 4DL	NW	Y	U + R	84	1	.012	.012	.017	.005
4IL	NW	Y	U + R	40	8	.200	.063	.122	.022
All	NW	MA + O	U + R	93	2	.022	.015	.017	.005

Final Model

$$\underset{\sim}{X} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 2 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 2 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 2 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \quad \underset{\sim}{b} = \begin{bmatrix} .017 \\ .028 \\ .049 \end{bmatrix}$$

Statistical Tests	D.F.	X^2_W
$b_2 \hat{=} 0$	1	43.92
$b_3 \hat{=} 0$	1	4.50
Model	2	57.51
Residual	15	6.67
Total	17	64.18

include the location effect and fit the illustrated two parameter model. The model fit, residual $\chi^2_W = 1.07$ with df = 2 and the location effect accounts for 82 percent of the variation. The predicted usage rate for rural drivers (9.7 percent) is almost three times that for urban drivers (3.4 percent).

White males on four-lane undivided highways, four-lane divided highways, and four-lane Interstates.

For white male drivers of foreign cars with N.C. registration on each of the three road types (four-lane undivided highways, four-lane divided highways, and four-lane Interstates), total variation was nonsignificant and therefore differences among the respective subpopulations do not appear to exist. Thus, our models for each of these highway types involve only an overall mean parameter yielding corresponding predicted usage rates of 10.0 percent, 10.5 percent, and 24.7 percent.

White females.

Total variation among subpopulations of white female drivers is nonsignificant, $\chi^2_N = 2.34$, df = 5, and so differences do not appear to exist. Therefore our final model involves only an overall mean parameter yielding a predicted usage rate of 8.9 percent.

Blacks.

Since total variation is nonsignificant ($\chi^2_W = .90$, df = 1) for black drivers, differences between the two road type groups (two-

TABLE C13

Road Type	Race-Sex	Age	Location	Number Observed	Number Users	Observed Usage Rate	Est. s.e.	Predicted Usage Rate	Est. s.e.
2L	W-M	Y	U	80	2	.025	.017	.034	.015
2L	W-M	Y	R	101	10	.099	.030	.097	.024
2L	W-M	MA + O	U	67	4	.060	.029	.034	.015
2L	W-M	MA + O	R	54	5	.093	.039	.097	.024

Final Model

$$\underset{\sim}{X} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad \underset{\sim}{b} = \begin{bmatrix} .034 \\ .062 \end{bmatrix}$$

Statistical Tests	D.F.	X^2_W
Model	1	4.95
Residual	2	1.07
Total	3	6.02

TABLE C14

Road Type	Race-Sex	Age	Location	Number Observed	Number Users	Observed Usage Rate	Est. s.e.	Predicted Usage Rate	Est. s.e.
4NL	W-M	Y	U	103	9	.087	.028	.100	.017
4NL	W-M	Y	R	74	8	.108	.036	.100	.017
4NL	W-M	MA + O	U	82	8	.098	.033	.100	.017
4NL	W-M	MA + O	R	48	6	.125	.048	.100	.017

Final Model

$$\underset{\sim}{X} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \underset{\sim}{b} = [.100]$$

Statistical Tests	D.F.	χ^2
Total variation	3	.54

TABLE C15

Road Type	Race-Sex	Age	Location	Number Observed	Number Users	Observed Usage Rate	Est. s.e.	Predicted Usage Rate	Est. s.e.
4DL	W-M	Y	U	70	5	.071	.031	.105	.016
4DL	W-M	Y	R	134	13	.097	.026	.105	.016
4DL	W-M	MA + O	U	51	9	.176	.053	.105	.016
4DL	W-M	MA + O	R	101	13	.129	.033	.105	.016

Final Model

$$\underset{\sim}{X} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \underset{\sim}{b} = [.105]$$

Statistical Tests	D.F.	χ^2
Total Variation	3	3.59

TABLE C16

Road Type	Race-Sex	Age	Location	Number Observed	Number Users	Observed Usage Rate	Est. s.e.	Predicted Usage Rate	Est. s.e.
4IL	W-M	Y	U + R	56	12	.214	.055	.247	.045
4IL	W-M	MA + O	U + R	35	11	.314	.078	.247	.045

Final Model

$$\underset{\sim}{X} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \underset{\sim}{b} = [.247]$$

Statistical Tests	D.F	χ^2
Total Variation	1	1.09

TABLE C17

Road Type	Race-Sex	Age	Location	Number Observed	Number Users	Observed Usage Rate	Est. s.e.	Predicted Usage Rate	Est. s.e.
2L	W-F	ALL	U	43	4	.093	.044	.088	.013
2L	W-F	ALL	R	90	9	.100	.032	.088	.013
4NL	W-F	ALL	U	91	6	.066	.026	.088	.013
4NL	W-F	ALL	R	61	6	.098	.038	.088	.013
4DL + 4IL	W-F	ALL	U	62	4	.065	.031	.088	.013
4DL + 4IL	W-F	ALL	R	138	16	.116	.027	.088	.013

Final Model

$$\underset{\sim}{X} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \underset{\sim}{b} = [.088]$$

Statistical Tests	D.F.	χ^2_W
Total Variation	5	2.57

lane and four-lane undivided highways vs four-lane divided and four-lane Interstates) do not appear to exist. A model involving only an overall mean was fitted to the data yielding a predicted usage rate of 3.3 percent.

Overall model for drivers of foreign cars with N.C. registration.

The individual final models for white males on two-lane highways, four-lane undivided highways, four-lane divided highways and four-lane Interstates were combined with those for white females and black drivers to form a preliminary overall model for drivers of foreign cars with N.C. registration. Statistical tests were performed to detect any structural differences or similarities among the six groups. As a result of these tests, we constructed the final model, which fits with residual χ^2_W = 10.17, df = 19 and has significant effect parameters. This model divides these drivers into three clusters with corresponding predicted usage rates of 3.4 percent, 9.6 percent and 24.7 percent.

Drivers of Foreign Cars with Out-of-State Registration.

Since total variation is significant with χ^2_W = 31.47, df = 10, it can be said that differences exist among the usage rates for foreign car drivers with out-of-state registration. We fit a model with an overall mean or common initial value, and two incremental effects. The model fits, with a residual χ^2_W = 3.00, df = 8, and accounts for

TABLE C18

Road Type	Race-Sex	Age	Location	Number Observed	Number Users	Observed Usage Rate	Est. s.e.	Predicted Usage Rate	Est. s.e.
2L + 4NL	B	All	U + R	79	2	.025	.018	.033	.016
4DL + 4IL	B	All	U + R	48	3	.063	.035	.033	.016

Final Model

$$\underset{\sim}{X} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \underset{\sim}{b} = [.033]$$

Statistical Tests	D.F.	X^2_W
Total Variation	1	.90

TABLE C19

Final Clusters of All Drivers of Foreign Cars with an N.C. License

Cluster	Road Type	Race-Sex	Age	Location	Predicted Usage Rate
I	2L All 4NL + 4DL	W-M W-F W-M	All All All	R U + R U + R	3.4%
II	2L All	W-M B	All All	U U + R	9.6%
III	4IL	W-M	All	U + R	24.7%

TABLE C20

Road Type	Race-Sex	Age	Location	Number Observed	Number Users	Observed Usage Rate	Est. s.e.	Predicted Usage Rate	Est. s.e.
2L	W-M	Y	U	80	2	.025	.017	.034	.011
2L	W-M	Y	R	101	10	.099	.030	.096	.008
2L	W-M	MA + O	U	67	4	.060	.029	.034	.011
2L	W-M	MA + O	R	54	5	.093	.039	.096	.008
2L	W-F	All	U	43	4	.093	.044	.096	.008
2L	W-F	All	R	90	9	.100	.032	.096	.008
4NL	W-M	Y	U	103	9	.087	.028	.096	.008
4NL	W-M	Y	R	74	8	.108	.036	.096	.008
4NL	W-M	MA + O	U	82	8	.098	.033	.096	.008
4NL	W-M	MA + O	R	48	6	.125	.048	.096	.008
4NL	W-F	All	U	91	6	.066	.026	.096	.008
4NL	W-F	All	R	61	6	.098	.038	.096	.008
4DL	W-M	Y	U	70	5	.071	.031	.096	.008
4DL	W-M	Y	R	134	13	.097	.026	.096	.008
4DL	W-M	MA + O	U	51	9	.176	.053	.096	.008
4DL	W-M	MA + O	R	101	13	.129	.033	.096	.008
4DL + 4IL	W-F	All	U	62	4	.065	.031	.096	.008
4DL + 4IL	W-F	All	R	138	16	.116	.027	.096	.008
4IL	W-M	Y	U + R	56	12	.214	.055	.247	.045
4IL	W-M	MA + O	U + R	35	11	.314	.078	.247	.045
2L + 4NL	B	All	U + R	79	2	.025	.018	.034	.011
4DL + 4IL	B	All	U + R	48	3	.063	.035	.034	.011

Preliminary Model

$$\underset{\sim}{X} = \begin{bmatrix} 1&0&0&0&0&0&0\\ 1&1&0&0&0&0&0\\ 1&0&0&0&0&0&0\\ 1&1&0&0&0&0&0\\ 0&0&1&0&0&0&0\\ 0&0&1&0&0&0&0\\ 0&0&0&1&0&0&0\\ 0&0&0&1&0&0&0\\ 0&0&0&1&0&0&0\\ 0&0&0&1&0&0&0\\ 0&0&1&0&0&0&0\\ 0&0&1&0&0&0&0\\ 0&0&0&0&1&0&0\\ 0&0&0&0&1&0&0\\ 0&0&0&0&1&0&0\\ 0&0&0&0&1&0&0\\ 0&0&1&0&0&0&0\\ 0&0&1&0&0&0&0\\ 0&0&0&0&0&1&0\\ 0&0&0&0&0&1&0\\ 0&0&0&0&0&0&1\\ 0&0&0&0&0&0&1 \end{bmatrix}, \quad \underset{\sim}{b} = \begin{bmatrix} .034\\ .062\\ .088\\ .100\\ .105\\ .247\\ .033 \end{bmatrix}$$

Statistical Tests	D.F.	χ^2_W
$b_1 + b_2 - b_3 \hat{=} 0,\ b_1 + b_2 - b_4 \hat{=} 0,\ b_1 + b_2 - b_5 \hat{=} 0,\ b_1 - b_7 \hat{=} 0$	4	.69
$b_2 \hat{=} 0$	1	4.95
Model	6	36.50
Residual	15	9.75

Final Model

$$\underset{\sim}{X} = \begin{bmatrix} 1&0&0\\ 1&1&0\\ 1&0&0\\ 1&1&0\\ 1&1&0\\ 1&1&0\\ 1&1&0\\ 1&1&0\\ 1&1&0\\ 1&1&0\\ 1&1&0\\ 1&1&0\\ 1&1&0\\ 1&1&0\\ 1&1&0\\ 1&1&0\\ 1&1&0\\ 1&1&0\\ 1&1&1\\ 1&1&1\\ 1&0&0\\ 1&0&0 \end{bmatrix}, \quad \underset{\sim}{b} = \begin{bmatrix} .034\\ .062\\ .151 \end{bmatrix}$$

Statistical Tests	D.F.	χ^2_W
$b_2 \hat{=} 0$	1	21.21
$b_3 \hat{=} 0$	1	10.93
Model	2	35.81
Residual	19	10.44

90 percent of the total variation. It divides the subpopulations into three clusters with predicted usage rates of 8.5 percent, 16.0 percent, and 34.9 percent.

<u>All Drivers of U.S. Cars</u>

Models for all U.S. car drivers and all foreign car drivers were arrived at by a first principles approach rather than through a direct concatenation of the two final models from their subpopulations defined by type of registration (N.C. and out-of-state). That is, overall models for the two car make groups were determined by a process similar to the one described in the previous four sections for earlier models where we attempted to define clusters that would account for most of the variation. Moreover, in the formulation of these models, special attention was applied to certain similarities between the clusters formed under the final models for North Carolina and out-of-state car registration drivers. A final model was defined for all U.S. car drivers which divides such drivers into six clusters with respective rates of .8 percent, 2.5 percent, 4.1 percent, 5.4 percent, 7.1 percent, and 12.2 percent. The model fits, with a residual χ^2_W = 14.88, df = 41 and accounts for 94 percent of the total variation.

U.S. car drivers with North Carolina registration comprise the definitive subpopulation, i.e., the one of most interest and for which there is the most data (12,044 drivers), and the clusters defined by its final model remained intact under the overall model. Clusters I,

TABLE C21

Final Clusters of All Drivers of Foreign Cars with Other Licenses

Cluster	Road Type	Race-Sex	Age	Location	Predicted Usage Rate
I	2L + 4NL + 4DL	W-F	All	U + R	8.5%
II	2L + 4NL	W-M	Y	U + R	
			MA + O	U	
	4DL	W-M	All	U	
	2L + 4NL + 4DL	B	All	U + R	16.0%
III	2L + 4NL	W-M	MA + O	R	
	4DL	W-M	All	R	
	4IL	All	All	U + R	34.9%

TABLE C22

Road Type	Race-Sex	Age	Location	Number Observed	Number Users	Observed Usage Rate	Est. s.e.	Predicted Usage Rate	Est. s.e.
2L + 4NL	W-M	Y	U	35	6	.171	.064	.160	.031
2L + 4NL	W-M	Y	R	29	4	.138	.064	.160	.031
2L + 4NL	W-M	MA + O	U	18	4	.222	.098	.160	.031
2L + 4NL	W-M	MA + O	R	27	9	.333	.091	.349	.036
4DL	W-M	All	U	28	4	.143	.066	.160	.031
4DL	W-M	Y	R	42	17	.405	.076	.349	.036
4DL	W-M	MA + O	R	29	8	.276	.083	.349	.036
2L + 4NL + 4DL	W-F	All	U + R	59	5	.085	.036	.085	.036
2L + 4NL + 4DL	B	All	U + R	25	4	.160	.073	.160	.031
4IL	All	Y	U + R	47	15	.319	.068	.349	.036
4IL	All	MA + O	U + R	25	11	.440	.099	.349	.036

Final Model

$$\underset{\sim}{X} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad \underset{\sim}{b} = \begin{bmatrix} .085 \\ .075 \\ .189 \end{bmatrix}$$

Source of Variation	D.F.	X^2_W
$b_2 \hat{=} 0$	1	2.46
$b_3 \hat{=} 0$	1	15.47
Model	2	28.47
Residual	8	3.00
Total	10	31.47

II, III, IV, and V, with corresponding usage rates of .8 percent, 2.4 percent, 4.1 percent. 5.3 percent, and 6.9 percent became all or part of clusters I, II, III, IV, and V in the overall U.S. model with smoothed rates of .8 percent, 2.5 percent, 4.1 percent, 5.4 percent, and 7.1 percent.

For U.S. car drivers with out-of-state registration, clusters II and III, with usage rates of 4.5 percent and 7.3 percent became part of clusters III and V with rates of 4.1 percent and 7.1 percent in the overall model, while cluster IV became cluster VI with the same rate of 12.2 percent. However, cluster I was redefined. It contained all the older drivers, who accounted for 73 percent of the cluster's total population. For these drivers, sample sizes were too small to detect any significant location effect. However, in combination with N.C. registration drivers, statistical tests indicate the urban-rural effect was significant for older drivers on two-lane and four-lane undivided highways. Furthermore, with the two registration types combined, cluster I, which had a usage rate of 1.7 percent was determined not to be a significant grouping of drivers and was broken up in such a way that older white urban drivers on two-lane and four-lane undivided highways were absorbed into cluster I of the overall model with a usage rate of 0.8 percent, and the rest, including certain subpopulations of black drivers, were assigned to cluster II with a usage rate of 2.5 percent.

TABLE C23

Final Clusters of U.S. Car Drivers

Cluster	Road Type	Race-Sex	Age	Location	Registration	Predicted Rate
I	2L + 4NL	W-M	O	U	N.C.	
	2L + 4NL	W-F	MA + O	U + R	N.C.	
	4DL	W-F	MA + O	U + R	N.C.	
	All	B	MA + O	U + R	N.C.	
	2L + 4NL	W	O	U	O	.8%
II	2L + 4NL	W-M	Y + MA	U	N.C.	
	2L + 4NL	W-M	O	R	N.C.	
	2L + 4NL	W-F	Y	U + R	N.C.	
	4DL	W-M	Y + MA	U	N.C.	
	4DL	W-M	O	U + R	N.C.	
	4DL	W-F	Y	U + R	N.C.	
	All	B	Y	U + R	N.C.	
	2L + 4NL	W	O	R	O	
	4DL	W	O	U + R	O	
	4IL	W	O	U + R	O	
	2L + 4NL + 4DL	B	Y	U + R	O	
	All	B	MA + O	U + R	O	2.5%
III	2L + 4NL	W-M	Y + MA	R	N.C.	
	4DL	W-M	MA	R	N.C.	
	2L + 4NL	W	Y	U + R	O	
	2L + 4NL + 4DL	W	MA	U	O	4.1%
IV	4IL	W-F	MA + O	U + R	N.C.	5.4%
V	4DL	W-M	Y	R	N.C.	
	4IL	W-M	Y+MA+O	U + R	N.C.	
	4IL	W-F	Y	U + R	N.C.	
	2L + 4NL	W	MA	R	O	
	4DL	W	Y	U + R	O	
	4DL	W	MA	R	O	
	4IL	W	MA	U + R	O	7.1%
VI	4IL	W	Y	U + R	O	
	4IL	B	Y	U + R	O	12.2%

TABLE C24

Registration	Road Type	Race-Sex	Age	Location	Number Observed	Number Users	Observed Usage Rate	Est. s.e.	Predicted Usage Rate	Est. s.e.
N.C.	2L + 4NL	W-M	Y	U	612	15	.025	.006	.025	.001
N.C.	2L + 4NL	W-M	Y	R	645	24	.037	.007	.041	.003
N.C.	2L + 4NL	W-M	MA	U	1213	32	.026	.005	.025	.001
N.C.	2L + 4NL	W-M	MA	R	1101	49	.045	.006	.041	.003
N.C.	2L + 4NL	W-M	O	U	247	3	.012	.007	.008	.002
N.C.	2L + 4NL	W-M	O	R	255	7	.027	.010	.025	.001
N.C.	2L + 4NL	W-F	Y	U	431	11	.026	.008	.025	.001
N.C.	2L + 4NL	W-F	Y	R	360	8	.022	.008	.025	.001
N.C.	2L + 4NL	W-F	MA + O	U	635	8	.013	.004	.008	.002
N.C.	2L + 4NL	W-F	MA + O	R	540	6	.011	.005	.008	.002
N.C.	4DL	W-M	Y	U	280	9	.032	.011	.025	.001
N.C.	4DL	W-M	Y	R	692	44	.064	.009	.071	.004
N.C.	4DL	W-M	MA	U	370	10	.027	.008	.025	.001
N.C.	4DL	W-M	MA	R	1090	45	.041	.006	.041	.003
N.C.	4DL	W-M	O	U	81	4	.049	.024	.025	.001
N.C.	4DL	W-M	O	R	220	5	.023	.010	.025	.001
N.C.	4DL	W-F	Y	U	156	2	.013	.009	.025	.001
N.C.	4DL	W-F	Y	R	331	9	.027	.009	.025	.001
N.C.	4DL	W-F	MA + O	U	209	1	.005	.005	.008	.002
N.C.	4DL	W-F	MA + O	R	457	3	.007	.004	.008	.002
N.C.	4IL	W-M	Y	U + R	249	20	.080	.017	.071	.004
N.C.	4IL	W-M	MA	U + R	491	36	.073	.012	.071	.004
N.C.	4IL	W-M	O	U + R	103	7	.068	.025	.071	.004
N.C.	4IL	W-F	Y	U + R	116	8	.069	.024	.071	.004
N.C.	4IL	W-F	MA + O	U + R	134	7	.052	.019	.054	.005
N.C.	All	B	Y	U + R	639	12	.019	.005	.025	.001
N.C.	All	B	MA + O	U + R	387	2	.005	.004	.008	.002
O	2L + 4NL	W	Y	U	179	9	.050	.016	.041	.003
O	2L + 4NL	W	Y	R	174	10	.057	.018	.041	.003
O	2L + 4NL	W	MA	U	280	11	.039	.012	.041	.003
O	2L + 4NL	W	MA	R	343	26	.076	.014	.071	.004
O	2L + 4NL	W	O	U	78	0	.000	.009	.008	.002
O	2L + 4NL	W	O	R	91	2	.022	.015	.025	.001
O	4DL	W	Y	U	47	5	.106	.045	.071	.004
O	4DL	W	Y	R	248	18	.073	.016	.071	.004
O	4DL	W	MA	U	94	4	.043	.021	.041	.003
O	4DL	W	MA	R	495	35	.071	.011	.071	.004
O	4DL	W	O	U	35	1	.029	.028	.025	.001
O	4DL	W	O	R	157	5	.032	.014	.025	.001
O	4IL	W	Y	U + R	171	19	.111	.024	.122	.022
O	4IL	W	MA	U + R	306	21	.069	.014	.071	.004
O	4IL	W	O	U + R	122	3	.025	.014	.025	.001
O	2L + 4NL + 4DL	B	Y	U + R	84	1	.012	.012	.025	.001
O	4IL	B	Y	U + R	40	8	.200	.063	.122	.022
O	All	B	MA + O	U + R	93	2	.022	.015	.025	.001

Final Model

$$
\underset{\sim}{X} = \begin{bmatrix}
1 & 1 & 0 & 0 \\
1 & 2 & 0 & 0 \\
1 & 1 & 0 & 0 \\
1 & 2 & 0 & 0 \\
1 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 \\
1 & 2 & 1 & 0 \\
1 & 1 & 0 & 0 \\
1 & 2 & 0 & 0 \\
1 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
1 & 2 & 1 & 0 \\
1 & 2 & 1 & 0 \\
1 & 2 & 1 & 0 \\
1 & 2 & 1 & 0 \\
1 & 1 & 1 & 0 \\
1 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 \\
1 & 2 & 0 & 0 \\
1 & 2 & 0 & 0 \\
1 & 2 & 0 & 0 \\
1 & 2 & 1 & 0 \\
1 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 \\
1 & 2 & 1 & 0 \\
1 & 2 & 1 & 0 \\
1 & 2 & 0 & 0 \\
1 & 2 & 1 & 0 \\
1 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 \\
1 & 2 & 1 & 1 \\
1 & 2 & 1 & 0 \\
1 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 \\
1 & 2 & 1 & 1 \\
1 & 1 & 0 & 0
\end{bmatrix}, \quad \underset{\sim}{b} = \begin{bmatrix} .008 \\ .017 \\ .029 \\ .052 \end{bmatrix}
$$

Statistical Tests	D.F.	χ^2_W
$b_2 \hat{=} 0$	1	92.73
$b_3 \hat{=} 0$	1	31.94
$b_4 \hat{=} 0$	1	5.09
Model	3	233.91
Residual	41	14.88

All Drivers of Foreign Cars

Using the first principles approach discussed in the section on all U.S. car drivers, we arrived at a final model for all foreign car drivers which divides the subpopulations into five clusters with respective usage rates of 3.3 percent, 9.2 percent, 15.2 percent, 27.3 percent, and 33.2 percent. The model fits, with a residual $\chi^2_W = 11.08$, df = 30, and accounts for 90 percent of the total variation. The clusters formed under the final models for both North Carolina registration drivers of foreign cars and out-of-state registration drivers of foreign cars remained intact under the new overall model. In the case of North Carolina registration foreign car drivers, clusters I and III with respective usage rates of 3.4 percent and 24.7 percent became clusters I and IV in the overall model, with corresponding smoothed rates of 3.2 percent and 27.3 percent. For out-of-state registration foreign car drivers, clusters II and III with respective rates of 16.0 percent and 34.9 percent became clusters III and V with rates adjusted downward to 15.1 percent and 33.2 percent. Cluster II of North Carolina registration foreign car drivers, with a usage rate of 9.6 percent was combined with cluster I of out-of-state registration foreign car drivers with a usage rate of 8.5 percent to become cluster II in the overall model with a usage rate of 9.2 percent.

TABLE C25

Final Clusters of All Foreign Car Drivers

Cluster	Road Type	Race-Sex	Age	Location	Residence	Predicted Usage Rate
I	2L	W-M	Y + MA + O	U	N.C.	
	All	B	All	U + R	N.C.	3.2%
II	2L	W-M	Y + MA + O	R	N.C.	
	All	W-F	Y + MA + O	U + R	N.C.	
	4NL	W-M	Y + MA + O	U + R	N.C.	
	4DL	W-M	Y	U + R	N.C.	
	2L + 4NL + 4DL	W-F	All	U + R	O	9.2%
III	4DL	W-M	MA + O	U + R	N.C.	
	2L + 4NL	W-M	Y	U + R	O	
	2L + 4NL	W-M	MA + O	U	O	
	4DL	W-M	All	U	O	
	2L + 4NL + 4DL	B	All	U + R	O	15.1%
IV	4IL	W-M	All	U + R	N.C.	27.3%
V	2L + 4NL	W-M	MA + O	R	O	
	4DL	W-M	Y + MA + O	R	O	
	4IL	All	All	U + R	O	33.2%

TABLE C26

Registration	Road Type	Race-Sex	Age	Location	Number Observed	Number Users	Observed Usage Rate	Est. s.e.	Predicted Usage Rate	Est. s.e.
N.C.	2L	W-M	Y	U	80	2	.025	.017	.032	.010
N.C.	2L	W-M	Y	R	101	10	.099	.030	.092	.007
N.C.	2L	W-M	MA + O	U	67	4	.060	.029	.032	.010
N.C.	2L	W-M	MA + O	R	54	5	.093	.039	.092	.007
N.C.	2L	W-F	All	U	43	4	.093	.044	.092	.007
N.C.	2L	W-F	All	R	90	9	.100	.032	.092	.007
N.C.	4NL	W-M	Y	U	103	9	.087	.028	.092	.007
N.C.	4NL	W-M	Y	R	74	8	.108	.036	.092	.007
N.C.	4NL	W-M	MA + O	U	82	8	.098	.033	.092	.007
N.C.	4NL	W-M	MA + O	R	48	6	.125	.048	.092	.007
N.C.	4NL	W-F	All	U	91	6	.066	.026	.092	.007
N.C.	4NL	W-F	All	R	61	6	.098	.038	.092	.007
N.C.	4DL	W-M	Y	U	70	5	.071	.031	.092	.007
N.C.	4DL	W-M	Y	R	134	13	.097	.026	.092	.007
N.C.	4DL	W-M	MA + O	U	51	9	.176	.053	.151	.014
N.C.	4DL	W-M	MA + O	R	101	13	.129	.033	.151	.014
N.C.	4DL + 4IL	W-F	All	U	62	4	.065	.031	.092	.007
N.C.	4DL + 4IL	W-F	All	R	138	16	.116	.027	.092	.007
N.C.	4IL	W-M	Y	U + R	56	12	.214	.055	.273	.029
N.C.	4IL	W-M	MA + O	U + R	35	11	.314	.078	.273	.029
N.C.	2L + 4NL	B	All	U + R	79	2	.025	.018	.032	.010
N.C.	4DL + 4IL	B	All	U + R	48	3	.063	.035	.032	.010
O	2L + 4NL	W-M	Y	U	35	6	.171	.064	.151	.014
O	2L + 4NL	W-M	Y	R	29	4	.138	.064	.151	.014
O	2L + 4NL	W-M	MA + O	U	18	4	.222	.098	.151	.014
O	2L + 4NL	W-M	MA + O	R	27	9	.333	.091	.332	.029
O	4DL	W-M	All	U	28	4	.143	.066	.151	.014
O	4DL	W-M	Y	R	42	17	.405	.076	.332	.029
O	4DL	W-M	MA + O	R	29	8	.276	.083	.332	.029
O	2L + 4NL + 4DL	W-F	All	U + R	59	5	.085	.036	.092	.007
O	2L + 4NL + 4DL	B	All	U + R	25	4	.160	.073	.151	.014
O	4IL	All	Y	U + R	47	15	.319	.068	.332	.029
O	4IL	All	MA + O	U + R	25	11	.440	.099	.332	.029

<u>Final Model</u>

$$\underset{\sim}{X} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 2 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 2 & 0 \\ 1 & 2 & 0 \\ 1 & 2 & 1 \\ 1 & 2 & 0 \\ 1 & 2 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 2 & 1 \\ 1 & 2 & 1 \end{bmatrix}, \quad \underset{\sim}{b} = \begin{bmatrix} .032 \\ .059 \\ .181 \end{bmatrix}$$

Statistical Tests	D.F.	X^2_W
$b_2 \hat{=} 0$	1	33.71
$b_3 \hat{=} 0$	1	35.79
Model	2	97.68
Residual	30	11.32

All Drivers

The final model for all drivers was also determined from the first principles approach, where we attempted to define clusters of driver subpopulations that accounted for most of the total variation. However, before deriving this model we combined certain subpopulations within U.S. and foreign car drivers. A reduced table (less rows) with larger cell sizes makes it more sensitive to detect across cluster structural similarities and differences. Moreover, it is easier to handle in terms of the computer program we were using. All the combinations were performed within clusters and on subpopulations that were previously determined to be homogeneous.

The final model determined for all drivers contains an overall mean and five incremental effects and divides the subpopulations into eight clusters with respective usage rates of .8 percent, 2.5 percent, 4.1 percent, 5.4 percent, 7.1 percent, 9.7 percent, 15.8 percent, 27.2 percent, and 33.3 percent. The model fits with a residual χ^2_W = 17.13, df = 49, and accounts for 96 percent of the total variation. Clusters formed from the final models of U.S. and foreign car drivers remained intact. For U.S. car drivers, clusters I, II, IV, and V with co-responding usage rates of .8 percent, 2.5 percent, 5.4 percent, and 7.1 percent became, respectively, clusters I, II, IV, and V in the overall model for all drivers with the same rates. For foreign car drivers, clusters III, IV, and V with corresponding rates of 15.2 percent, 27.3 percent, and 33.2 percent became clusters VII, VIII, and IX in the overall model with smoothed rates of 15.8 percent, 27.2 percent, and 33.3 percent. Cluster III for U.S. car drivers and cluster I for foreign car drivers, with respective rates of 4.1 percent and 3.3 percent were combined to form cluster III with a rate of 4.1 percent. Also, cluster VI for U.S. car drivers and cluster II for foreign car drivers with respective rates of 12.2 percent and 9.2 percent were combined to form cluster VI, with an adjusted rate of 9.7 percent.

TABLE C27

Final Clusters of All Drivers

Cluster	Car Make	Registration	Road Type	Race-Sex	Age	Location	Predicted Usage rate
I	U.S.	N.C. + O	2L + 4NL	W	O	U	
	U.S.	N.C.	2L + 4NL + 4DL	W-F	MA + O	U + R	
	U.S.	N.C.	All	B	MA + O	U + R	.8%
II	U.S.	N.C.	2L + 4NL + 4DL	W-M	Y + MA	U	
	U.S.	N.C. + O	2L + 4NL	W-M	O	R	
	U.S.	N.C. + O	4DL	W-M	O	U + R	
	U.S.	N.C.	2L + 4NL + 4DL	W-F	Y	U + R	
	U.S.	N.C.	All	B	Y	U + R	
	U.S.	O	All	W-F	O	R	
	U.S.	O	4DL + 4IL	W	O	U + R	
	U.S.	O	2L + 4NL + 4DL	B	Y	U + R	
	U.S.	O	All	B	MA + O	U + R	2.5%
III	U.S.	N.C.	2L + 4NL	W-M	Y + MA	R	
	U.S.	N.C.	4DL	W-M	MA	R	
	U.S.	O	2L + 4NL	W	Y	U + R	
	U.S.	O	2L + 4NL + 4DL	W	MA	U	
	F	N.C.	2L	W-M	All	U	
	F	N.C.	All	B	All	U + R	4.1%
IV	U.S.	N.C.	4IL	W-F	MA + O	U + R	5.4%

V	U.S.	N.C.	4DL	W-M	Y	R	
	U.S.	N.C.	4IL	W-M	All	U + R	
	U.S.	N.C.	4IL	W-F	Y	U + R	
	U.S.	0	2L + 4NL + 4DL	W	MA	R	
	U.S.	0	4DL	W	Y	U + R	
	U.S.	0	4IL	W	MA	U + R	7.1%
VI	U.S.	0	4IL	All	Y	U + R	
	F	N.C.	2L	W-M	All	R	
	F	N.C.	2L	W-F	All	U + R	
	F	N.C.	4NL	W	All	U + R	
	F	N.C.	4DL	W-M	Y	U + R	
	F	N.C.	4DL + 4IL	W-F	All	U + R	
	F	0	2L + 4NL + 4DL	W-F	All	U + R	9.7%
VII	F	N.C.	4DL	W-M	MA + O	U + R	
	F	0	2L + 4NL + 4DL	W-M	All	U	
	F	0	2L + 4NL	W-M	Y	R	
	F	0	2L + 4NL + 4DL	B	All	U + R	15.8%
VIII	F	N.C.	4IL	W-M	All	U + R	27.2%
IX	F	0	4DL	W-M	Y	R	
	F	0	2L + 4NL + 4DL	W-M	MA + O	R	
	F	0	4IL	All	MA + O	U + R	33.3%

TABLE C28

Vehicle Type	Registration	Road Type	Race-Sex	Age	Location	Number Observed	Number Users	Observed Usage Rate	Est. s.e.	Predicted Usage rate	Est. s.e.
U.S.	N.C.	2L + 4NL	W-M	Y	U	612	15	.025	.006	.025	.001
U.S.	N.C.	2L + 4NL	W-M	Y	R	645	24	.037	.007	.041	.003
U.S.	N.C.	2L + 4NL	W-M	MA	U	1213	32	.026	.005	.025	.001
U.S.	N.C.	2L + 4NL	W-M	MA	R	1101	49	.045	.006	.041	.003
U.S.	N.C.	2L + 4NL	W-M	O	U	247	3	.012	.007	.008	.002
U.S.	N.C.	2L + 4NL	W-M	O	R	255	7	.027	.010	.025	.001
U.S.	N.C.	4DL	W-M	Y	U	280	9	.032	.011	.025	.001
U.S.	N.C.	4DL	W-M	Y	R	692	44	.064	.009	.071	.004
U.S.	N.C.	4DL	W-M	MA	U	370	10	.027	.008	.025	.001
U.S.	N.C.	4DL	W-M	MA	R	1090	45	.041	.006	.041	.003
U.S.	N.C.	4DL	W-M	O	U	81	4	.049	.024	.025	.001
U.S.	N.C.	4DL	W-M	O	R	220	5	.023	.010	.025	.001
U.S.	N.C.	2L + 4NL + 4DL	W-F	Y	U + R	1278	30	.023	.004	.025	.001
U.S.	N.C.	2L + 4NL + 4DL	W-F	MA + O	U + R	1841	18	.010	.002	.008	.002
U.S.	N.C.	4IL	W-M	All	U + R	843	63	.075	.009	.071	.004
U.S.	N.C.	4IL	W-F	Y	U + R	116	8	.069	.024	.071	.004
U.S.	N.C.	4IL	W-F	MA + O	U + R	134	7	.052	.019	.054	.005
U.S.	N.C.	All	B	Y	U + R	639	12	.019	.005	.025	.001
U.S.	N.C.	All	B	MA + O	U + R	387	2	.005	.004	.008	.002
U.S.	Other	2L + 4NL	W	Y	U	179	9	.050	.016	.041	.003
U.S.	Other	2L + 4NL	W	Y	R	174	10	.057	.018	.041	.003
U.S.	Other	2L + 4NL	W	MA	U	280	11	.039	.012	.041	.003
U.S.	Other	2L + 4NL	W	MA	R	343	26	.076	.014	.071	.004
U.S.	Other	2L + 4NL	W	O	U	78	0	.000	.009	.008	.002
U.S.	Other	2L + 4NL	W	O	R	91	2	.022	.015	.025	.001
U.S.	Other	4DL	W	Y	U	47	5	.106	.045	.071	.004
U.S.	Other	4DL	W	Y	R	248	18	.073	.016	.071	.004
U.S.	Other	4DL	W	MA	U	94	4	.043	.021	.041	.003
U.S.	Other	4DL	W	MA	R	495	35	.071	.012	.071	.004

U.S.	Other	4DL	W	O	U	35	1	.029	.028	.025	.001
U.S.	Other	4DL	W	O	R	157	5	.032	.014	.025	.001
U.S.	Other	4IL	W	Y	U + R	171	19	.111	.024	.097	.008
U.S.	Other	4IL	W	MA	U + R	306	21	.069	.014	.071	.004
U.S.	Other	4IL	W	O	U + R	122	3	.025	.014	.025	.001
U.S.	Other	2L + 4NL + 4DL	B	Y	U + R	84	1	.012	.012	.025	.001
U.S.	Other	4IL	B	Y	U + R	40	8	.200	.063	.097	.008
U.S.	Other	All	B	MA + O	U + R	93	2	.022	.015	.025	.001
Foreign	N.C.	2L	W-M	All	U	147	6	.041	.016	.041	.003
Foreign	N.C.	2L	W-M	All	R	155	15	.097	.024	.097	.008
Foreign	N.C.	2L	W-F	All	U + R	133	13	.098	.026	.097	.008
Foreign	N.C.	4NL	W	All	U + R	459	43	.094	.014	.097	.008
Foreign	N.C.	4DL	W-M	Y	U + R	204	18	.088	.020	.097	.008
Foreign	N.C.	4DL	W-M	MA + O	U + R	152	22	.145	.029	.158	.020
Foreign	N.C.	4DL + 4IL	W-F	All	U + R	200	20	.100	.021	.097	.008
Foreign	N.C.	4IL	W-M	Y	U + R	56	12	.214	.055	.272	.031
Foreign	N.C.	4IL	W-M	MA + O	U + R	35	11	.314	.078	.272	.031
Foreign	N.C.	All	B	All	U + R	127	5	.039	.017	.041	.003
Foreign	Other	2L + 4NL + 4DL	W-M	All	U	81	14	.173	.042	.158	.020
Foreign	Other	2L + 4NL	W-M	Y	R	29	4	.138	.064	.158	.020
Foreign	Other	4DL	W-M	Y	R	42	17	.405	.076	.333	.029
Foreign	Other	2L + 4NL + 4DL	W-M	MA + O	R	56	17	.304	.061	.333	.029
Foreign	Other	2L + 4NL + 4DL	W-F	All	U + R	59	5	.085	.036	.097	.008
Foreign	Other	2L + 4NL + 4DL	B	All	U + R	25	4	.160	.073	.158	.020
Foreign	Other	4IL	All	Y	U + R	47	15	.319	.068	.333	.029
Foreign	Other	4IL	All	MA + O	U + R	25	11	.440	.099	.333	.029

Final Model

$$
\underset{\sim}{X} = \begin{bmatrix}
1 & 1 & 0 & 0 & 0 & 0 \\
1 & 2 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 \\
1 & 2 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 \\
1 & 2 & 1 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 \\
1 & 2 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 \\
1 & 2 & 1 & 0 & 0 & 0 \\
1 & 2 & 1 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 \\
1 & 2 & 0 & 0 & 0 & 0 \\
1 & 2 & 0 & 0 & 0 & 0 \\
1 & 2 & 0 & 0 & 0 & 0 \\
1 & 2 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 \\
1 & 2 & 1 & 0 & 0 & 0 \\
1 & 2 & 1 & 0 & 0 & 0 \\
1 & 2 & 0 & 0 & 0 & 0 \\
1 & 2 & 1 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 \\
1 & 2 & 1 & 1 & 0 & 0 \\
1 & 2 & 1 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 \\
1 & 2 & 1 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 \\
1 & 2 & 0 & 0 & 0 & 0 \\
1 & 2 & 1 & 1 & 0 & 0 \\
1 & 2 & 1 & 1 & 0 & 0 \\
1 & 2 & 1 & 1 & 0 & 0 \\
1 & 2 & 1 & 1 & 0 & 0 \\
1 & 2 & 1 & 1 & 1 & 0 \\
1 & 2 & 1 & 1 & 0 & 0 \\
1 & 2 & 1 & 1 & 0 & 1 \\
1 & 2 & 1 & 1 & 0 & 1 \\
1 & 2 & 0 & 0 & 0 & 0 \\
1 & 2 & 1 & 1 & 1 & 0 \\
1 & 2 & 1 & 1 & 1 & 0 \\
1 & 2 & 1 & 1 & 1 & 1 \\
1 & 2 & 1 & 1 & 1 & 1 \\
1 & 2 & 1 & 1 & 0 & 0 \\
1 & 2 & 1 & 1 & 1 & 0 \\
1 & 2 & 1 & 1 & 1 & 1 \\
1 & 2 & 1 & 1 & 1 & 1
\end{bmatrix}, \quad \underset{\sim}{b} = \begin{bmatrix} .008 \\ .016 \\ .029 \\ .026 \\ .061 \\ .175 \end{bmatrix}
$$

Statistical Tests	D.F.	χ^2_W
$b_2 \hat{=} 0$	1	89.55
$b_3 \hat{=} 0$	1	32.45
$b_4 \hat{=} 0$	1	8.41
$b_5 \hat{=} 0$	1	8.46
$b_6 \hat{=} 0$	1	32.12
Model	5	432.43
Residual	49	17.13

APPENDIX D

Shoulder Harness Usage:
Three-Point Versus Four-Point Systems

This appendix is concerned with a comparison of shoulder harness usage for U.S. model vehicles with 3-point systems vs. those with 4-point systems. The vehicles included in this part of the study were restricted to those with N.C. registration for which a vehicle identification number could ultimately be traced for the purpose of determining whether the seat-belt-shoulder-harness system was 3-point or 4-point. For the overall samples which were derived from this process, the following results were obtained

	Number Observed	Number Users	Observed Usage Rate
3-point system	1550	49	3.2%
4-point system	5993	142	2.4%

which, on the basis of χ^2_p = 3.13 with df = 1, suggest that a weakly (α = .10) significant relationship exists between shoulder harness usage and system type.

On the other hand, we have already noted (see Appendix A) that several variables pertaining to the driver's demographic characteristics and the road's environmental characteristics have significant effects on shoulder harness usage. In particular, age, race and sex of the driver as well as nature of the location and road type at the observation site were important in this respect. For this reason, a more appropriate approach for undertaking the comparison of 3-point systems with 4-point systems is that used by Campbell (1970) and Koch and Reinfurt

(1974) for studying the degree of injury severity in certain specific makes and models of accident involved automobiles. The basic statistic is analogous to those previously described by Cochran (1954) and Mantel and Haenszel (1959).

Let $n_{hi_1i_2i_3i_4i_5j}$ denote the observed frequency of the h-th usage category for drivers of the i_1-th age, i_2-th sex, i_3-th race at the i_4-th location on the i_5-th road type with vehicles equipped with the j-th system type. The subscripts h, i_1, i_2, i_3, i_4, i_5, and j have levels defined as follows:

h = 1 User
2 Non-user

i_1 = 1 Younger age
2 Mature age
3 Older age

i_2 = 1 Male
2 Female

i_3 = 1 White
2 Black

i_4 = 1 Urban
2 Rural

i_5 = 1 2-Lane
2 4-Lane
3 4-Div. Lane
4 Interstate

j = 1 3-point
2 4-point

Hence, there are 96 combinations of driver and site characteristics which are used in this analysis.

Define $r_{hi_1i_2i_3i_4i_5} = \sum_{j=1}^{2} n_{hi_1i_2i_3i_4i_5j}$ to represent the frequency of the h-th usage category for observations involving the (i_1, i_2, i_3, i_4, i_5)-th driver x site situation in the combined 3-point and 4-point reference population. Also, define $r_{\cdot i_1 i_2 i_3 i_4 i_5 \cdot} = \sum_{h=1}^{2} r_{h i_1 i_2 i_3 i_4 i_5 \cdot}$. This quantity represents the total number of observations corresponding to $(i_1, i_2, i_3, i_4, i_5)$.

The quantity

$$P_{h i_1 i_2 i_3 i_4 i_5} = (r_{h i_1 i_2 i_3 i_4 i_5 \cdot} / r_{\cdot i_1 i_2 i_3 i_4 i_5 \cdot})$$

represents the conditional probability of the h-th usage category for the $(i_1, i_2, i_3, i_4, i_5)$-th situation within the combined reference population. If usage of 3-point systems and 4-point systems are the same for all of the 96 situations corresponding to the respective $(i_1, i_2, i_3, i_4, i_5)$, then the expected frequency for usage of the j-th system in the $(i_1, i_2, i_3, i_4, i_5)$-th situation is given by

$$m_{h i_1 i_2 i_3 i_4 i_5 j} = n_{\cdot i_1 i_2 i_3 i_4 i_5 j} P_{h i_1 i_2 i_3 i_4 i_5}$$

where $n_{\cdot i_1 i_2 i_3 i_4 i_5 j} = \sum_{h=1}^{2} n_{h i_1 i_2 i_3 i_4 i_5 j}$ represents the frequency of the $(i_1, i_2, i_3, i_4, i_5)$-th situation for vehicles equipped with the j-th system. In order to obtain an overall comparison of usage in 3-point and 4-point systems, both the observed frequencies $n_{h i_1 i_2 i_3 i_4 i_5 j}$ and the expected frequencies $m_{h i_1 i_2 i_3 i_4 i_5 j}$ are summed over the totality of situations $(i_1, i_2, i_3, i_4, i_5)$ to determine

$$n_{hj} = \sum_{i_1=1}^{3} \sum_{i_2=1}^{2} \sum_{i_3=1}^{2} \sum_{i_4=1}^{2} \sum_{i_5=1}^{4} n_{hi_1i_2i_3i_4i_5j}$$

$$m_{hj} = \sum_{i_1=1}^{3} \sum_{i_2=1}^{2} \sum_{i_3=1}^{2} \sum_{i_4=1}^{2} \sum_{i_5=1}^{4} m_{hi_1i_2i_3i_4i_5j}$$

If $n_{hj} \cong m_{hj}$, then it can be said that there is essentially no difference between the two systems with respect to shoulder harness usage.

For a specific usage category h and system category j, the statistical significance of the difference between n_{hj} and m_{hj} can be evaluated by means of a χ^2-test where

$$\chi^2_T = (n_{hj} - m_{hj})^2 / v_{hj}$$

and v_{hj} is an estimate of the variance of $(n_{hj} - m_{hj})$. An appropriate choice for v_{hj} is

$$v_{hj} = \sum_{i_1=1}^{3} \sum_{i_2=1}^{2} \sum_{i_3=1}^{2} \sum_{i_4=1}^{2} \sum_{i_5=1}^{4} \mathrm{var}\{n_{hi_1i_2i_3i_4i_5j}\}$$

where

$$\mathrm{var}\{n_{hi_1i_2i_3i_4i_5j}\} =$$

$$m_{hi_1i_2i_3i_4i_5j}\{(1 - \frac{r_{hi_1i_2i_3i_4i_5\cdot}}{r_{\cdot i_1i_2i_3i_4i_5\cdot}})\ (1 - \frac{n_{\cdot i_1i_2i_3i_4i_5j} - 1}{r_{\cdot i_1i_2i_3i_4i_5\cdot} - 1})\}.$$

If the hypothesis that $n_{hj} \cong m_{hj}$ is true, then χ^2_T has approximately the chi square distribution with df = 1. This may be used as the basis for determining significant differences between n_{hj} and m_{hj}.

The pertinent data for comparing 3-point and 4-point systems in terms of this test statistic are given in Table D1 where the rows correspond to the 96 combinations of driver and site characteristics. The columns are arranged for the particular comparison of 4-point systems (i.e., j = 2) to the reference population with respect to the usage (i.e., H = 1) response category. Thus, the columns of this table correspond to the number observed in the combined reference population ($r_{\cdot \underset{\sim}{i} \cdot} = r_{\cdot i_1 i_2 i_3 i_4 i_5 \cdot}$), the number of users in the combined reference population ($r_{1 \underset{\sim}{i} \cdot} = r_{1 i_1 i_2 i_3 i_4 i_5 \cdot}$), the number observed with the 4-point system ($n_{\cdot \underset{\sim}{i} 2} = n_{\cdot i_1 i_2 i_3 i_4 i_5 2}$), the number of users with the 4-point system ($n_{1 \underset{\sim}{i} 2} = n_{1 i_1 i_2 i_3 i_4 i_5 2}$), the expected number of users with the 4-point system from the reference population ($m_{1 \underset{\sim}{i} 2} = m_{1 i_1 i_2 i_3 i_4 i_5 2}$), and the estimate of variance for the observed minus the expected difference ($v_{1 \underset{\sim}{i} 2} = v_{1 i_1 i_2 i_3 i_4 i_5 2}$).

The resulting value of χ^2_T with respect to the totality of the 96 driver x site situations is 1.63, which is nonsignificant compared to a chi square distribution with one degree of freedom. This suggests that there is no significant difference between the 3-point and 4-point systems regarding shoulder harness usage across 96 subpopulations. The χ^2 statistic was also computed within subgroups defined by levels of age, sex, race, road type, and road location in their own right. Significance (α = .05) was determined only within urban road locations where χ^2 = 4.58. In view of other findings we feel that this result is probably a peculiarity of the particular sample considered in this report as opposed to reflecting a real difference between the 3-point and 4-point systems.

TABLE D1

Observed Data For Comparison of 3-Point and 4-Point Systems for Respective Driver x Site Situations $\underset{\sim}{i} = (i_1, i_2, i_3, i_4, i_5)$

Race-Sex	Age	Location	Road	Number Observed in Combined Reference Population $r_{\cdot\underset{\sim}{i}\cdot}$	Number Users in Combined Reference Population $r_{1\underset{\sim}{i}\cdot}$	Number Observed with 4-point System $n_{\cdot\underset{\sim}{i}2}$	Number Users with 4-point System $n_{1\underset{\sim}{i}2}$	Expected Number Users with 4-point System from Reference Population $m_{1\underset{\sim}{i}2}$	Estimate of Variance for Observed minus Expected Difference $v_{1\underset{\sim}{i}2}$
W-M	Y	U	2L	191	3	138	2.0	2.16	0.5920
W-M	Y	U	4NL	211	2	158	0.0	1.49	0.3726
W-M	Y	U	4DL	181	7	144	5.0	5.56	1.0944
W-M	Y	U	4IL	12	1	6	0.0	0.50	0.2292
W-M	Y	R	2L	241	9	190	7.0	7.09	1.4454
W-M	Y	R	4NL	154	5	116	3.0	3.76	0.8992
W-M	Y	R	4DL	401	18	311	15.0	13.96	2.9925
W-M	Y	R	4IL	145	13	113	11.0	10.13	2.0354
W-M	MA	U	2L	357	5	266	4.0	3.72	0.9363
W-M	MA	U	4NL	447	14	354	10.0	11.08	2.2345
W-M	MA	U	4DL	233	7	189	5.0	5.67	1.0400
W-M	MA	U	4IL	24	5	17	2.0	3.54	0.8178
W-M	MA	R	2L	349	15	273	11.0	11.73	2.4453
W-M	MA	R	4NL	247	5	201	5.0	4.06	0.7424
W-M	MA	R	4DL	593	15	437	8.0	11.05	2.8344
W-M	MA	R	4IL	272	18	204	14.0	13.50	3.1517
W-M	0	U	2L	105	0	85	0.0	0.0	0.0
W-M	0	U	4NL	69	1	58	1.0	0.84	0.1321
W-M	0	U	4DL	56	3	46	2.0	2.46	0.4165
W-M	0	U	4IL	3	1	3	1.0	1.00	0.0
W-M	0	R	2L	90	0	77	0.0	0.0	0.0
W-M	0	R	4NL	59	2	51	2.0	1.72	0.2265
W-M	0	R	4DL	128	2	100	2.0	1.56	0.3365
W-M	0	R	4IL	63	2	51	2.0	1.61	0.2986
W-F	Y	U	2L	147	1	116	0.0	0.78	0.1653
W-F	Y	U	4NL	138	5	108	5.0	3.91	0.8198
W-F	Y	U	4DL	99	0	84	0.0	0.0	0.0
W-F	Y	U	4IL	5	0	3	0.0	0.0	0.0
W-F	Y	R	2L	140	4	118	4.0	3.37	0.5147
W-F	Y	R	4NL	83	0	66	0.0	0.0	0.0
W-F	Y	R	4DL	191	4	154	2.0	3.22	0.6117
W-F	Y	R	4IL	72	5	58	4.0	4.02	0.7288
W-F	MA	U	2L	198	1	156	1.0	0.78	0.1663
W-F	MA	U	4NL	190	1	161	1.0	0.84	0.1287
W-F	MA	U	4DL	132	1	110	0.0	0.83	0.1378
W-F	MA	U	4IL	2	0	2	0.0	0.0	0.0
W-F	MA	R	2L	198	0	167	0.0	0.0	0.0
W-F	MA	R	4NL	115	1	95	1.0	0.82	0.1424
W-F	MA	R	4DL	244	2	198	1.0	1.62	0.3035
W-F	MA	R	4IL	78	4	59	4.0	3.02	0.6992
W-F	0	U	2L	45	2	42	2.0	1.86	0.1189

W-F	0	U	4NL	23	0	21	0.0	0.0	0.0
W-F	0	U	4DL	24	0	20	0.0	0.0	0.0
W-F	0	U	4IL	2	0	2	0.0	0.0	0.0
W-F	0	R	2L	27	0	22	0.0	0.0	0.0
W-F	0	R	4NL	22	0	16	0.0	0.0	0.0
W-F	0	R	4DL	61	0	50	0.0	0.0	0.0
W-F	0	R	4IL	19	1	16	1.0	0.84	0.1260
B-M	Y	U	2L	0	0	0	0.0	0.0	0.0
B-M	Y	U	4NL	3	1	3	1.0	0.87	0.1091
B-M	Y	U	4DL	1	0	0	0.0	0.0	0.0
B-M	Y	U	4IL	136	0	0	0.0	0.0	0.0
B-M	Y	R	2L	66	0	57	0.0	0.0	0.0
B-M	Y	R	4NL	71	1	62	0.0	0.66	0.2074
B-M	Y	R	4DL	20	1	18	1.0	0.71	0.1973
B-M	Y	R	4IL	3	0	3	0.0	0.0	0.0
B-M	MA	U	2L	39	0	31	0.0	0.0	0.0
B-M	MA	U	4NL	15	1	10	1.0	0.86	0.1113
B-M	MA	U	4DL	39	0	28	0.0	0.0	0.0
B-M	MA	U	4IL	26	0	23	0.0	0.0	0.0
B-M	MA	R	2L	43	0	37	0.0	0.0	0.0
B-M	MA	R	4NL	38	0	33	0.0	0.0	0.0
B-M	MA	R	4DL	16	1	14	1.0	0.85	0.1215
B-M	MA	R	4IL	0	0	0	0.0	0.0	0.0
B-M	0	U	2L	14	0	10	0.0	0.0	0.0
B-M	0	U	4NL	8	0	7	0.0	0.0	0.0
B-M	0	U	4DL	27	0	23	0.0	0.0	0.0
B-M	0	U	4IL	8	0	8	0.0	0.0	0.0
B-M	0	R	2L	7	0	7	0.0	0.0	0.0
B-M	0	R	4NL	2	0	2	0.0	0.0	0.0
B-M	0	R	4DL	0	0	0	0.0	0.0	0.0
B-M	0	R	4IL	0	0	0	0.0	0.0	0.0
B-F	Y	U	2L	2	1	1	0.0	0.83	0.1360
B-F	Y	U	4NL	0	0	0	0.0	0.0	0.0
B-F	Y	U	4DL	3	0	3	0.0	0.0	0.0
B-F	Y	U	4IL	0	0	0	0.0	0.0	0.0
B-F	Y	R	2L	48	0	40	0.0	0.0	0.0
B-F	Y	R	4NL	26	0	22	0.0	0.0	0.0
B-F	Y	R	4DL	15	0	11	0.0	0.0	0.0
B-F	Y	R	4IL	0	0	0	0.0	0.0	0.0
B-F	MA	U	2L	6	0	5	0.0	0.0	0.0
B-F	MA	U	4NL	6	0	6	0.0	0.0	0.0
B-F	MA	U	4DL	18	0	15	0.0	0.0	0.0
B-F	MA	U	4IL	4	0	4	0.0	0.0	0.0
B-F	MA	R	2L	25	0	24	0.0	0.0	0.0
B-F	MA	R	4NL	14	0	14	0.0	0.0	0.0
B-F	MA	R	4DL	5	0	5	0.0	0.0	0.0
B-F	MA	R	4IL	0	0	0	0.0	0.0	0.0
B-F	0	U	2L	10	0	9	0.0	0.0	0.0
B-F	0	U	4NL	5	0	4	0.0	0.0	0.0
B-F	0	U	4DL	17	0	16	0.0	0.0	0.0
B-F	0	U	4IL	4	0	4	0.0	0.0	0.0
B-F	0	R	2L	1	0	0	0.0	0.0	0.0
B-F	0	R	4NL	2	0	2	0.0	0.0	0.0
B-F	0	R	4DL	0	0	0	0.0	0.0	0.0
B-F	0	R	4IL	0	136	0	0.0	0.0	0.0

APPENDIX E

Interaction of Right-Front Passenger and Driver as Related to Shoulder Harness Usage

This appendix is concerned with the relationship of right front passenger usage of shoulder harnesses to driver usage. For this purpose, our sample contained 5174 eligible vehicles with N.C. registration (whose age, race, sex known) in which the shoulder harness usage status of both the driver and right front passengers was observed. Of the 174 such vehicles in which the driver was wearing a shoulder harness, 89 (or 51.1 percent) of these carried a right front passenger who was wearing a shoulder harness. Moreover, this high usage rate was relatively unaffected by both driver variables like age, race, and sex as well as site variables like location and road type. On the other hand, for the 5000 eligible observed vehicles in which the driver was not wearing a shoulder harness, only 25 (or 0.5 percent) of these carried a right front passenger who was wearing a shoulder harness. In this context, vehicle type as well as road type were of some importance as reflected by the following table

Vehicle Type	Road Type	Number Observed	Number Users	Observed Usage Rate
U.S.	2L + 4NL + 4DL	3974	9	0.2%
U.S.	4LI	513	5	1.0%
U.S.	All	513	11	2.1%

for which $\chi^2_p = 36.18$, df = 2. Thus, significant differences exist among these rates. However, further examination of this table indicated that the U.S.-4LI sample and the Foreign sample were not significantly different ($\chi^2_p = 2.29$, df = 1). As a result, this analysis suggests that right front passenger usage for vehicles in which the driver is not wearing a shoulder harness can be characterized by two clusters. One of these corresponds to U.S. model vehicles on non-Interstate road type for which the usage rate was 0.2 percent; the other was the combined group of U.S. model vehicles on Interstate roads together with foreign model vehicles for which the usage rate was 1.6 percent.

In view of these findings, it would seem that highway safety programs which tended to encourage increased driver usage of shoulder harness systems would correspondingly induce increased right front passenger usage.

APPENDIX F

Shoulder Harness Utilization Data Collection Form

SHOULDER HARNESS UTILIZATION DATA COLLECTION FORM

(1) ROAD TYPE ☐
1. 2-Lane
2. 4-Lane
3. 4-Lane Divided
4. Interstate

(2-3) COUNTY ☐☐

(4) LOCATION ☐
1. urban
2. rural

(5) POSITION ☐
1. stationary
2. enroute

(6-7) TIME ☐☐

(8) PAVEMENT CONDITION ☐
1. dry
2. wet

(9) DAY OF WEEK ☐
1. sun.-7. sat.

(10-13) DATE ☐☐☐☐
mo. day

(14-15) POSTED SPEED LIMIT ☐☐

RECORDER'S NAME ______________________

DRIVER			R.F. PASSENGER		INSTATE LICENSE						
RACE & AGE 1-wm 2-wf 3-nwm 4-nwf 9-unk (16)	AGE 1-y 2-m 3-o 9-unk (17)	HARNESS IN USE? 9-unk (18)	PRESENT? 9-unk (19)	HARNESS? 9-unk (20)	9-unk (21-26)	G (27)	F (28)	FOR (29)	O (30)	UNK (31)	DRIV. TNG. (32)