DOT HS-801 833

A STATISTICAL ANALYSIS OF SEAT BELT EFFECTIVENESS IN 1973-75 MODEL CARS INVOLVED IN TOWAWAY CRASHES

Contract No. DOT-HS-5-01255

February 1976

Interim Report

PREPARED FOR:

U.S. DEPARTMENT OF TRANSPORTATION

NATIONAL HIGHWAY TRAFFIC SAFETY ADMINISTRATION

WASHINGTON, D.C. 20590

Document is available to the public through
the National Technical Information Service,
Springfield, Virginia 22161

Technical Report Documentation Page

1. Report No.: DOT HS-801 833

2. Government Accession No.:

3. Recipient's Catalog No.:

4. Title and Subtitle: A STATISTICAL ANALYSIS OF SEAT BELT EFFECTIVENESS IN 1973-75 MODEL CARS INVOLVED IN TOWAWAY CRASHES.

5. Report Date: February 1976

6. Performing Organization Code:

7. Author(s): Donald W. Reinfurt, Claudio Z. Silva, Yosef Hochberg

8. Performing Organization Report No.:

9. Performing Organization Name and Address:
The Highway Safety Research Center
University of North Carolina,
Chapel Hill, N.C. 27514

10. Work Unit No. (TRAIS):

11. Contract or Grant No.: DOT-HS-5-01255

12. Sponsoring Agency Name and Address:
Department of Transportation
National Highway Traffic Safety Administration
400 Seventh Street, S.W.
Washington, D.C. 20590

13. Type of Report and Period Covered:
Interim Report
July 1, 1975 - Oct. 31, 1975

14. Sponsoring Agency Code:

15. Supplementary Notes:

16. Abstract

Using a probability sample of towaway accidents involving 1973-75 model cars and investigated by special teams in 5 different geographic regions, several multivariate analyses are examined which estimate injury rates and corresponding effectiveness relating to various levels of belt usage. The effective sample size is 10,758 occupants (drivers or right front seat passengers) for which belt status and AIS injury information along with age and seat position, car weight, crash configuration and vehicle damage severity is available.

Analyses using log-linear models and weighted least squares procedures are carried out in considerable detail. The limitations and/or advantages of each method are pointed out and the results of various sensitivity investigations presented.

A "Fact Book" of tables of interest from the overall file is included.

17. Key Words:
seat belt effectiveness
standardized injury rates
weighted least squares
log-linear models

18. Distribution Statement:
Document is available to the public through the National Technical Information Service, Springfield, Virginia 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	97	

Form DOT F 1700.7 (8-72)

TECHNICAL SUMMARY

The many safety belt effectiveness studies in the literature agree on the positive benefits of these systems but vary considerable in their estimates of the magnitude of the effectiveness. Reasons for this disagreement include: (1) differing reporting thresholds for the accident data upon which the studies were based; (2) a variety of injury criteria even when using the K, A, B, C, O scale, due to state and regional differences; (3) differential attempts to control for certain variables which interact with belt usage, ranging from no attempt to control for vehicle damage severity, driver age, etc. to somewhat limited atttempts that might control for several variables but possibly not some of their important interactions; and (4) varying investigative biases and inaccuracies in the data (especially police-reported accident data).

An additional problem with available information on safety belt effectiveness is that generally there are no rigorous estimates of the precision of the measures presented. All of these difficulties present serious problems for the policy makers faced with interpreting the results of the studies.

This study, which is part of the Restraint Systems Evaluation Program (RSEP) of the National Highway Traffic Safety Administration, has attempted to overcome these many problems. At present, there is detailed information on some 8,000 towaway accidents involving 1973-75 model passenger cars. A reasonably uniform reporting threshold can be expected since the accidents are towaway accidents. In addition, the limitation of the data to the 1973-75 model years assures that safety features in the vehicles are reasonably comparable as well as guaranteeing uniformity in type of restraint system available to the outboard front seat occupants of these case vehicles. The Level 2 data combines information from police reports with subject and witness interviews, hospital information, and investigation of the vehicle. National representativeness is strived for by utilizing NHTSA-sponsored teams in western New York, Michigan, Miami, San Antonio, and Los Angeles. And, finally, the effects of some of the most important confounding variables are accounted for in the multivariate analyses employed. To the extent possible, the corresponding estimates of the precision of the resulting effectiveness measures are derived.

In order to maximize the likelihood of obtaining detailed information on injured occupants, a stratified probability sample of towaway accidents has been obtained. Occupants of vehicles in which at least one occupant was transported to a treatment facility were sampled at 100 percent. Otherwise, vehicles were selected at a 50 percent rate using the odd/even status of the license plate terminal digit as the randomizing mechanism.

On the basis of the available 10,758 weighted observations for which complete information was available on belt usage and injury level within the various combinations of crash configuration, vehicle damage severity, vehicle weight, and occupant age/seat position, 56.3 percent of the occupants were unrestrained, 16.9 percent wore a lap belt only and 26.8 percent wore both lap and shoulder belts. As the belt systems would generally be 3-point systems, it is not surprising to begin seeing greater usage of both belts than the lap belt alone - even in accidents. Belt usage by vehicle model year is given in Table 1. As expected, lap and shoulder belt usage jumped considerably

Table 1.

Model Year	None	Lap	Lap and Shoulder	Total
1973	3354 (64.1%)[1]	1547 (29.6%)	329 (6.3%)	5230 (48.6%)[2]
1974	2351 (48.3%)	235 (4.8%)	2281 (46.9%)	4867 (45.2%)
1975	353 (53.4%)	33 (5.0%)	275 (41.6%)	661 (6.1%)
Total	6058 (56.3%)	1815 (16.9%)	2885 (26.8%)	10758

[1] Row percent [2] Column percent

with the 1974 model vehicles. At this point in time, an ignition interlock system was introduced which prevented the motorist from starting the car without first buckling up. The percentages for "none" and "lap" then primarily indicate defeat of the system or possibly observational errors.

Also of interest is the restraint usage by injury (AIS) distribution for the sample (see Table 2). If "injured" is defined as "AIS $\geq$ 2", then 9.7 percent of the sample was injured.

Table 2.

AIS Level	None	Lap	Lap and Shoulder	Total
0 Not injured	2511 (41.4%)[1]	879 (48.4%)	1563 (54.2%)	4953 (46.0%)
1 Minor	2801 (46.2%)	782 (43.1%)	1178 (40.8%)	4761 (44.3%)
2 Moderate	560 (9.2%)	124 (6.8%)	109 (3.8%)	793 (7.4%)
3 Severe	106 (1.7%)	20 (1.1%)	19 (0.7%)	145 (1.3%)
4 Serious non-fatal	25 (0.4%)	4 (0.2%)	5 (0.2%)	34 (0.3%)
5 Critical non-fatal	2 (0.0%)	2 (0.1%)	1 (0.0%)	5 (0.0%)
6 Fatal	53 (0.9%)	4 (0.2%)	10 (0.3%)	67 (0.6%)
Total	6058 (56.3%)[2]	1815 (16.9%)	2885 (26.8%)	10758

[1]Column percent [2]Row percent

Crude injury rates derived from Table 2 amount to 12.3 percent, 8.5 percent, and 5.0 percent for the unrestrained (U), lap (L), and lap and shoulder (LS) belt categories, respectively.

Defining belt effectiveness as the percentage reduction in injury as one becomes progressively more restrained, we have overall crude effectiveness measures of 30.9 percent, 59.3 percent and 41.2 percent for U vs. L, U vs. LS, and L vs. LS, respectively. These overall injury rates and effectiveness measures provide unadjusted baseline estimates.

To what extent does belt usage vary according to vehicle size or crash configuration? Certainly, to make a fair comparison between the belt systems, it is important to control for the more important variables which interact with belt usage. Due to limitations on the quality and distribution of the data, it was decided to post-stratify (or control for) crash configuration, vehicle damage severity, vehicle weight, and occupant age/seat position. The distribution of the available sample for each of these variables is given in Table 3.

To appropriately control for these variables in a multivariate analysis procedure for categorical data, several procedures are examined and the results described in considerable detail, since each is not without limitations. As each yields fairly similar results, the limiting assumptions become more tolerable.

The primary analysis procedure implemented used a log-linear model along with weighted least squares for categorical data to collapse the initial strata to an eventual 16 strata from which the required estimates are derived, along with the desired precision estimates. Matrix inversion for deriving the precision estimates necessitates this extensive collapsing. Subsequent work has suggested that certain covariance terms are unimportant. By setting these terms equal to zero, the collapsing requirements will be considerably less stringent.

It should be noted that the collapsing criteria are population conditions under which the standardized injury rates are invariant under strata collapsing of "similar" cells. The decision to collapse certain strata is made under a hypothesis-testing framework and, as such, is affected by random variability. However, additional data will make this less of a potential problem.

Table 3.

Crash Configuration	Percent
1. Head-on with vehicle	6.7
2. Rear-end, striking	14.8
3. Rear-end, stuck	7.1
4. Angle, striking	22.7
5. Angle, struck in left side	14.0
6. Angle, struck in right side	14.2
7. Rollover	1.5
8. Sideswipe	3.0
9. Head-on with fixed object	11.0
10. Skidded sideways into fixed object	5.0
Damage Severity	
1. Minor	45.9
2. Moderate	39.2
3. Moderately severe	10.6
4. Severe	4.3
Vehicle Size	
1. Subcompact (<2700 lbs)	29.9
2. Compact (2700-3599)	25.0
3. Intermediate (3600-4100)	23.6
4. Full-sized (> 4100)	21.5
Occupant Age/Seating Position	
1. Driver, 10-55	68.2
2. Driver, over 55	7.3
3. Front-seat passenger, 10-55	22.2
4. Front-seat passenger, over 55	2.4

Using the combined procedure on the 10,758 observations now available, the adjusted injury rates become 12.0 percent, 9.3 percent, and 5.1 percent for unrestrained, lap belted, and lap and shoulder belted occupants, respectively. Again, with belt effectiveness defined as the percentage decrease in injury (AIS $\geq$ 2) as one becomes progressively more restrained, the overall effectiveness measures become 21.9 percent, 57.4 percent, and 45.5 percent for U vs. L, U vs. LS, and L vs. LS, respectively. The corresponding 95 percent confidence intervals are given by (11.7%, 32.1%), (51.3%, 63.5%), and (35.9%, 55.1%).

It is of interest to note that the primary effect of controlling for crash configuration, damage severity, vehicle weight and occupant age/seat position is to <u>increase</u> the crude injury rate for lap belted occupants from 8.5 percent to 9.3 percent. This results in considerably reduced effectiveness of the lap belt. It should be noted that other multivariate procedures have the same effect on the lap belt rates; similarly, when examining various subsets of the data (e.g., compact vehicles, severe damage, etc.), the adjusted lap belt injury rate consistently exceeds the crude rate. See Table 4 for examples of crude and adjusted injury rates.

It is to be expected that accounting for each of the control variables will differentially affect the overall injury rates and therefore the effectiveness estimates; likewise for various combinations of the control variables. To examine this effect, a detailed sensitivity analysis was carried out. In essence, it is aimed at the question: "What is the effect of controlling for vehicle damage? crash configuration?

Table 4.

	Injury Rates					
	Crude			Adjusted		
	U	L	LS	U	L	LS
Car Weight						
Compact (<3600 lbs)	12.6%	9.8%	5.5%	12.2%	11.1%	5.9%
Full (≥ 3600 lbs)	12.0%	6.9%	4.1%	11.6%	7.2%	4.1%
Damage Severity						
Moderate	9.1%	7.1%	3.5%	9.0%	7.3%	3.4%
Severe	29.2%	18.5%	14.2%	28.6%	21.2%	14.6%

damage by configuration? etc."

Although sensitivity across various subsets of the data was also examined, attention here is focused at the overall effectiveness measures. Each change in percentages cited in Table 5 represents the difference between the crude effectiveness estimates (31.1 percent, 59.5 percent and 41.2 percent for U vs. L, U vs. LS, and L vs. LS, respectively) and the estimates derived with the subset of control variables cited. For example, accounting for crash configuration reduces the unadjusted effectiveness estimate of lap belts by 5.5 percent whereas accounting simultaneously for crash configuration and damage reduces the unadjusted estimate by 10.6 percent.

Generally, it would seem that controlling for vehicle camage is most important with crash configuration next in importance. Clearly, controlling for age/seating position has the least effect on the crude effectiveness estimates.

Table 5.

	U vs. L	U vs. LS	L vs. LS
Crash configuration (C)	-5.5%	-2.7%	+0.7%
Vehicle weight (W)	-0.6%	+1.5%	+2.8%
Vehicle damage (D)	-3.5%	-1.2%	+1.2%
Age/seating position (P)	+0.4%	-0.1%	-0.4%
C x W	-6.0%	-1.5%	+2.7%
C x D	-10.6%	-6.1%	+0.2%
C x P	-0.6%	-0.5%	-0.3%
W x D	-4.2%	-0.6%	+2.5%
W x P	+0.3%	+1.0%	+1.2%
D x P	-4.0%	-1.5%	+1.2%
C x W x D	-6.3%	-2.1%	+2.2%
C x W x P	-0.9%	+0.7%	+1.7%
C x D x P	-6.1%	-3.6%	+0.1%
W x D x P	-4.4%	-0.1%	+3.5%
C x W x D x P	-9.2%	-2.1%	+4.3%

ACKNOWLEDGMENTS

Sincere appreciation is expressed to Dr. Charles J. Kahane for helpful input throughout, to Dr. Gary G. Koch for statistical consultation, to Ms. Debbie Shaver for computer programming, and to Mr. Amitabh Dutt, Ms. Anita Leung, and Mr. Dennis Ryan for assistance with the "Fact Book" preparation. Thanks are also due to our typists Ms. Barbara Crockett and Ms. Peggy James and to Ms. Jane Stutts and Mr. Michael Sadof for coordinating efforts during the final days of report preparation.

TABLE OF CONTENTS

I. INTRODUCTION

The great variety of studies on the subject of safety belt effectiveness have one thing in common - they virtually all agree that these active restraint systems available in all recent model cars sold in the United States are effective in reducing injuries and deaths in motor vehicle collisions. One important aspect in which they disagree is the magnitude of this effectiveness. As alternatives to these systems are being considered, it is most important to know, as nearly as possible, the "true" effectiveness of lap belt and lap and shoulder belt systems, and this implies knowledge about the precision of these estimates derived from a well-controlled field study of accidents.

As described in detail in Kahane, Lee, and Smith (1975), most studies of safety belt effectiveness have been based solely on existing traffic accident records provided by reporting police agencies. This data source generally provides the necessary quantity of data but lacks much of the needed data quality. Clearly, even a state police reporting system cannot be considered nationally representative as, among other things, it would overrepresent rural crashes. Generally such sources do not provide information on certain important variables or else not in sufficient detail to be used in an appropriate analysis. As these variables (e.g., specific crash configuration, vehicle weight) have an important effect on injury severity, information on them must be available in adequate detail. Also, one of the most important variables, injury, is typically described by the K, A, B, C, O scale, which is extraordinarily broad, ill-defined and very subjective, making it most unsatisfactory for analysis purposes.

In addition, there are often numerous investigative biases and inaccuracies in data from traffic accidents as, for example, serious conflict between police-reported and occupant-reported belt usage (see Hochberg and Reinfurt [1974]). Furthermore, reporting thresholds differ so greatly (even within some states) that a given study may be based on a rather non-homogeneous or biased sample of accident reports.

Clearly studies based on in-depth accident investigations avoid most of the above-mentioned pitfalls. However, they would not meet the requirement of being nationally representative nor would they provide a large random sample upon which to base subsequent statistical inference.

This study is based on an intermediate level of data referred to as Level 2 accident data. It combines information provided from police reports with subject and witness interviews, hospital information, and investigation of the vehicle. The data derives from five NHTSA-sponsored teams destributed across the United States (namely, Western New York, Michigan, Miami, San Antonio (Texas), and Los Angeles). Of interest are towaway accidents involving 1973 and newer model passenger cars. As "towaway" is reasonably well-defined, the reporting threshold should be consistent across the five teams. By limiting study to 1973 and newer model cars, there is a guarantee that similar belt systems are available in all cars and that the presence or absence of other safety features is comparable for all cars in the sample.

Working within certain time constraints, it was decided to carry out stratified random sampling in each of the areas in order to obtain an effective sample size of some 15,000 occupants. As only the outboard front seat occupants have both lap and shoulder belts available for use, the study is limited to these two seat positions. With respect to the stratification, all vehicles where hospital treatment was involved were sampled at 100 percent. The remaining vehicles were sampled at 50 percent - exceptions to this scheme are delineated in Kahane et al. (1975). Occupants of these vehicles are included in the sample on the odd/even status of the terminal digit of the license plate. This stratification provides additional precision in the resulting effectiveness estimates through an increased effective sample size and allows detailed information on all of the occupants of special interest (namely, those generally more seriously injured). In addition, that particular subgroup is easier to track down for follow-up interview.

To the extent possible, information was collected for each sampled occupant on some 168 variables. Refer to Appendix A for a complete listing of these variables. It should be noted that there is extensive important information on vehicle damage through the C.D.C. (Collision Deformation Classification), including object contacted and inches of crush, along with detailed injury information through the O.I.C. (Occupant Injury Classification) which utilizes the A.I.S. (Abbreviated Injury Scale).

As can be seen from Appendix A, there is detailed information on virtually all of the crash variables which should affect injury severity, including information on the occupant (e.g., age, sex, height, seating position, belt use), vehicle (e.g., make and model (weight), mileage, extent of damage), and environment and crash situation (e.g., accident type, crash configuration, road type).

In this interim report, a "Fact Book" about towaway accidents of new cars is presented. The tables therein include some 16,000 weighted observations now available (the analysis results are based on 10,758 observations initially available) and show differential belt usage as a function of vehicle size and/or model year, crash configuration, damage severity, seat position and occupant age. Likewise, for unrestrained occupants, the corresponding injury severity distributions are presented. Finally, belt effectiveness estimates for AIS $\geq$ 2 are presented for the overall sample as well as by vehicle size (compact vs full-sized), crash configuration (3 groups) and damage severity (moderate vs severe).

The major effort of this project is to appropriately compare standardized injury rates for various belt groups (unrestrained, lap, lap and shoulder) and the corresponding effectiveness measures for the overall sample as well as for selected subsets, such as occupants of compact cars, various crash configurations, etc. In the process, estimates of the precision of these injury rate and effectiveness measures are obtained wherever possible. The post-stratification variables used as control variables in the analysis are essentially those suggested in Kahane et al. (1975), namely, crash configuration, damage severity, vehicle size, and occupant age and seating position. Obviously any analysis is constrained by the number of factor level combinations and the distribution of the sample across these combinations. For this reason, the post-stratification variables are revised to the variables and corresponding levels presented in Table 1.

Thus, the analysis techniques developed compare injury rates and corresponding effectiveness measures for the three belt usage categories - overall and for selected subsets - controlling for the interacting effects on injury of the variables given in Table 1.

An alternative to using the categorical variable AIS to define an occupant's injury severity is to use the associated direct costs of medical bills, lost wages, etc., due to the injuries sustained.

The final report will present the results of utilizing estimates of direct costs from a variety of sources as the dependnet variable in a multivariate analysis of belt effectiveness.

Table 1. Post-stratification variables.

Crash Configuration

1. Head-on with vehicle
2. Rear end, striking
3. Rear end, struck
4. Angle, striking
5. Angle, struck in left side
6. Angle, struck in right side
7. Rollover
8. Sideswipe
9. Head-on with fixed object
10. Side of vehicle into fixed object

Damage Severity

1. Minor (e.g., 12-FDEW-1, 12-FYEW-1, 12-FLEW-1, 12-FLEE-1, 12-FLEE-2)
2. Moderate (e.g., 12-FDEW-2, 12-FYEW-2, 12-FLEW-2, 12-FLEW-3, 12-FLEE-3, 12-FLEE-4)
3. Moderately severe (e.g., 12-FDEW-3, 12-FYEW-3, 12-FLEW-4, 12-FLEE-5)
4. Severe (e.g., 12-FDEW-4, 12-FYEW-4, 12-FLEW-5, 12-FLEE-6)

Vehicle Size

1. Subcompact (< 2700 lbs.)
2. Compact (2700 - 3599)
3. Intermediate (3600 - 4100)
4. Full-sized (> 4100)

Occupant Age-Seating Position

1. 10 - 55 driver
2. Over 55 driver
3. 10 - 55 front-seat passenger
4. Over 55 front-seat passenger

II. THE DATA; ANALYSIS PLAN

The Data

At present, there is detailed information on 10,758 "occupants" on which the analyses are based. The basic observations have been weighted by the appropriate inverse sampling fractions and are such that there is no missing data for the six variables of interest (belt usage, injury, crash configuration, damage severity, vehicle size, and occupant age/seating position).

As indicated previously, the data consist of detailed occupant information (see Appendix A) for towaway crashes involving 1973-75 model cars. These crashes occurred in 1974 and 1975 in five geographic regions across the United States (namely, Western New York State, Michigan, Miami, San Antonio, and Los Angeles). The data were collected primarily by special NHTSA-sponsored teams of accident specialists combining information from police reports, occupant and witness interviews, hospital or other injury information, and investigation of the vehicle.

For the multivariate analysis, attention is focused on belt usage (3 levels), AIS injury (initially 7 levels), crash configuration (10 levels), damage severity (4 levels), vehicle weight (4 levels) and occupant age/seating position (4 levels). See Table 1 for the description of the levels of the post-stratification variables.

Belt usage determination derives from a combination of information from the police report, occupant interview, investigation of the vehicle, and occasionally location of injuries.

The AIS injury severity for a given occupant is defined to be the maximum severity for the first three injuries (i.e., max (var 135, var 141, var 147); see Appendix A) unless either the police injury code or the treatment mortality code indicates a fatality (i.e., var (129) = 1 or var (130) = 7). In this case, the AIS code is 6 indicating a fatality. Unless otherwise stated in this report, "injured" corresponds to AIS $\geq$ 2 (i.e., at least moderate injury). As more data become available, it will be important to examine, in detail, AIS $\geq$ 3 and AIS = 6. For AIS $\geq$ 2 corresponding to injury, the belt usage by injury level for the weighted sample is given in Table 2.

Table 2. Belt usage by injury level distribution.

Belt Usage \ Injury Level	Not Injured	Injured	Total
None	5312 (87.7%)[1]	746 (12.3%)	6058 (56.3%)[2]
Lap	1661 (91.5%)	154 (8.5%)	1815 (16.9%)
Lap + Shoulder	2741 (95.0%)	144 (5.0%)	2885 (26.8%)
Total	9714 (90.3%)	1044 (9.7%)	10758

[1] Row percent

[2] Column percent

Overall, 9.7% of the sample were injured (AIS $\geq$ 2), 56.3% were unrestrained, 16.9% wore a lap belt only, and 26.8% wore both lap and shoulder belts. As the belt systems would generally be 3-point systems, it is not surprising to begin seeing greater usage of both belts even in accidents.

Note that Table 2 provides crude, unconditional injury rates for each belt category. Thus, for this file of towaway crashes, the overall injury rates are $\hat{R}_1$=.123, $\hat{R}_2$=.085, and $\hat{R}_3$=.05 for the unrestrained (U), lap belt (L), and lap and shoulder (LS) belt categories, respectively. Defining effectiveness as the percentage reduction in injury as one becomes progressively more restrained, we have overall effectiveness measures of

$$\hat{E}_{1,2} = \frac{12.3 - 8.5}{12.3} = .309$$

$$\hat{E}_{1,3} = \frac{12.3 - 5.0}{12.3} = .593$$

$$\hat{E}_{2,3} = \frac{8.5 - 5.0}{8.5} = .412$$

for U vs L, U vs LS, and L vs LS, respectively. These overall injury rates and effectiveness measures provide unconditional baseline estimates for subsequent comparisons.

Crash configuration (see Table 1) was determined using variables 22, 24, 60, 61, and 63 as given in Appendix A. With respect to crash configuration, the category, "other non-collision," was excluded since there were fewer than 20 such accidents and since it did not logically combine with any of the other crash configuration categories. The distribution of crash configuration by injury level and belt usage is given in Table 3.

Vehicle damage has 4 levels (see Table 1) and is defined using variables 1, 22, 24, 60, 61, 63, and 64 as given in Appendix A and hence primarily utilizes the Collision Deformation Classification (CDC). The distribution of damage categories by injury level and belt usage is given in Table 4.

Vehicle weight also has 4 levels (see Table 1) and is defined using the vehicle make/model code (variables 39, 40 in Appendix A). Table 5 shows the distribution of vehicle weight by injury level and belt usage.

Table 3. Crash configuration by injury level and belt usage.

Crash Configuration \ Belt Usage	Not Injured U	Not Injured L	Not Injured LS	Injured U	Injured L	Injured LS	Total
1. Head-on with vehicle	376	91	142	73	19	21	722 (6.7%)
2. Rear-end striking	803	240	459	60	20	15	1597 (14.8%)
3. Rear-end, struck	313	150	260	18	8	10	759 (7.1%)
4. Angle, striking	1273	389	579	140	32	24	2437 (22.7%)
5. Angle, struck in left side	709	253	407	106	13	17	1505 (14.0%)
6. Angle, struck in right side	721	244	422	97	21	26	1531 (14.2%)
7. Rollover	86	6	44	20	3	3	162 (1.5%)
8. Sideswipe	143	49	121	10	2	1	326 (3.0%)
9. Head-on with fixed object	634	149	202	155	24	14	1178 (11.0%)
10. Skidded sideways into fixed object	254	90	105	67	12	13	541 (5.0%)
Total	5312	1661	2741	746	154	144	10758

Table 4. Vehicle damage by injury level and belt usage.

Damage Category \ Belt Usage	Not Injured			Injured			Total
	U	L	LS	U	L	LS	
1. Minor	2538	799	1336	182	46	37	4938 (45.9%)
2. Moderate	2083	681	1055	279	67	49	4214 (39.2%)
3. Moderately Severe	530	127	253	177	27	27	1141 (10.6%)
4. Severe	161	54	97	108	14	31	465 (4.3%)
Total	5312	1661	2741	746	154	144	10758

Table 5. Vehicle weight by injury level and belt usage.

Vehicle Weight \ Belt Usage	Not Injured			Injured			Total
	U	L	LS	U	L	LS	
1. Subcompact (<2700 lb)	1441	486	943	238	49	63	3220 (29.9%)
2. Compact (2700-3599 lb)	1295	385	765	158	46	37	2686 (25.0%)
3. Intermediate (3600-4100 lb)	1331	362	602	184	31	31	2541 (23.6%)
4. Full-sized (>4100 lb)	1245	428	431	166	28	13	2311 (21.5%)
Total	5312	1661	2741	746	154	144	10758

Finally, since no drivers and very few right front seat occupants were under 10 years of age, it was decided to drop that age category and to combine age and seating position into a single variable, occupant age/seating position (see Table 1). The distribution of the weighted sample for occupant age/seating position by injury level and belt usage is given in Table 6.

Table 6. Occupant age/seating position by injury level and belt usage.

Belt Usage / Occupant Age/ Seating Position	Not Injured			Injured			Total
	U	L	LS	U	L	LS	
1. YD Young driver	3471	1192	1974	484	116	95	7332 (68.2%)
2. OD Old driver	342	140	217	63	9	14	785 (7.3%)
3. YP Young passenger	1381	299	490	164	26	28	2388 (22.2%)
4. OP Old passenger	118	30	60	35	3	7	253 (2.4%)
Total	5312	1661	2741	746	154	144	10758

As these 6 variables are used in the multivariate analysis which follows, the detailed sampling distributions are presented.

Notation

Unless otherwise indicated, the following notation is used in this report:

n_{hij} = number of individuals in stratum h with belt usage i and injury level j

where $h = 1,2,\ldots,d$
$i = 1,2,3$
$j = 1,2$

with
$$i = \begin{cases} 1 \text{ if no belt (N)} \\ 2 \text{ if lap belt only (L)} \\ 3 \text{ if lap and shoulder belt (LS)} \end{cases}$$

$$j = \begin{cases} 1 \text{ if injured (AIS} \geq 2) \\ 2 \text{ otherwise} \end{cases}$$

$n_{hi\cdot} = \sum_j n_{hij}$ = number in stratum h with belt usage i

$n_{h\cdot j} = \sum_i n_{hij}$ = number in stratum h with injury j

$n_{\cdot ij} = \sum_h n_{hij}$ = number with belt usage i and injury level j

$n_{h\cdot\cdot} = \sum_{i,j} n_{hij}$ = number in stratum h

$n_{\cdot\cdot\cdot} = \sum_{h,i,j} n_{hij}$ = total number in sample

and

$$\hat{R}_i = \sum_h w_h \, p_{hi1}$$

$$= \sum_h \left(\frac{n_{h\cdot\cdot}}{n_{\cdot\cdot\cdot}}\right) \cdot \left(\frac{n_{hi1}}{n_{hi\cdot}}\right)$$

= estimated overall injury rate for restraint system i, $i = 1,2,3$

$$\hat{E}_{i,i'} = \frac{\hat{R}_i - \hat{R}_{i'}}{\hat{R}_i}$$

= estimated injury-reducing effect of belt system i' compared to belt system i,i < i'

Additional notational conveniences are achieved by the following:

C = crash configuration

S = damage severity

W = vehicle weight

P = occupant age/seating position

I = injured

$\bar{I}$ = not injured

Overall Analysis Plan

The main goal of the analysis is to derive standardized injury rates, effectiveness measures and corresponding standard errors for the various belt usage categories -- both for the overall (weighted) sample and for a variety of subsets of interest (e.g., compact cars, head-on collisions). Chapters III and IV of this report describe the two procedures used thus far to accomplish this goal.

As automobile accidents are extremely complex events involving a large number of factors, any analysis that fails to take these factors into account can be grossly misleading. Also, the variables involved are primarily categorical and thus categorical methods must be utilized. The variety of traditional chi-squares type procedures are inadequate due to the multi-dimensionality of the problem.

In recent years, considerable research has been carried out in this area of analysis of complex contingency tables. Most of the methods use models which express functions of the observed cell frequencies (say, number of unbelted occupants with severe injuries in cell (i,j,k,l,m)) in terms of combinations of a variety of independent variables (say, car weight, crash configuration, age, seating position, damage severity). The log-linear model of Goodman (1970, 1971) expresses the logarithm of the expected value of the function of the cell frequencies in terms of a linear combination of the main effects and interactions of a variety of independent variables. Maximum likelihood methods then provide estimates of the adjusted rates of interest plus tests of significance for the importance of the various main effects and interactions.

Alternatively, the weighted least squares approach of Grizzle, Starmer, Koch (1969) expresses the expected value of either linear or log-linear functions of the observed cell proportions in terms of a linear combination of effects of a variety of independent variables. Weighted least squares methods (directly analogous to those used in the familiar general linear models procedures for continuous variables) provide estimates of the fit of the model and estimates of the functions of interest and their corresponding standard errors.

Neither of these procedures is without its limitations. For example, the log-linear model analysis (Goodman [1970, 1971]; Appendix C) allows a large number of factor-level combinations but fails to provide standard errors of the derived estimates. Weighted least squares procedures (Grizzle, Starmer, and Koch [1969]; Appendix D) provide estimates and their standard errors but, as matrix inversion is required, are limited to relatively few factor-level combinations. This requires extensive collapsing of the post-stratification variables which sacrifices important information from the collapsed strata.

As this has been an evolutionary process and as each method has its limitations, the analysis for each method is carried out in detail along with a sensitivity analysis. The sensitivity analysis shows the effect on the resulting estimates of collapsing on various subsets of the post-stratification variables; i.e., it shows what happens to the crude estimates as additional control variables are progressively introduced into the analysis.

Figure 1 shows the overall flow of the analysis procedures. A description of the log-linear model procedure utilized in Chapter III is illustrated in Figure 2 while Figure 3 details the flow of the weighted least squares and log-linear models used in Chapter IV.

Results of these procedures are examined, the limitations reviewed, and future plans outlined in Chapter VI.

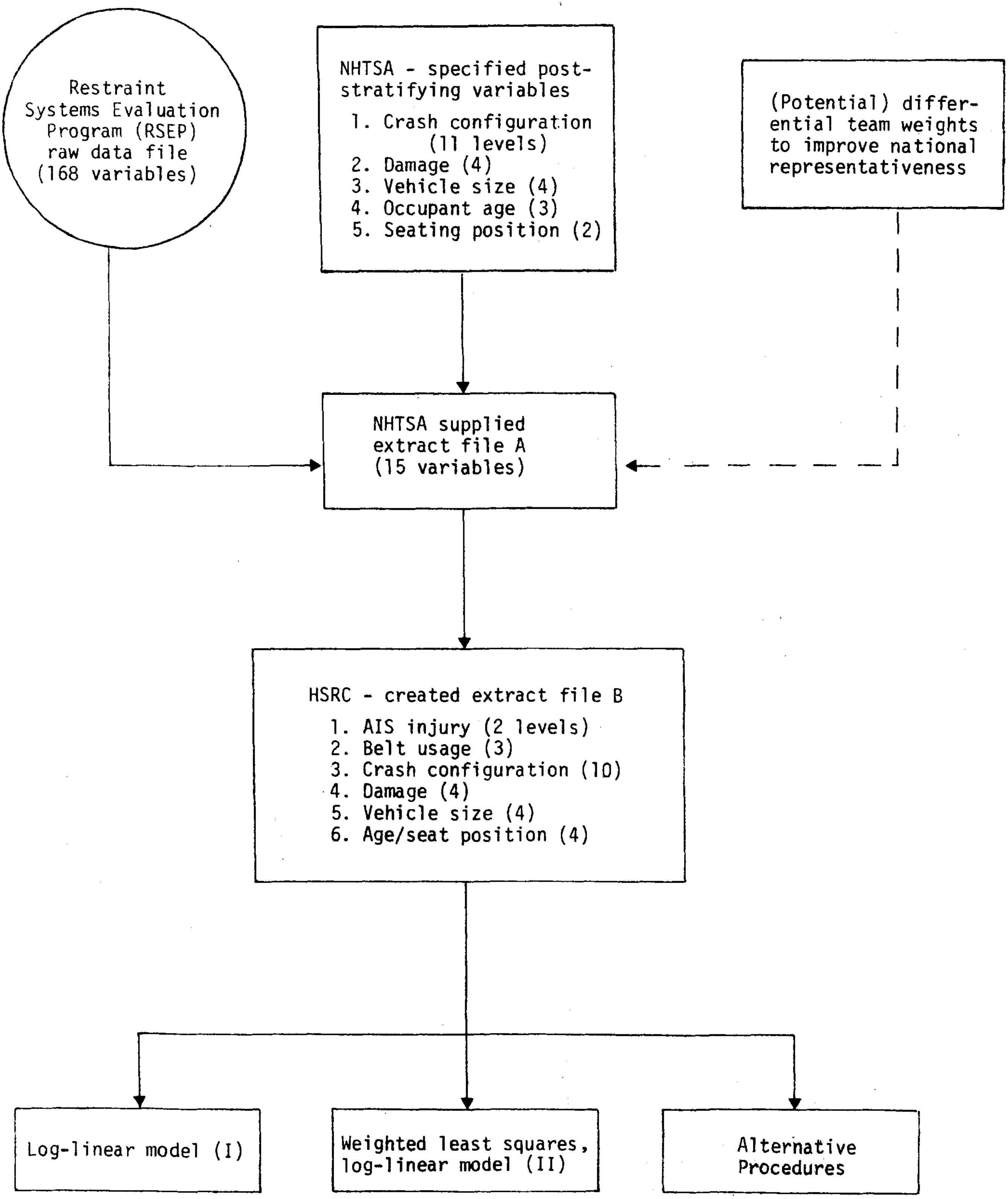

Figure 1. Mathematical Modelling for Determining True Belt Effectiveness.

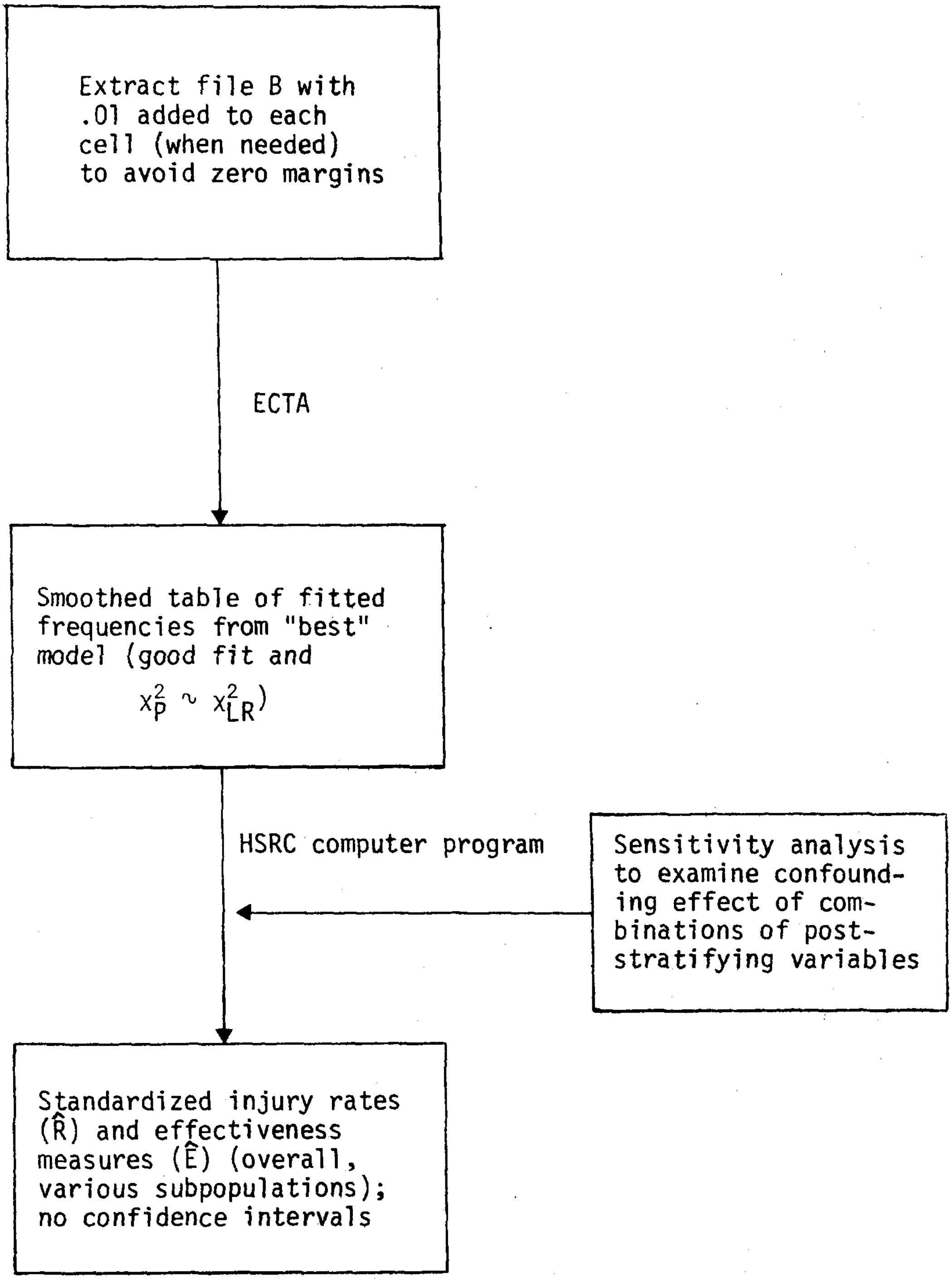

Figure 2. Log-linear model (I)

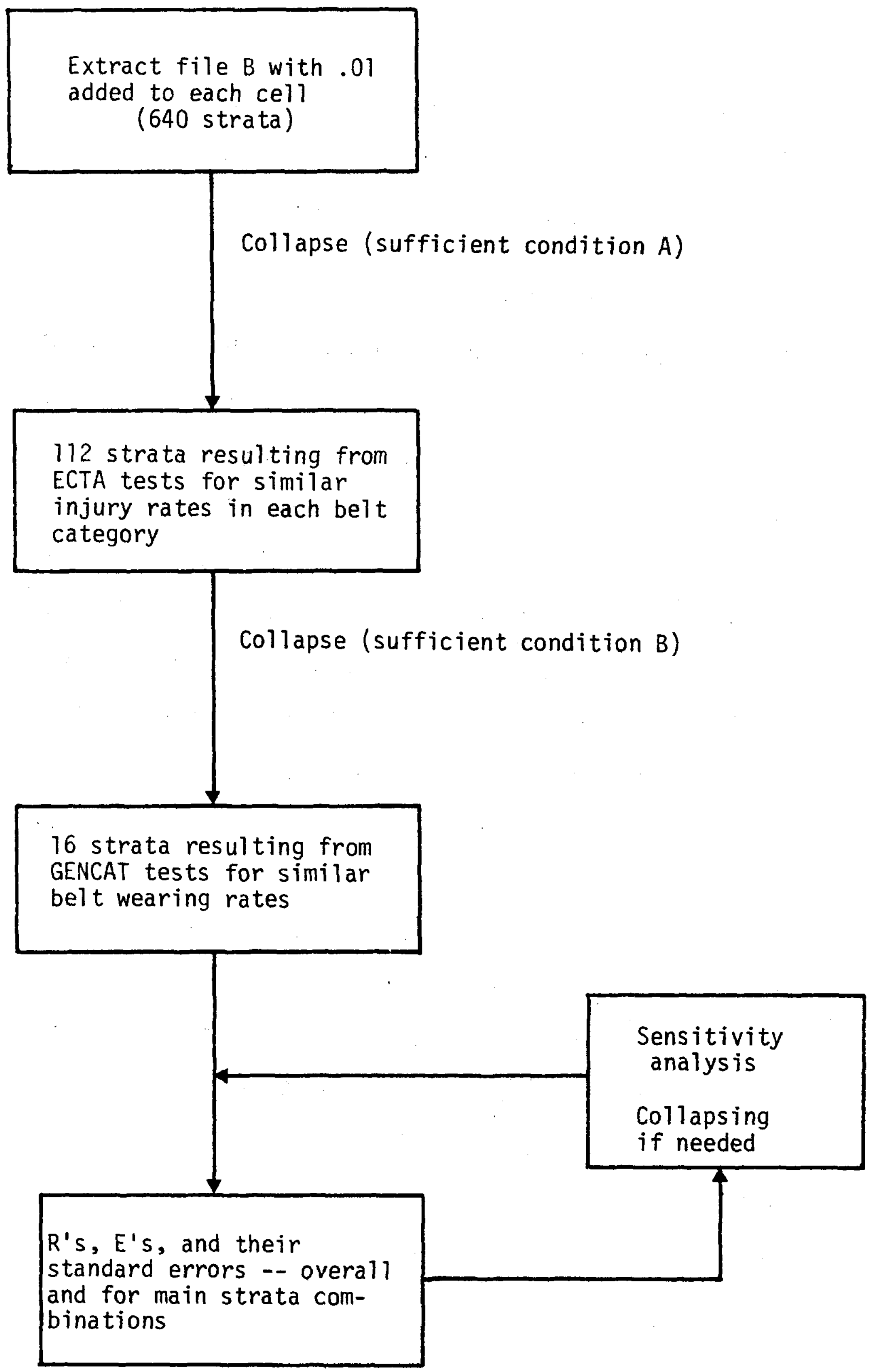

Figure 3. Weighted least squares, log-linear model (II)

III. ESTIMATION OF STANDARDIZED INJURY RATES AND TRUE EFFECTIVENESS MEASURES USING LOG-LINEAR MODELS FOR MULTI-DIMENSIONAL CONTINGENCY TABLES.

Estimates Using Smoothed Frequencies

The original data set contains a large number of zero cells which makes it difficult to carry out direct estimation of the overall injury rates and the true effectiveness measures. These zero cells are due to the large number of strata ($4\times10\times4\times4 = 640$) created by the post-stratification on vehicle size (W), crash configuration (C), damage (S), and occupant age/seat position (P). Recalling that in each of these strata there is a 2×3 (injury $\times$ belt usage) frequency distribution, it is easy to realize that many of these 3840 cells will be empty even with a (weighted) sample of 10,758 cases.

Thus, a first step is to fit a model to the observed injury $\times$ belt usage distributions, and use these fitted or smoothed values as input for an HSRC computer program which then derives the required standardized injury rates and effectiveness measures.

To smooth the data, we use the ECTA (Everyman's Contingency Table Analysis) computer program which is based on an underlying log-linear model of the table cell frequencies. See Appendix C and Goodman (1970, 1971) for details. In this case, the model assumes the form

$$\xi_{\mu,l_1,l_2,l_3,l_4} = \mu + \lambda^{W}_{l_1} + \lambda^{C}_{l_2} + \lambda^{S}_{l_3} + \lambda^{P}_{l_4} + \ldots + \lambda^{WCSP}_{l_1 l_2 l_3 l_4}$$

where

$$\xi_{u,l_1,l_2,l_3,l_4} = \ln(F_{u,l_1,l_2,l_3,l_4})$$

$$= \ln(E[f_{u,l_1,l_2,l_3,l_4}])$$

with

f_{u,l_1,l_2,l_3,l_4} = frequency in the u-th category of injury × belt usage for W (weight at level l_1, C (crash configuration) at level l_2, S (damage severity) at level l_3, and P (age/seat position) at level l_4.

The estimation of the parameters λ and the fitted values are accomplished by ECTA using an iterative proportional fitting procedure. Basically, ECTA adjusts the table to fit certain prescribed margins preserving the interaction structure in the original table specified by these margins. For each proposed model, ECTA calculates the Pearson chi-square statistic where

$$\chi_P^2 = \sum \frac{(f-\hat{F})^2}{\hat{F}} ,$$

the likelihood ratio chi-square statistic where $\chi_{LR}^2 = 2 \sum f \ln(f/\hat{F})$, the corresponding degrees of freedom, and for χ_{LR}^2 the p-value.

The decision was made to accept a model as adequate for our purposes if the corresponding p-value is reasonably large and if (within such a class of models) $\chi_P^2 \doteq \chi_{LR}^2$. A useful option incorporated in ECTA allows one to add a small quantity (e.g., .01) to each cell of the multi-dimensional table before beginning the iterative process. We have used this option whenever necessary to avoid zero marginals and to accelerate convergence of the iterative process.

Using this option and the criterion mentioned previously, we have found the appropriate model to be the one which requires fitting the following set of margins:

{WCSP}, {IWCR}, {IWSR}, {IWPR}, {ICSR}, {ICPR}, {ISPR} .

Note that this model ignores the highest order interactions.

For this model we have χ^2_{LR} = 1793 with p > .5 (d.f. = 2565) and χ^2_p = 4657. With the fitted or smoothed values given by this model used as input, we obtain the estimated standardized injury rates ($\hat{R}$'s) and the estimated true effectiveness measures ($\hat{E}$'s) presented in Table 7. The subsets considered in this table were chosen to match those produced by the collapsing required in using the GENCAT program in Chapter IV; i.e., the collapsing does not affect the ECTA modeling procedure. The overall $\hat{R}$'s and $\hat{E}$'s are repeated in Table 35 in Chapter VI to facilitate comparison with the results of the other estimation procedures.

Table 7 shows that, for each belt usage level, the injury rates are similar between the different subsets, except for Damage Severity. For Damage Severity the injury rate for SEV is about three times the injury rate for MOD. This possibly reflects on the appropriateness of collapsing the four levels of Damage Severity.

The effectiveness measures are similar across all the table for U vs. LS and for L vs. LS. (Note the "reversion" in the last entry.) However, for U vs. L, the differences are quite large.

<u>Senitivity Analysis for the Log-Linear Model Approach</u>

The following question should be raised at this point: How would the $\hat{R}$'s and $\hat{E}$'s change if we had information on only three, two, or one of the post-stratifying variables? To answer this question,

Table 7. Injury rates and true effectiveness measure using ECTA smoothing and HSRC program: Overall and selected subsets.*

		Overall	Car Weight		Damage Severity		Crash Configuration			Occupant Age/Seat Position			
			COMP	FULL	MOD	SEV	A	B	C	YD	OD	YP	OP
Injury Rate	U	.11929	.12228	.11566	.08906	.28987	.11814	.10995	.16004	.11648	.15571	.10222	.24352
	L	.09885	.11143	.08354	.07816	.21556	.10507	.07508	.15024	.10367	.08383	.08858	.10315
	L+S	.05504	.05942	.04972	.03811	.15058	.05628	.04733	.07620	.04967	.08082	.05574	.12089
Effectiveness	U/L	.1714	.0887	.2777	.1224	.2564	.1107	.3172	.0612	.1100	.4616	.1334	.5764
	U/L-S	.5386	.5140	.5702	.5721	.4805	.5236	.5695	.5239	.5736	.4810	.4547	.5036
	L/L-S	.4431	.4667	.4049	.5124	.3015	.4644	.3696	.4928	.5209	.0360	.3707	.1720

* Estimates for the subsets defined as follows:

Weight: COMP = Subcompact + Compact
FULL = Intermediate + Full

Severity: MOD = Minor + Moderate
SEV = Moderately Severe + Severe

Crash Configuration: A = Head-on with Vehicle + Head-on with Fixed Object + Rear End Striking + Angle Striking
B = Rear End, Struck + Angle, Struck in Left Side + Angle, Struck in Right Side
C = Rollover + Sideswipe + Side of Vehicle into Fixed Object

Occupant Age/Seat Position: YD = Young Driver
YP = Young Passenger
OD = Old Driver
OP = Old Passenger

we have applied our modelling and estimation procedure to each possible selection of three, two or one post-stratifying variables out of the original four (i.e., adding across the remaining variables).

With fewer post-stratifying variables, there are fewer strata, fewer cells and, therefore, less chance of having a zero cell. In fact, for some of these "smaller " models, it isn't necessary to add a small constant to all the cells before beginning the iterative process. On the other hand, a broad classification like that corresponding to only one or two post-stratifying variables may demand the use of a saturated model to obtain a good fit. Results in terms of $\hat{R}$'s and $\hat{E}$'s are displayed in Tables 8 and 9 , respectively.

Table 8 shows overall that the $\hat{R}$'s are only slightly affected when the information is gradually reduced from four to one post-stratifying variables (small reductions for $\hat{R}_1$, small increases for $\hat{R}_2$ and $\hat{R}_3$). The $\hat{E}$'s (see Table 9) are more strongly affected by these changes. Despite the large variability, $\hat{E}_1$ and $\hat{E}_2$ show a clearly increasing trend as the number of factors decreases. This is consistent with the preceeding remark on the $\hat{R}$'s. $\hat{E}_3$ doesn't show any special pattern. (Note reversions, especially in the last row of Table 9.)

Table 8. Sensitivity analysis of injury rate estimates using ECTA smoothing and HSRC program: Overall and selected subsets.

Subset			Variables in the Model														
			WCSP	WCS	WCP	WSP	CSP	WC	WS	WP	CS	CP	SP	W	C	S	P
Overall		U	.11929	.11950	.12074	.12250	.11763	.12036	.12110	.12427	.11695	.11961	.12145	.12354	.11941	.12089	.12370
		L	.09885	.10295	.08822	.08930	.10169	.08719	.08856	.08576	.10038	.09093	.08747	.08588	.08923	.08757	.08475
		L/S	.05504	.05691	.05267	.04872	.05606	.05208	.04896	.04817	.05534	.05400	.05071	.04812	.05269	.05045	.05017
Car Weight	COMP	U	.12228		.12343	.12573		.12330	.12417	.12751				.12674			
		L	.11143		.09866	.10295		.09744	.10134	.09826				.09847			
		L/S	.05942		.05821	.05468		.05805	.05591	.05524				.05512			
	FULL	U	.11566		.11748	.11857		.11677	.11737	.12033				.11964			
		L	.08354		.07552	.07267		.07470	.07300	.07054				.07056			
		L/S	.04972		.04594	.04145		.04482	.04050	.03957				.03959			
Damage Severity	MOD	U	.08906	.08938		.09153	.08771		.09067		.08742		.09101			.09049	
		L	.07816	.07616		.07139	.07892		.07154		.07819		.07054			.07062	
		L/S	.03811	.03800		.03392	.03812		.03388		.03697		.03515			.03498	
	SEV	U	.28987	.29070		.29881	.28770		.29450		.28508		.29484			.29411	
		L	.21556	.25528		.19123	.23115		.18553		.22673		.18387			.18417	
		L/S	.15058	.16443		.13294	.15800		.13484		.15996		.13939			.13863	

Table 8. Continued

Subset			WCSP	WCS	WCP	WSP	CSP	WC	WS	WP	CS	CP	SP	W	C	S	P
			Variables in the Model														
Crash Configuration	A	U	.11814	.11723	.11956		.11642	.11926			.11617	.11824			.11818		
		L	.10507	.11092	.10034		.10258	.10284			.10479	.09801			.10047		
		L/S	.05628	.05742	.05379		.05668	.05286			.05607	.05501			.05340		
	B	U	.10995	.11143	.11142		.10869	.11110			.10764	.11053			.11029		
		L	.07508	.07394	.06037		.08441	.05620			.07862	.06662			.06148		
		L/S	.04733	.04777	.04682		.04832	.04582			.04728	.04775			.04672		
	C	U	.16004	.16227	.16186		.15746	.16077			.15572	.16101			.16015		
		L	.15024	.16390	.12105		.16019	.11117			.15514	.13969			.12676		
		L/S	.07620	.08763	.06778		.08094	.07067			.08084	.07119			.07057		
Seat Position	D	U	.12031		.12251	.12386	.11894			.12619		.12150	.12309				.12559
		L	.10173		.09197	.09122	.10327			.08699		.09321	.08857				.08595
		L/S	.05271		.05023	.04514	.05387			.04546		.05137	.04792				.04734
	P	U	.11617		.11532	.11834	.11360			.11836		.11383	.11639				.11789
		L	.09007		.07672	.08340	.09685			.08198		.08392	.08406				.08105
		L/S	.06217		.06016	.05970	.06276			.05651		.06208	.05929				.05888
Age	Y	U	.11297		.11536	.11647	.11166			.11896		.11439	.11633				.11839
		L	.09996		.08977	.09227	.10364			.08748		.09245	.08917				.08655
		L/S	.05116		.04832	.04655	.05254			.04606		.05005	.04837				.04792
	O	U	.17752		.17098	.17886	.17332			.17395		.16845	.16940				.17340
		L	.08863		.07381	.06150	.08357			.06966		.07671	.07155				.06784
		L/S	.09077		.09322	.06894	.08889			.06798		.09088	.07261				.07130

Table 9. Sensitivity analysis of true effectiveness estimates using ECTA smoothing and HSRC program: Overall and selected subsets.

Subset			WCSP	WCS	WCP	WSP	CSP	WC	WS	WP	CS	CP	SP	W	C	S	P
			Variables in the Model														
Overall		U/L	.17138	.18374	.26932	.27105	.13550	.27558	.26873	.30991	.14169	.23980	.27981	.30480	.25275	.27563	.31489
		U/LS	.53858	.54046	.56377	.60233	.52345	.56727	.59572	.61234	.52680	.54858	.58245	.61048	.55877	.58267	.59440
		L/LS	.44315	.43702	.40298	.45445	.44875	.40265	.44716	.43825	.44868	.40619	.42022	.43970	.40953	.42367	.40798
Car Size	COMP	U/L	.08873	.09916	.20061	.18115		.20969	.18387	.22938				.22301			
		U/LS	.51403	.51910	.52841	.56508		.52918	.54972	.56677				.56505			
		L/LS	.46671	.46617	.41006	.46887		.40426	.44828	.43781				.44021			
	FULL	U/L	.22768	.29330	.35717	.38709		.36027	.37801	.41379				.41026			
		U/LS	.57016	.56813	.60898	.65039		.61621	.65496	.67113				.66906			
		L/LS	.40492	.33890	.39173	.42959		.40008	.44527	.43899				.43884			
Damage Severity	MOD	U/L	.12237	.13698		.22003	.10026		.21100		.10564		.22490			.21964	
		U/LS	.57206	.57612		.62941	.56535		.62629		.57714		.61384			.61348	
		L/LS	.51239	.51223		.52487	.51691		.52634		.52719		.50179			.50469	
	SEV	U/L	.25636	.27637		.36004	.19656		.37000		.20466		.37638			.37380	
		U/LS	.48054	.47787		.55509	.45083		.54212		.43888		.52725			.52865	
		L/LS	.30146	.27846		.30479	.31648		.27321		.29448		.24193			.24727	

Table 9. Continued

Subset			Variables in the Model														
			WCSP	WCS	WCP	WSP	CSP	WC	WS	WP	CS	CP	SP	W	C	S	P
Crash Configuration	A	U/L	.11066	.09310	.16081		.11890	.13766			.09797	.17104			.14984		
		U/LS	.52365	.52416	.55014		.51315	.55679			.51738	.53476			.54812		
		L/LS	.46437	.47531	.46393		.44745	.48603			.46496	.43877			.46847		
	B	U/L	.31718	.36073	.45816		.22346	.49413			.26963	.39724			.44260		
		U/LS	.56952	.57680	.57975		.55545	.58754			.56073	.56083			.57640		
		L/LS	.36955	.33800	.22440		.42753	.18465			.39857	.28336			.24003		
	C	U/L	.06123	.12308	.25212		-.01733	.30854			.00368	.13329			.20851		
		U/LS	.52387	.51809	.58125		.48595	.56043			.48082	.55785			.55936		
		L/LS	.49281	.45045	.44008		.49471	.36429			.47890	.49038			.44328		
Seat Position	D	U/L	.15442		.24925	.26353				.31068		.23278	.28045				.31561
		U/LS	.56187		.58996	.63554				.63975		.57722	.61073				.62307
		L/LS	.48186		.45382	.50513				.47738		.44895	.45901				.44925
	P	U/L	.22512		.33476	.29524				.30739		.26280	.27774				.31255
		U/LS	.46482		.47837	.49552				.52255		.45465	.49054				.50053
		L/LS	.30935		.21588	.28419				.31065		.26024	.29463				.27344
Age	Y	U/L	.11520		.22180	.20779				.26467		.19181	.23348				.26993
		U/LS	.54710		.58110	.60032				.61283		.56245	.58417				.59528
		L/LS	.48813		.46170	.49548				.47348		.45860	.45752				.44639
	O	U/L	.50073		.56829	.65616				.59953		.54462	.57764				.60977
		U/LS	.48868		.45475	.61456				.60921		.46049	.57138				.58861
		L/LS	-.02414		-.26299	-.12100				.02416		-.18475	-.01482				-.05102

IV. ESTIMATION OF INJURY RATES AND TRUE EFFECTIVENESS MEASURES AND THEIR STANDARD ERRORS USING WEIGHTED LEAST SQUARES.

Introduction

An alternative to the log-linear model approach is provided by the GSK approach (Grizzle, Starmer and Koch [1969]), i.e., by the weighted least squares analysis of categorical data. Appendix D describes how to estimate any compound function (combination of linear, logarithmic and exponential transformations) of categorical data and its standard error using this powerful approach.

The U.N.C. Biostatistics Department has recently developed a very general computer program, GENCAT, that can handle a broad scope of categorical data problems. However, the standard version of GENCAT cannot work with more than 80 functions simultaneously. It will be seen later that at least five functions (adequately defined) per stratum are needed to compute $\hat{R}$ and $\hat{E}$. Therefore, we must reduce drastically the number of strata by judicious collapsing.

Collapsing Criteria

Under which conditions would it be valid to collapse various strata? That is, under which circumstances would it be algebraically equivalent (in terms of the evaluation of the R's) to treat two strata as two separate entities or as one unique entity? The following are sufficient conditions for collapsing:

Criterion A: Collapse strata h and h' if, for each belt usage level, the "population injury rates" are equal; i.e.,

$$\frac{n_{h11}}{n_{h1\cdot}} = \frac{n_{h'11}}{n_{h'1\cdot}}\,,\ \frac{n_{h21}}{n_{h2\cdot}} = \frac{n_{h'21}}{n_{h'2\cdot}} \text{ and } \frac{n_{h31}}{n_{h3\cdot}} = \frac{n_{h'31}}{n_{h'3\cdot}} \tag{4.1}$$

Criterion B: Collapse strata h and h' if they have the same "population belt usage distribution"; i.e.,

$$\frac{n_{h1\cdot}}{n_{h\cdot\cdot}} = \frac{n_{h'1\cdot}}{n_{h'\cdot\cdot}}\,,\ \frac{n_{h2\cdot}}{n_{h\cdot\cdot}} = \frac{n_{h'2\cdot}}{n_{h'\cdot\cdot}} \text{ and } \frac{n_{h3\cdot}}{n_{h\cdot\cdot}} = \frac{n_{h'3\cdot}}{n_{h'\cdot\cdot}} \tag{4.2}$$

The sufficiency of each of these criteria can readily be seen. Under Criterion A, the "contribution" of strata h and h' to, say, R_1 is (aside from the constant $\frac{1}{n_{\cdots}}$)

$$\frac{n_{h11}}{n_{h1\cdot}}(n_{h\cdot\cdot}) + \frac{n_{h'11}}{n_{h'1\cdot}}(n_{h'\cdot\cdot}) = \frac{n_{h11}}{n_{h1\cdot}}(n_{h\cdot\cdot} + n_{h'\cdot\cdot})$$

$$= \frac{n_{h11} + n_{h'11}}{n_{h1\cdot} + n_{h'1\cdot}}(n_{h\cdot\cdot} + n_{h'\cdot\cdot}) \tag{4.3}$$

Expression (4.3) follows from Criterion A and the composition property for proportions. This equality is an identity under Criterion A and its right-hand side is the contribution of the collapsed strata h + h' to R_1. Similarly, R_2 and R_3 would remain unchanged if we collapse h and h' provided that Criterion A is true.

Under Criterion B, the contribution of strata h and h' to R_1 is

$$n_{h11}\left(\frac{n_{h\cdot\cdot}}{n_{h1\cdot}}\right) + n_{h'11}\left(\frac{n_{h'\cdot\cdot}}{n_{h'1\cdot}}\right) = (n_{h11} + n_{h'11})\left(\frac{n_{h\cdot\cdot}}{n_{h'\cdot\cdot}}\right)$$

since the first equality in (4.2) implies $\dfrac{n_{h\cdot\cdot}}{n_{h1\cdot}} = \dfrac{n_{h'\cdot\cdot}}{n_{h'1\cdot}}$.

Also $\dfrac{n_{h\cdot\cdot}}{n_{h1\cdot}} = \dfrac{n_{h\cdot\cdot} + n_{h'\cdot\cdot}}{n_{h1\cdot} + n_{h'1\cdot}}$. Thus

$$n_{h11}\left(\frac{n_{h\cdot\cdot}}{n_{h1\cdot}}\right) + n_{h'11}\left(\frac{n_{h'\cdot\cdot}}{n_{h'1\cdot}}\right) = (n_{h11} + n_{h'11})\,\frac{n_{h\cdot\cdot} + n_{h'\cdot\cdot}}{n_{h1\cdot} + n_{h'1\cdot}} \tag{4.4}$$

where the right-hand side of (4.4) is the contribution of the collapsed strata h + h' to R_1. Likewise for R_2 and R_3.

Marginal Collapsing Using ECTA and Criterion A

Both of the collapsing criteria are "population criteria." Therefore, we cannot verify them but must resort to statistical tests using the sampling information. The null hypothesis will be that the rates mentioned above have differences not significantly different from zero.

One important feature of ECTA is that, if we have an n-level factor, we can associate its n-1 degrees of freedom with n-1 "effects" or comparisons of interest by utilizing appropriate design matrices $\underline{X}$. For example, the following design matrices are useful for examining the potential for collapsing various combinations of levels of weight, of damage severity, of age/seat position and of crash configuration:

$$\underline{X} = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & -1 & -1 \end{bmatrix} \text{ for W, S and P}$$

$$\underline{X} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & -1 & 0 & 0 & -1 & 0 & 1 & 0 \end{bmatrix} \text{ for C.}$$

In this way we are comparing, for example, injury rates within each belt category for level 1 vs. level 2 of W, level 3 vs. level 4 of W and levels 1 and 2 vs. levels 3 and 4 of W. To use Criterion A, the file is divided into three subsets corresponding to the belt usage levels with a saturated model fitted to each. The tests corresponding to the specified design matrices are then carried out by ECTA yielding standardized λ test statistics which, under the null hypotheses, are approximately normally distributed.

Thus if we find that the standardized λ for one of these comparisons is sufficiently small simultaneously for unrestrained, for lap belt and for lap and shoulder belt users, we can safely collapse the levels involved in this comparison.

Proceeding in this way, we found initially that we could collapse levels 1 and 2 and levels 3 and 4 of W. Similar results hold for S. For C we can collapse levels 1 and 9 (head-on with vehicle, head-on with fixed object), 2 and 4 (rear-end striking, angle striking) and 8 and 10 (sideswipe, skidded sideways into fixed object). No marginal collapsing can be justified for P.

In summary, using ECTA and Criterion A we can reduce the W levels from 4 to 2 (say, COMP and FULL), the damage severity levels from 4 to 2 (say, MOD and SEV) and crash configuration levels from 10 to 7. Therefore, the number of strata is reduced from 640 to 112 (=2×2×4×7). Of course, these results are subject to all the possible sampling variability consequences.

Collapsing of Strata using GENCAT and Criterion B

Following the original collapsing using Criterion A, it is important to explore the possible collapsing of more specific strata ("internal" collapsing). This can be done using the GENCAT program (see Appendix D) and collapsing Criterion B. Because this would imply working with 672 (= 112×6) cells which is beyond the capacity of the standard GENCAT, it will be necessary to study separately the four subsets of the 112 strata corresponding to the different car weight (W) by damage severity (S) combinations (COMPMOD, COMPSEV, FULLMOD, FULLSEV). Each of these subsets has 28 strata consisting of combinations of crash configuration (C) and age/seat position (P).

In each of these four runs we have $r = 3$ responses (belt usage levels) and $s = 28$ populations (strata) (see Appendix D). Note again, that the zero cells are replaced by .01. Therefore the initial $\underset{\sim}{p}$ vector will contain the 84 entries corresponding to the 28 belt usage distributions; i.e.,

$$\underset{\sim}{p}' = \left[\frac{n_{11\cdot}}{n_{1\cdot\cdot}} \quad \frac{n_{12\cdot}}{n_{1\cdot\cdot}} \quad \frac{n_{13\cdot}}{n_{1\cdot\cdot}} \;\vdots\; \cdots\cdots \;\vdots\; \frac{n_{28,1\cdot}}{n_{28,\cdot\cdot}} \quad \frac{n_{28,2\cdot}}{n_{28,\cdot\cdot}} \quad \frac{n_{28,3\cdot}}{n_{28,\cdot\cdot}} \right]$$

Linear dependence within these distributions will be eliminated using the linear transformation $\underset{\sim}{A}\underset{\sim}{p}$ where $\underset{\sim}{A}$ is a 56×84 block-diagonal matrix with main blocks $\underset{\sim}{A}^* = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. This will pick up the sample proportions of unbelted and lap-belt users.

As seen in Appendix D, the linear model $\underset{\sim}{A}\underset{\sim}{p} = \underset{\sim}{X}\underset{\sim}{\pi}$, with $\underset{\sim}{X} = \underset{\sim}{I}_{56}$ = identity matrix and $\underset{\sim}{\pi}$ representing the population proportions of unbelted and lap belt users, provides the framework for testing hypotheses of the form $\underset{\sim}{C}\underset{\sim}{\pi} = \underset{\sim}{0}$. For example, to compare the belt usage distribution corresponding to crash configuration 1 vs. the distribution corresponding to crash configuration 2 <u>within</u> old drivers, it is convenient to use:

$$\underset{2\times 56}{\underset{\sim}{C}} = \begin{array}{c} \left[\begin{array}{c|cc|cc|cc|cc|cc|cc|cc|c|c} \underset{2\times14}{\underset{\sim}{0}} & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \underset{2\times14}{\underset{\sim}{0}} & \underset{2\times14}{\underset{\sim}{0}} \\ & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & & \end{array}\right] \\ \begin{array}{ccccccccccc} \text{YD} & & \underbrace{1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7}_{\text{OD}} & & \text{YP} & & \text{OP} \end{array} \end{array}$$

crash configuration

The first row of $\underset{\sim}{C}\underset{\sim}{\pi}$ contains the desired comparison for unbelted older drivers and the second row for the corresponding lap belt users. For example, if the run corresponding to COMPMOD provides evidence to reject H_0: $\underset{\sim}{C}\underset{\sim}{\pi} = \underset{\sim}{0}$, then it is adequate to collapse the strata "COMPMOD, driver over 55, crash configuration 1" and "COMPMOD, driver over 55, crash configuration 2".

Working in this way, it was possible to reduce the number of strata from 112 to 59. Further, because certain strata had very small frequencies, they were collapsed within one "weight × damage severity" category when similar collapsing had been obtained in the other corresponding categories. This, in turn, reduced the number of strata to 52.

A further step was to compare (using the same technique described above) the belt usage distribution of YD vs. OD, YP vs. OP, etc., for a given crash configuration within a given weight × damage severity combination (e.g., COMPMOD, COMPFULL, etc). These comparisons result in the final collection of 16 strata described in Table 10. With 5 functions per stratum, the GENCAT is then able to analyze the resulting information.

Use of GENCAT To Estimate R's, E's and Their Standard Errors

The collapsing described previously provides 16 (= d) strata. This d-value is large enough (for the standard GENCAT program) to require three separate runs to estimate $\{R_1, R_2, E_{12}\}$, $\{R_1, R_3, E_{13}\}$ and $\{R_2, R_3, E_{23}\}$, respectively.

To obtain $\hat{R}_1$, $\hat{R}_2$,$\hat{E}_{12}$ and their standard errors, the following information is taken from each stratum: n_{h11}, n_{h12}, n_{h21}, n_{h22}, $n_{h3\cdot}$; i.e., the number of injured and non-injured non-belt users, the number of injured and non-injured lap-belt users and the number of lap and shoulder belt users. Using these 5 responses per stratum, the set-up (in the terminology of Appendix D) is s = 1 population and r = 5d (= 80) responses. GENCAT will divide these numbers by n... (= total number of cases) in generating the vector ($\underset{\sim}{p}$) of 80 relative frequencies.

An initial linear transformation defined by the block-diagonal

Table 10. The final 16 strata for GENCAT.*

Stratum Number	Vehicle Weight	Damage Severity	Age/Seat Position	Crash Configuration	Frequency U I	Frequency U $\bar{I}$	Frequency L I	Frequency L $\bar{I}$	Frequency LS I	Frequency LS $\bar{I}$
1	COMP	MOD	OD	1+2	9	52	0	25	1	43
2	COMP	MOD	OP	1+2	9	24	0	5	2	17
3	COMP	MOD	YD+YP	1+2	144	1412	45	396	32	782
4	COMP	MOD	YD+OD+YP+OP	3+4+5	56	654	18	272	15	488
5	COMP	MOD	YD+OD+YP+OP	6+7	23	216	4	74	6	148
6	COMP	SEV	YD+OD	1+2	51	81	10	22	14	38
7	COMP	SEV	YP+OP	1+2	18	32	1	0	4	8
8	COMP	SEV	YD+OD+YP+OP	3+4+5	54	199	10	63	17	154
9	COMP	SEV	YD+OD+YP+OP	6+7	32	66	7	14	9	30
10	FULL	MOD	YD+OD+YP+OP	1+2	136	1396	32	394	14	460
11	FULL	MOD	YD+OD+YP+OP	3+4+5	57	702	8	265	14	367
12	FULL	MOD	YD+OD	6+7	23	120	5	36	2	65
13	FULL	MOD	YP+OP	6+7	4	45	1	13	0	21
14	FULL	SEV	YD+OD+YP+OP	1+2	61	89	7	27	7	34
15	FULL	SEV	YD+OD+YP+OP	3+4+5	54	188	6	47	7	80
16	FULL	SEV	YD+OD+YP+OP	6+7	15	36	0	8	0	6

*Variable categories are defined as follows:

Vehicle Weight
- COMP = Subcompact + Compact
- FULL = Intermediate + Full

Damage Severity
- MOD = Minor + Moderate
- SEV = Moderately Severe + Severe

Age/Seat Position
- YD = Young Driver
- OD = Old Driver
- YP = Young Passenger
- OP = Old Passenger

Crash Configuration
- 1 = Head-on with Vehicle + Head-on with Fixed Object
- 2 = Rear End, Striking + Angle, Striking
- 3 = Rear End, Struck
- 4 = Angle, Struck in Left Side
- 5 = Angle, Struck in Right Side
- 6 = Rollover
- 7 = Sideswipe + Side of Vehicle into Fixed Object

matrix $\underset{\sim}{A}$ (5d × 5d) with basic blocks

$$\underset{\sim}{A}^* = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

will generate an (80×1) vector $\underset{\sim}{A}\underset{\sim}{p}$ with the following entries for each stratum

$$\begin{bmatrix} p_{h11} \\ p_{h12} + p_{h12} \\ p_{h21} \\ p_{h21} + p_{h22} \\ \sum_{i,j} p_{hij} \end{bmatrix} = \begin{bmatrix} p_{h11} \\ p_{h1\cdot} \\ p_{h21} \\ p_{h2\cdot} \\ p_{h\cdot\cdot} \end{bmatrix}$$

Next, consider a block-diagonal matrix $\underset{\sim}{K}$ (2d × 5d) with basic blocks

$$\underset{\sim}{K}^* = \begin{bmatrix} 1 & -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & 1 \end{bmatrix}$$

Then $\underset{\sim}{K}\ \underset{\sim}{\ln}\ \underset{\sim}{A}\underset{\sim}{p}$ will be a (32 × 1) vector with the following entries for each stratum

$$\begin{bmatrix} \ln\left(\frac{p_{h11}}{p_{h1\cdot}}\right) p_{h\cdot\cdot} \\ \ln\left(\frac{p_{h21}}{p_{h2\cdot}}\right) p_{h\cdot\cdot} \end{bmatrix}$$

After taking exponentials and using

$$\underset{3\times 2d}{\underset{\sim}{M}} = \left[\begin{array}{cc:c:cc} 1 & 0 & & 1 & 0 \\ 0 & 1 & \cdots\cdots & 0 & 1 \\ 1 & -1 & & 1 & -1 \end{array}\right]$$

, the result will be

$$\underset{\sim}{M}\ \underset{\sim}{\exp}(\underset{\sim}{K}\ \underset{\sim}{\ln}\ \underset{\sim}{Ap}) = \begin{bmatrix} \sum_h \left(\frac{p_{h11}}{p_{h1\cdot}}\ p_{h\cdot\cdot}\right) \\ \sum_h \left(\frac{p_{h21}}{p_{h2\cdot}}\ p_{h\cdot\cdot}\right) \\ \sum_h \left(\frac{p_{h11}}{p_{h1\cdot}} - \frac{p_{h21}}{p_{h2\cdot}}\right) p_{h\cdot\cdot} \end{bmatrix} = \begin{bmatrix} \hat{R}_1 \\ \hat{R}_2 \\ \hat{R}_1 - \hat{R}_2 \end{bmatrix}$$

From the corresponding covariance matrix we can obtain s.e.$(\hat{R}_1)$ and s.e.$(\hat{R}_2)$. Finally, again taking logarithms and using exponentiation with $\underset{\sim}{L} = [-1 \quad 0 \quad 1]$, we can express $\hat{E}_{12}$ as

$$\underset{\sim}{\exp}[\underset{\sim}{L}\ \underset{\sim}{\ln}\ \underset{\sim}{M}(\underset{\sim}{\exp}\ \underset{\sim}{K}\ \underset{\sim}{\ln}\ \underset{\sim}{Ap})] = \frac{\hat{R}_1 - \hat{R}_2}{\hat{R}_1} = \hat{E}_{12}$$

The corresponding estimated standard error of $\hat{E}_{12}$ is obtained using the formulation in Appendix D. A similar procedure will give $\hat{R}_1$, $\hat{R}_3$, $\hat{E}_{13}$ and $\hat{R}_2$, $\hat{R}_3$, $\hat{E}_{23}$. The results are displayed in Table 11 and in Table 35 (see Chapter VI).

Sometimes it may be of interest to estimate the R's and E's on a subset of k (> 1) of the "original" d strata. To do this using GENCAT, it is necessary to modify $\underset{\sim}{K}$, $\underset{\sim}{M}$, and $\underset{\sim}{L}$ to keep track of the relative weight of that subset and adjust the estimates accordingly. Here $\underset{\sim}{K}$ will be a block diagonal (3k × 5d) matrix with

$$\underset{\sim}{K}^* = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \text{; likewise,}$$

Table 11. GENCAT estimates of injury rates, effectiveness measures and their standard errors for selected subsets of the 16 final strata.

	Overall	Weight: Comp	Weight: Full	Damage Severity: Mod	Damage Severity: Sev	Crash Configuration: A	Crash Configuration: B	Crash Configuration: C
Injury rates:								
U	.11969 (.004104)*	.12234 (.005744)	.11647 (.005824)	.09047 (.004032)	.28622 (.014407)	.14353 (.00775)	.10926 (.005847)	.11130 (.008536)
L	.093459 (.006720)	.11092 (.009762)	.072208 (.008482)	.072678 (.006567	.21188 (.024816)	.11497 (.012912)	.086140 (.009862)	.081674 (.012897)
LS	.050958 (.004122)	.058865 (.005591)	.041332 (.006098)	.03430 (.003699)	.14588 (.017654)	.063514 (.009206)	.047535 (.005569)	.042367 (.007198)
Effectiveness measures								
U vs L	.219170 (.061833)	.09338 (.089742)	.38002 (.082829)	.19667 (.080933)	.25972 (.093951)	.19900 (.098591)	.21164 (.099490)	.26618 (.128160)
U vs LS	.57425 (.037145)	.51885 (.050486)	.64512 (.055075)	.62086 (.044235)	.49032 (.066517)	.55750 (.067642)	.56495 (.055908)	.61934 (.070519)
L vs LS	.45476 (.058645)	.46928 (.068109)	.42760 (.110267)	.52805 (.066405)	.31150 (.114943)	.44756 (.100044)	.44816 (.090331)	.48127 (.119511)

(*) Standard error

$$\underset{4\times 3d}{\underline{M}} = [\underline{M}^* \vdots \underline{0} \vdots \underline{M}^* \vdots \underline{0} \vdots \;......]$$

with

$$\underline{M}^* = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

for each stratum of interest and $\underset{4\times 3}{\underline{0}}$ otherwise.

Finally, using

$$\underline{L} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \end{bmatrix}$$

will result in

$$\underline{\exp}(\underline{L}[\underline{\ln}\ \underline{M}(\underline{\exp}\ \underline{K}\ \underline{\ln}\ \underline{Ap})]) = \begin{bmatrix} \hat{R}^*_1 \\ \hat{R}^*_2 \\ \hat{R}^*_{12} \end{bmatrix}$$

where * is used to indicate that these estimates refer to only k ($\leq$ d) "interesting" strata. Selected results are shown in Table 11.

<u>Sensitivity Analysis Using GENCAT</u>

In a manner similar to that with ECTA, a sensitivity analysis was carried out using not all 4 but only 1, 2 or 3 of the post-stratification variables.(see Tables 12 and 13). Whenever the number of corresponding strata was small enough, the three levels of belt usage were treated simultaneously using all the information of each strata, i.e., n_{h11}, n_{h12}, n_{h21}, n_{h22}, n_{h31}, n_{h32}.

When some degree of collapsing was required to make possible the use

Table 12. Sensitivity analysis of injury rates estimates using GENCAT: Overall and selected subpopulations.

Variables in the Model

Population		Variables	WCSP	CSP	WCP	WSP	WCS	WC	WS	CS	WP	CP	SP	P	C	W	S	Unadjusted Injury Rate
Overall		U	.11969 (.0041)	.11990 (.0041)	.12338 (.0042)	.12146 (.0041)	.11960 (.0041)	.12063 (.0042)	.12086 (.0041)	.11757 (.0040)	.12398 (.0043)	.12324 (.0042)	.12143 (.0041)	.12370 (.0042)	.12032 (.0041)	.12354 (.0042)	.12089 (.0041)	.12314 (.0042)
		L	.09346 (.0067)	.08998 (.0068)	.08609 (.0066)	.08908 (.0068)	.08997 (.0068)	.09031 (.0070)	.08830 (.0067)	.09353 (.0071)	.08505 (.0066)	.08559 (.0080)	.08848 (.0068)	.08475 (.0066)	.08955 (.0070)	.08588 (.0066)	.08757 (.0067)	.08485 (.0065)
		L/S	.05096 (.0041)	.05285 (.0043)	.04913 (.0041)	.04931 (.0040)	.05092 (.0041)	.05068 (.0042)	.04970 (.0040)	.05477 (.0044)	.04900 (.0041)	.05059 (.0041)	.05096 (.0041)	.05017 (.0041)	.05203 (.0043)	.04812 (.0040)	.05045 (.0040)	.04991 (.0041)
Car Weight	COMP	U	.12234 (.0057)		.12655 (.0060)	.12452 (.0058)	.12219 (.0057)	.12387 (.0058)	.12405 (.0058)		.12694 (.0060)					.12674 (.0059)		.12644 (.0059)
		L	.11092 (.0098)		.09951 (.0097)	.10279 (.0099)	.10458 (.0100)	.10383 (.0104)	.10209 (.0098)		.09801 (.0096)					.09847 (.0096)		.09334 (.0096)
		L/S	.05887 (.0056)		.05622 (.0055)	.05604 (.0054)	.05878 (.0056)	.05711 (.0056)	.05613 (.0054)		.05502 (.0054)					.05513 (.0054)		.05531 (.0054)
	FULL	U	.11647 (.0058)		.11952 (.0060)	.11774 (.0059)	.11644 (.0058)	.11668 (.0059)	.11697 (.0058)		.12037 (.0060)					.11964 (.0060)		.11962 (.0060)
		L	.07221 (.0090)		.06976 (.0087)	.07240 (.0091)	.07219 (.0090)	.07384 (.0091)	.07151 (.0090)		.06928 (.0087)					.07056 (.0089)		.06949 (.0087)
		L/S	.04133 (.0061)		.04049 (.0060)	.04116 (.0060)	.04135 (.0061)	.04285 (.0064)	.04188 (.0061)		.04167 (.0062)					.03959 (.0059)		.04085 (.0060)
Damage Severity	MOD	U	.09047 (.0040)	.09037 (.0040)		.09114 (.0041)	.09040 (.0040)		.09084 (.0040)	.08832 (.0040)			.09105 (.0041)				.09049 (.0040)	.09071 (.0040)
		L	.07268 (.0066)	.07209 (.0065)		.07156 (.0065)	.07261 (.0066)		.07125 (.0065)	.07576 (.0069)			.07068 (.0065)				.07062 (.0064)	.07094 (.0064)
		L/S	.03430 (.0037)	.03494 (.0037)		.03395 (.0036)	.03433 (.0037)		.03436 (.0037)	.03537 (.0038)			.03467 (.0037)				.03498 (.0037)	.03472 (.0037)
	SEV	U	.28622 (.0144)	.12248 (.0067)		.29425 (.0146)	.28599 (.0144)		.29191 (.0146)	.28142 (.0141)			.29456 (.0146)				.29411 (.0146)	.29201 (.0146)
		L	.21188 (.0248)	.07838 (.0100)		.18895 (.0265)	.18891 (.0259)		.18548 (.0260)	.19480 (.0271)			.18993 (.0269)				.18417 (.0260)	.18468 (.0260)
		L/S	.14588 (.0177)	.05775 (.0069)		.13680 (.0170)	.14545 (.0176)		.13711 (.0170)	.16531 (.0193)			.14383 (.0176)				.13863 (.0167)	.14216 (.0173)

Table 12. Continued.

Variables Population			WCSP	CSP	WCP	WSP	WCS	WC	WS	CS	WP	CP	SP	P	C	W	S	Unadjusted Injury Rate
Age	Y	U		.11380 (.0042	.11760 (.0044)	.11545 (.0042)					.11868 (.0044)	.11783 (.0044)	.11575 (.0043)	.11839 (.0044)				.11782 (.0043)
		L		.09292 (.0073)	.08879 (.0071)	.09153 (.0073)					.08726 (.0070)	.08701 (.0070)	.09076 (.0073)	.08655 (.0070)				.08696 (.0070)
		L/S		.05055 (.0044)	.04634 (.0041)	.04686 (.0041)					.04647 (.0042)	.04817 (.0043)	.04874 (.0043)	.04792 (.0042)				.04755 (.0042)
	O	U		.17700 (.0159)	.17745 (.0164)	.17771 (.0160)					.17361 (.0160)	.17394 (.0161)	.17456 (.0157)	.17340 (.0160)				.17563 (.0161)
		L		.06241 (.0175)	.06083 (.0169)	.06619 (.0185)					.06439 (.0179)	.07237 (.0512)	.06714 (.0193)	.06784 (.0192)				.06593 (.0184)
		L/S		.07432 (.0185)	.07523 (.0162)	.07221 (.0153)					.07266 (.0156)	.07316 (.0155)	.07181 (.0150)	.07130 (.0150)				.07047 (.0148)
Seat Position	D	U									.12575 (.0050)	.12502 (.0050)	.12294 (.0049)	.12559 (.0050)				.12546 (.0050)
		L									.08678 (.0074)	.08726 (.0074)	.08972 (.0076)	.08595 (.0074)				.08579 (.0073)
		L/S									.04548 (.0043)	.04830 (.0045)	.04799 (.0044)	.04734 (.0044)				.04739 (.0044)
	P	U									.11855 (.0079)	.11779 (.0078)	.11679 (.0077)	.11789 (.0078)				.11720 (.0078)
		L									.07973 (.0142)	.08048 (.0234)	.08467 (.0150)	.0810[illegible] (.0144)				.08100 (.0144)
		L/S									.05983 (.0100)	.05760 (.0096)	.06011 (.0097)	.05888 (.0097)				.05993 (.0098)

Table 13. Sensitivity analysis of effectiveness estimates using GENCAT: Overall and selected subpopulations.

Variables in the Model

Variable	Population	WCSP	CSP	WCP	WSP	WCS	WC	WS	CS	WP	CP	SP	P	C	W	S	Unadjusted Injury Rate
Overall	U/L	.21917 (.0618)	.24954 (.0623)	.30219 (.0588)	.26656 (.0611)	.24773 (.0622)	.25137 (.0637)	.26940 (.0609)	.20453 (.0662)	.31400 (.0581)	.30549 (.0693)	.27135 (.0612)	.31489 (.0582)	.25573 (.0633)	.30480 (.0587)	.27563 (.0603)	.31098 (.0581)
	U/LS	.57425 (.0371)	.55924 (.0388)	.60183 (.0356)	.59406 (.0356)	.57426 (.0372)	.57991 (.0377)	.58878 (.0361)	.53418 (.0399)	.60478 (.0355)	.58955 (.0364)	.58030 (.0365)	.59440 (.0359)	.56754 (.0384)	.61048 (.0348)	.58267 (.0360)	.59467 (.0357)
	L/LS	.45476 (.0586)	.41268 (.0654)	.42940 (.0643)	.44652 (.0616)	.43406 (.0624)	.43885 (.0639)	.43714 (.0627)	.41441 (.0642)	.42387 (.0654)	.40901 (.0736)	.42401 (.0641)	.40798 (.0667)	.41895 (.0656)	.43970 (.0632)	.42387 (.0636)	.41174 (.0659)
Car Weight COMP	U/L	.09338 (.0837)		.21365 (.0850)		.14417 (.0903)	16174 (.0927)	.17701 (.0876)		.22794 (.0839)					.22301 (.0840)		.22219 (.0841)
	U/LS	.51885 (.0505)		.55575 (.0481)		.51895 (.0505)	.53899 (.0499)	.54755 (.0479)		.56658 (.0470)					.56505 (.0469)		.56255 (.0472)
	L/LS	.46928 (.0681)		.43505 (.0778)		.43792 (.0751)	.45004 (.0769)	.45023 (.0743)		.43862 (.0777)					.44021 (.0770)		.43759 (.0774)
Car Weight FULL	U/L	.38002 (.0828)		.41631 (.0787)		.38000 (.0829)	.36719 (.0841)	.38866 (.0823)		.42448 (.0782)					.41026 (.0801)		.41903 (.0786)
	U/LS	.64512 (.0551)		.66122 (.0533)		.64490 (.0551)	.63278 (.0579)	.64200 (.0553)		.65381 (.0543)					.66907 (.0520)		.65846 (.0532)
	L/LS	.42760 (.1103)		.41959 (.1130)		.42726 (.1104)	.41970 (.1124)	.41439 (.1129)		.39848 (.1173)					.43884 (.1095)		.41211 (.1139)
Damage Severity MOD	U/L	.19667 (.0809)	.20229 (.0805)		.21484 (.0792)	.19681 (.0810)		.21571 (.0792)	.14711 (.0864)			.22375 (.0790)				.21964 (.0786)	.21802 (.0790)
	U/LS	.62086 (.0442)	.61333 (.0446)		.62747 (.0432)	.62025 (.0443)		.62174 (.0439)	.60178 (.0462)			.61925 (.0439)				.61348 (.0443)	.61726 (.0440)
	L/LS	.52805 (.0664)	.51528 (.0677)		.52554 (.0666)	.52720 (.0666)		.51770 (.0677)	.53309 (.0657)			.50950 (.0689)				.50469 (.0698)	.51055 (.0682)
Damage Severity SEV	U/L	.25972 (.0940)	.36004 (.0880)		.35785 (.0957)	.33944 (.0960)		.36461 (.0946)	.30781 (.1021)			.35519 (.0968)				.37388 (.0936)	.36754 (.0946)
	U/LS	.49032 (.0665)	.52852 (.0614)		.53508 (.0621)	.49141 (.0664)		.53032 (.0628)	.41259 (.0742)			.51169 (.0645)				.52864 (.0612)	.51318 (.0640)
	L/LS	.31150 (.1149)	.26326 (.1280)		.27600 (.1355)	.23006 (.1399)		.26079 (.1383)	.15137 (.1535)			.24271 (.1416)				.24727 (.1397)	.23027 (.1433)

Table 13. Continued.

Crash Configuration	Variable / Population	WCSP	CSP	WCP	WSP	WCS	WC	WS	CS	WP	CP	SP	P	C	W	S	Unadjusted Injury Rate
A	U/L	.19900	.14673	.18799	.17452	.14003	.14358		.12119		.21462			.15287			.19090
		(.0986)	(.0901)	(.0871)	(.0877)	(.0906)	(.0912)		(.0921)		(.0850)			(.0904)			(.0969)
	U/LS	.55750	.55135	.59635	.54999	.56417	.55941		.51708		.58557			.54353			.58272
		(.0676)	(.0528)	(.0500)	(.0476)	(.0521)	(.0547)		(.0565)		(.0509)			(.0556)			(.0509)
	L/LS	.44756	.47478	.50290	.45485	.49320	.48554		.45019		.47232			.46116			.48127
		(.1000)	(.0764)	(.0749)	(.0737)	(.0743)	(.0775)		(.0792)		(.0794)			(.0802)			(.0770)
B	U/L	.21164	.39891	.45944	.38505	.40857	.44564		.36672		.43909			.44620			.45927
		(.0995)	(.0971)	(.0878)	(.0830)	(.0955)	(.0896)		(.1021)		(.0913)			(.0905)			(.0830)
	U/LS	.56495	.57104	.58870	.65079	.57477	.58003		.55594		.57141			.57640			.58756
		(.0559)	(.0628)	(.0613)	(.0540)	(.0633)	(.0627)		(.0652)		(.0636)			(.0630)			(.0612)
	L/LS	.44816	.28636	.23912	.43213	.28102	.24241		.29879		.23590			.24003			.23366
		(.0903)	(.1424)	(.1533)	(.1098)	(.1443)	(.1522)		(.1397)		(.1536)			(.1528)			(.1529)
C	U/L	.26618	.31428	.39291		.30465	.22086		.15530		.35611			.22213			.37253
		(.1282)	(.1656)	(.1503)		(.1639)	(.2084)		(.2119)		(.3216)			(.2035)			(.1553)
	U/LS	.61934	.56119	.65725		.61671	.66462		.55258		.65051			.64462			.64582
		(.0705)	(.1246)	(.0875)		(.0935)	(.0850)		(.1075)		(.0893)			(.0899)			(.0895)
	L/LS	.48127	.36008	.43543		.44878	.56956		.47032		.45722			.54313			.43554
		(.1195)	(.2233)	(.1863)		(.1231)	(.1481)		(.1708)		(.2961)			(.1551)			(.1855)

Table 13. Continued.

Variable	Population	WCSP	CSP	WCP	WSP	WCS	WC	WS	CS	WP	CP	SP	P	C	W	S	Unadjusted Injury Rate
Age Y	U/L		.18345	.24490	.20723					.26478	.26159	.21595	.26893				.26194
			(.0703)	(.0665)	(.0691)					(.0652)	(.0654)	(.0688)	(.0649)				(.0551)
	U/LS		.55577	.60598	.59412					.60842	.59116	.57895	.59528				.59645
			(.0414)	(.0381)	(.0386)					(.0381)	(.0393)	(.0397)	(.0380)				(.0385)
	L/LS		.45597	.47813	.40802					.46740	.44632	.46299	.44639				.45323
			(.0634)	(.0626)	(.0606)					(.0643)	(.0663)	(.0634)	(.0663)				(.0551)
Age O	U/L		.64740	.65720	.62752					.62912	.58396	.61537	.60077				.62458
			(.1033)	(.1003)	(.1094)					(.1008)	(.2963)	(.1157)	(.1161)				(.1102)
	U/LS		.53909	.57607	.59369					.58148	.57938	.58066	.58081				.59875
			(.1107)	(.0992)	(.0929)					(.0975)	(.0971)	(.0932)	(.0942)				(.0921)
	L/LS		-.19090	-.23667	-.09082					-.12844	-.01101	-.69427	-.05102				-.06879
			(.4456)	(.4331)	(.3820)					(.3959)	(.7466)	(.3000)	(.3697)				(.3735)
Position D	U/L									.30987	.30203	.27021	.31561				.31617
										(.0651)	(.0657)	(.0602)	(.0646)				(.0646)
	U/LS									.63034	.61362	.60966	.62307				.62226
										(.0371)	(.0393)	(.0391)	(.0383)				(.0384)
	L/LS									.47596	.44643	.46513	.44925				.44761
										(.0667)	(.0700)	(.0670)	(.0698)				(.0700)
Position P	U/L									.32750	.31676	.27503	.31256				.30981
										(.1276)	(.2037)	(.1362)	(.1306)				(.1314)
	U/LS									.49536	.51104	.48532	.50053				.48950
										(.0906)	(.0874)	(.0892)	(.0884)				(.0903)
	L/LS									.24961	.28435	.29007	.27344				.26142
										(.1830)	(.2396)	(.1694)	(.1760)				(.1787)

of this "economic" set-up, the insight gained from the collapsing effort described previously provided the guideline for such collapsing.

Additional Considerations

During this analysis, it was noted that the covariances between the injury rates are extremely small (usually of the order of 10^{-5} or 10^{-6}). This suggested repeating the modelling assuming that the injury rates are stochastically independent. How would this assumption affect the precision estimates?

Assuming that $\hat{R}_1$, $\hat{R}_2$ are independent and have variances v_1, v_2, an estimate of the variance of $\hat{E}_{1,2}$ is given by

$$\left(\frac{\hat{R}_1-\hat{R}_2}{\hat{R}_1}\right)^2 \left[\frac{v_1+v_2}{(\hat{R}_1-\hat{R}_2)^2} - \frac{2v_1}{\hat{R}_1(\hat{R}_1-\hat{R}_2)} + \frac{v_1}{\hat{R}_1^2}\right] \qquad (3.1)$$

(derived from (D.2) in Appendix D). From (3.1) we have for overall estimates s.e.$(\hat{E}_{1,2})$ = .0622 (.0618 in the original analysis), s.e. $(\hat{E}_{1,3})$ = .0374 (vs. .0371) and s.e.$(\hat{E}_{2,3})$ = .0590 (vs. .0586). Moreover, if we compare the last two columns of Table 35, we can see that the loss of precision induced by this assumption never exceeds .0013 and it is usually of order .0005.

This loss can be counterbalanced by the fact that, without the need of keeping track of covariances, the GENCAT approach would require two functions per strata (instead of 5) and consequently the number of strata can be increased from 16 to 40 and this would be clearly beneficial in terms of requiring less collapsing.

Current efforts are directed at further exploring the consequences of making these assumptions.

V. INTERIM "FACT BOOK" TABLES

The tables in this section are based on the expanded data base of approximately 16,000 weighted observations. These tables describe differential belt usage and occupant injury as a function of vehicle size, model year, vehicle damage, crash configuration and occupant age and seating position (vehicle, crash, and occupant characteristics).

The variables have been described previously and the values of certain variables have been grouped to facilitate comparisons.

Table 14. Overall belt effectiveness

AIS $\geq$ 2

	Effectiveness	Standard Error
Lap vs unrestrained	21.9%	.0618
Lap and shoulder vs unrestrained	57.4%	.0371
Lap and shoulder vs lap	45.5%	.0586

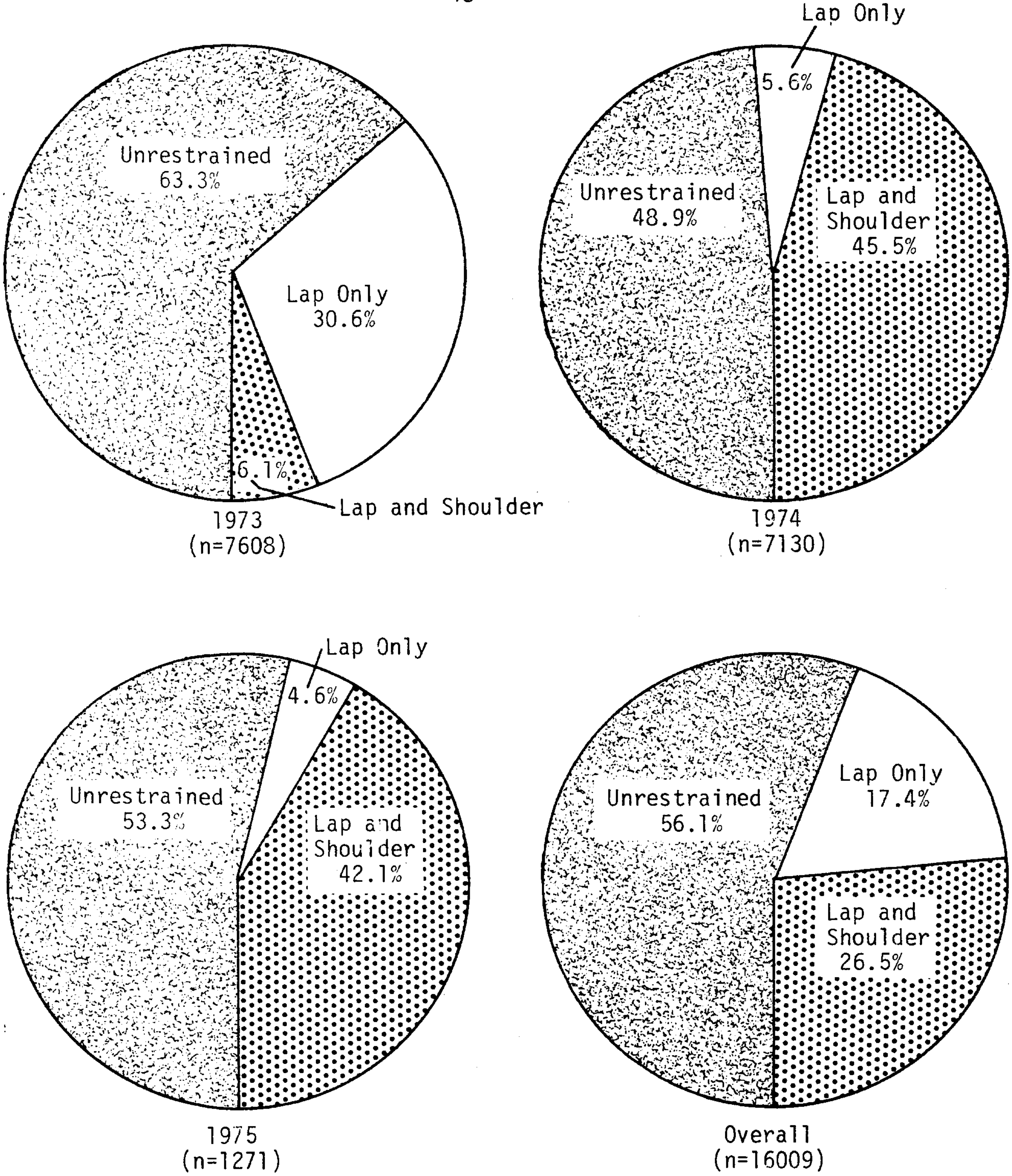

Figure 4 . Restraint system usage distribution by model year.

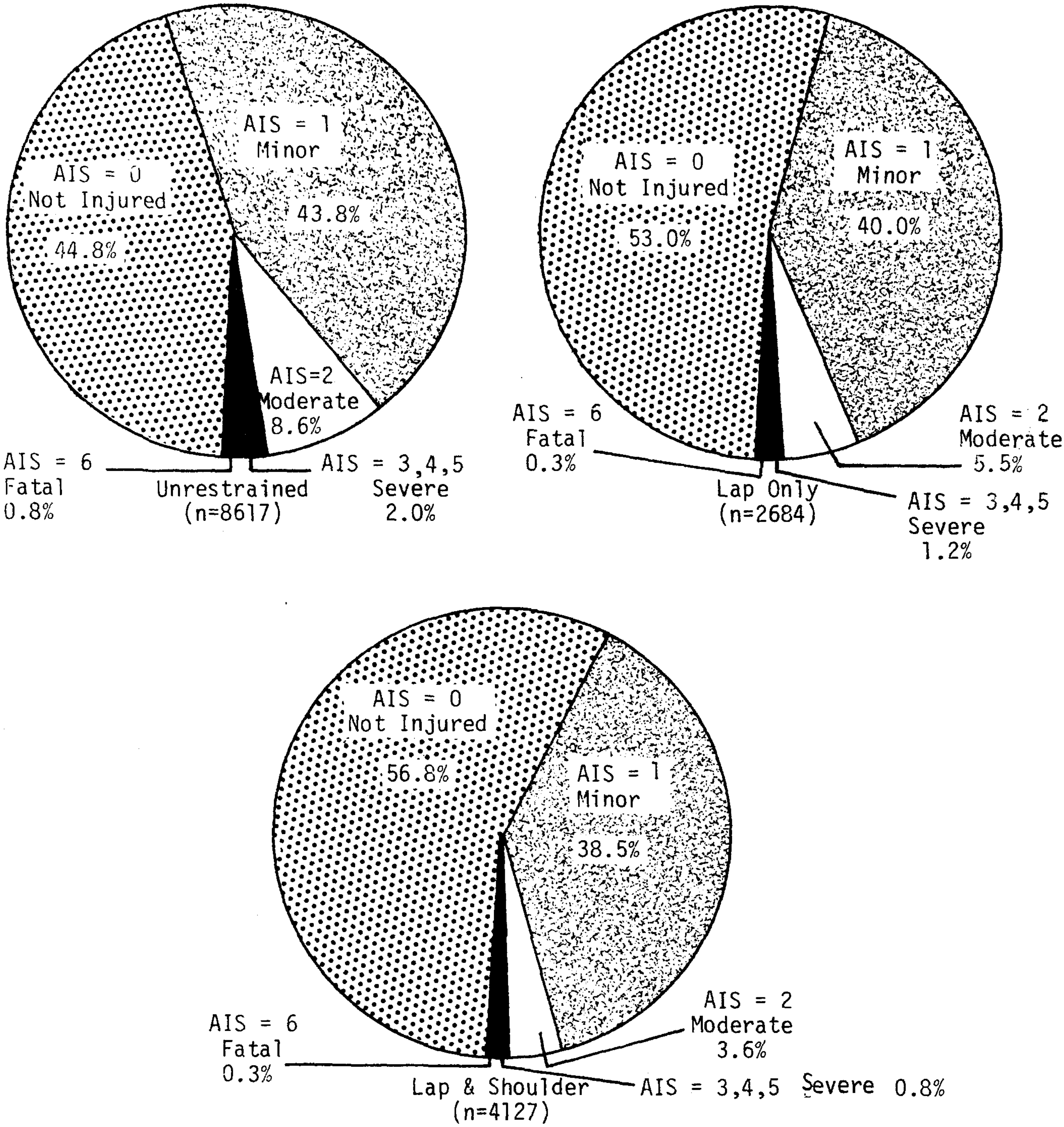

Figure 5. AIS distribution by restraint system usage.

Table 15. Belt effectiveness by vehicle size

AIS $\geq$ 2

		Effectiveness	Standard Error
COMP* (< 3600 lbs.)	U vs L	9.3%	.0897
	U vs LS	51.9%	.0505
	L vs LS	46.9%	.0681
FULL ($\geq$ 3600 lbs.)	U vs L	38.0%	.0828
	U vs LS	64.5%	.0551
	L vs LS	42.8%	.1103

*COMP = (subcompact) + (compact)

FULL = (intermediate) + (full-sized)

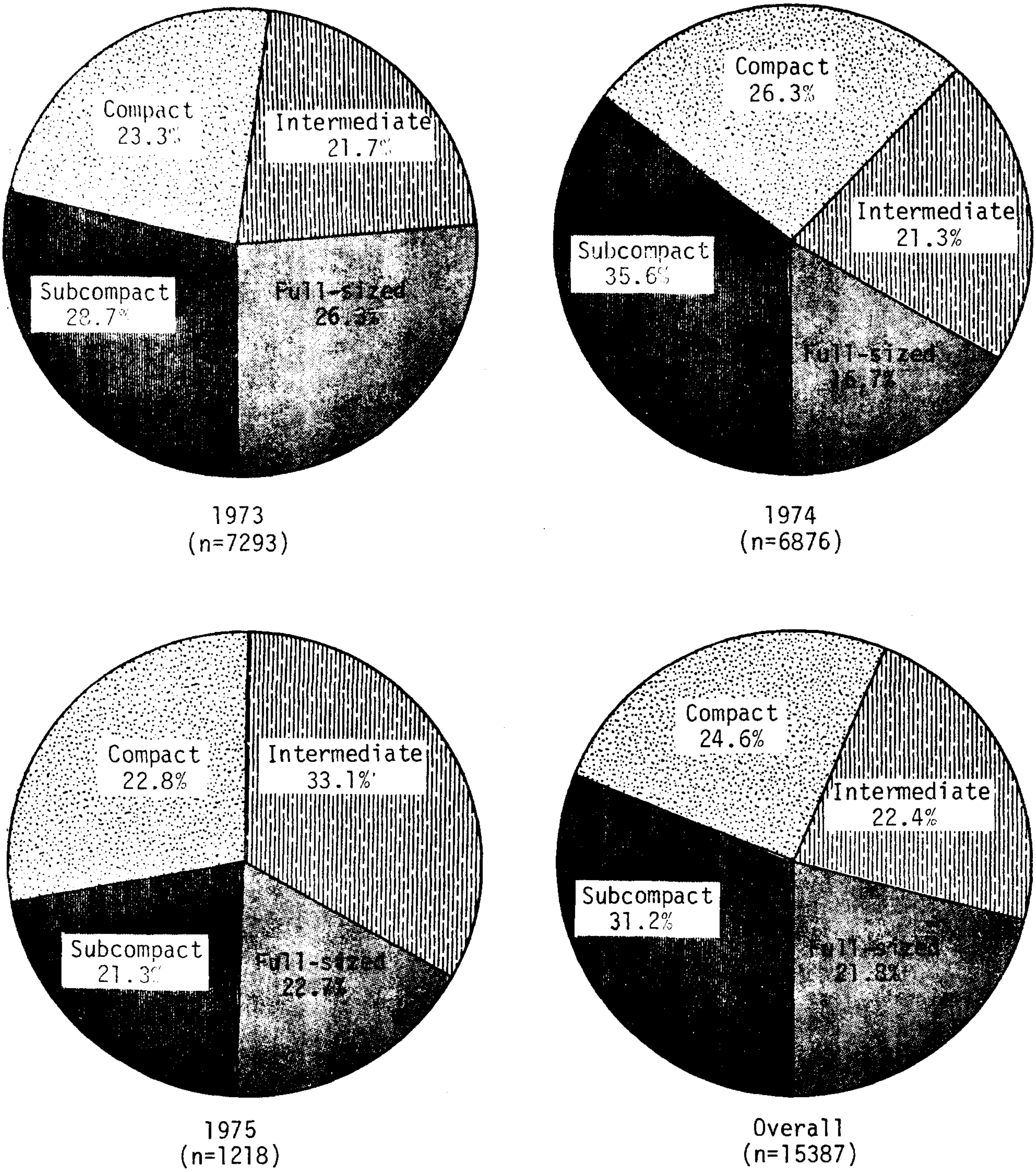

Figure 6. Vehicle size distribution by model year.

Table 16. Percent restraint use by vehicle size and model year.

Vehicle Size \ Model Year	1973			1974			1975			Overall		
	No Restraint	Lap Only	Lap & Shoulder	No Restraint	Lap Only	Lap & Shoulder	No Restraint	Lap Only	Lap & Shoulder	No Restraint	Lap Only	Lap & Shoulder
Sub-compact (<2700 lbs)	58.0	31.9	10.1	49.0	3.3	47.7	44.6	2.7	52.7	52.7	15.7	31.6
Compact (2700-3599 lbs)	62.0	31.5	6.5	48.1	4.5	47.5	48.2	2.2	49.6	54.3	16.4	29.3
Intermediate (3600-4100 lbs)	67.9	27.9	4.2	50.8	6.0	43.2	56.8	7.9	35.2	59.3	16.3	24.4
Full-sized (>4100 lbs)	67.6	30.0	2.4	47.3	8.5	44.1	60.3	4.0	35.7	60.1	20.5	19.5
Overall	63.6	30.4	6.0	48.9	5.0	46.1	53.0	4.6	42.4	56.2	17.0	26.8

Table 17. AIS distribution of unrestrained occupants by vehicle size.

Vehicle Size \ AIS	0 No Injury	1 Minor	2 Moderate	3,4,5 Severe	6 Fatal
Subcompact (<2700 lbs)	41.4	45.8	9.8	2.2	0.8
Compact (2700-3599 lbs)	44.7	45.1	7.8	1.7	0.7
Intermediate (3600-4100 lbs)	44.2	43.9	9.0	2.0	0.8
Full-Sized (>4100 lbs)	47.4	41.8	7.9	2.1	0.8
Overall	44.2	44.3	8.7	2.0	0.8

Table 18. Vehicle size distribution of unrestrained occupants by AIS.

Vehicle Size / AIS	Subcompact (<2700 lbs)	Compact (2700-3599 lbs)	Intermediate (3600-4100 lbs)	Full-Sized (>4100 lbs)
0 Not injured	27.6	24.1	23.6	24.8
1 Minor	30.5	24.3	23.4	21.9
2 Moderate	33.1	21.5	24.3	21.0
3,4,5 Severe	32.4	20.0	23.5	24.1
6 Fatal	30.3	20.0	24.6	24.6
Overall	29.5	23.8	23.5	23.2

Table 19. Belt effectiveness by crash configuration

AIS $\geq$ 2

		Effectiveness	Standard Error
	U vs L	19.9%	.0986
A*	U vs LS	55.8%	.0676
	L vs LS	44.8%	.1000
	U vs L~~S~~	21.2%	.0995
B	U vs LS	56.5%	.0559
	L vs LS	44.8%	.0903
	U vs L	26.6%	.1282
C	U vs LS	61.9%	.0705
	L vs LS	48.1%	.1195

*A = (head-on with vehicle) + (head-on with fixed object) + (rear-end striking) + (angle striking)

B = (rear-end, struck) + (angle, struck in left side) + (angle, struck in right side)

C = (rollover) + (sideswipe) + (crashed sideways into fixed object)

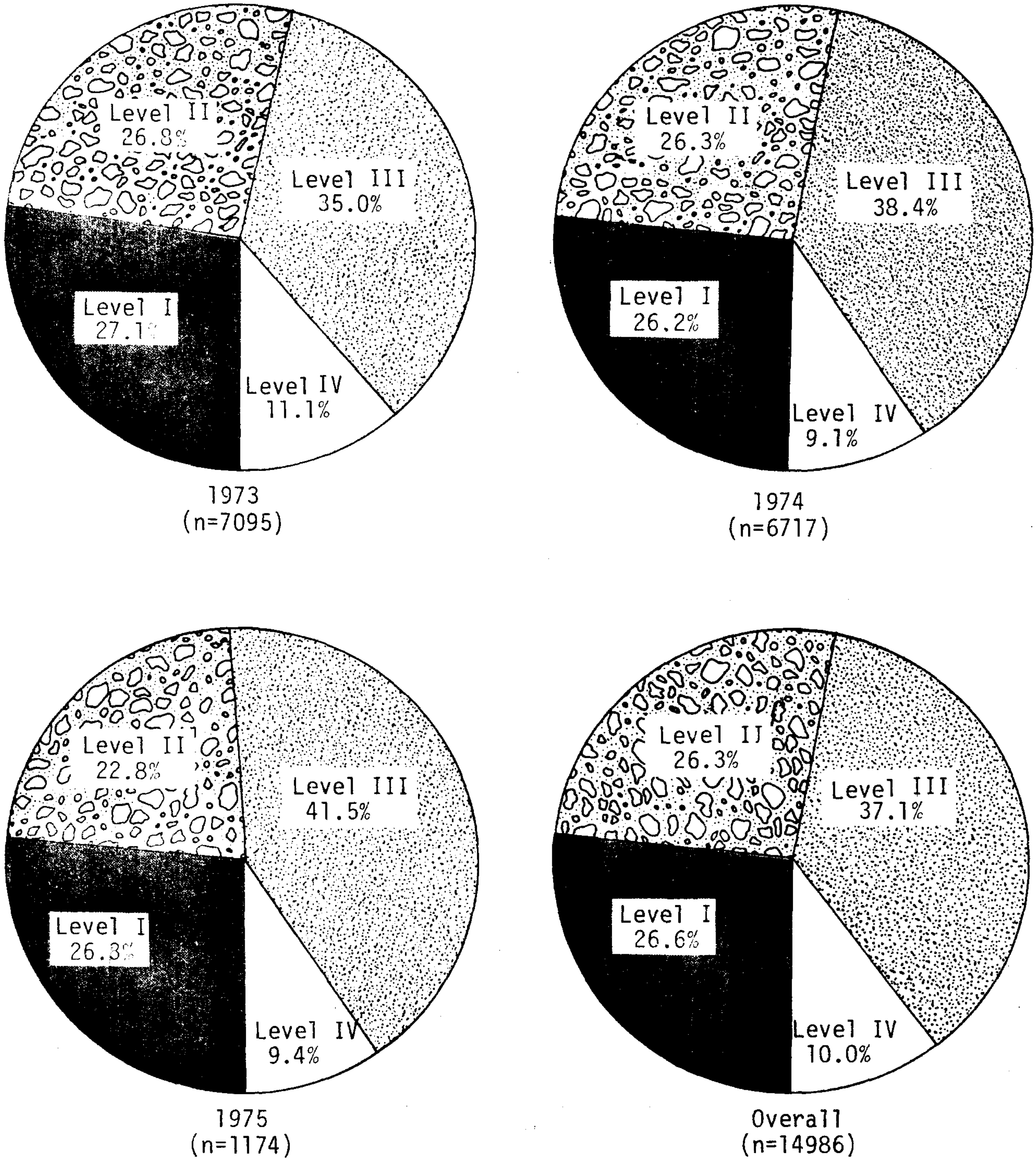

Figure 7 . Crash configuration levels* by model year.

*Crash configurations were divided into four levels as follows:

Level I = (Head-on with another vehicle) + (Head-on with fixed object) + (Rollover) + (Crashed sideways into fixed object)
Level II = (Angle, struck in left) + (Angle, struck in right)
Level III = (Rear-end striking) + (Angle, striking)
Level IV = (Rear-end struck) + (Sideswipe)

Table 20. Percent restraint use by crash configuration and model year.

Crash Configuration* \ Model Year	1973			1974			1975			Overall		
	No Restraint	Lap Only	Lap & Shoulder	No Restraint	Lap Only	Lap & Shoulder	No Restraint	Lap Only	Lap & Shoulder	No Restraint	Lap Only	Lap & Shoulder
Level I	69.8	25.5	4.7	57.6	5.5	37.0	59.2	6.8	34.0	63.6	15.2	21.2
Level II	63.5	30.5	6.0	46.0	4.7	49.3	49.3	4.9	45.9	54.7	17.2	28.2
Level III	63.9	30.9	5.2	49.3	5.6	45.1	54.6	3.9	41.5	56.3	16.8	26.9
Level IV	52.3	34.5	13.2	39.5	5.2	55.2	47.3	2.7	50.0	46.7	20.3	32.9
Overall	64.1	29.7	6.2	49.7	5.3	45.0	53.9	4.8	41.3	56.9	16.8	26.3

* See figure 7 for detailed description of crash configuration levels.

Table 21. Percent restraint use by crash configuration levels* and model year.

Crash Configuration* \ Model Year	1973			1974			1975			Overall		
	No Restraint	Lap Only	Lap & Shoulder	No Restraint	Lap Only	Lap & Shoulder	No Restraint	Lap Only	Lap & Shoulder	No Restraint	Lap Only	Lap & Shoulder
Level I	68.9	28.7	2.4	55.0	4.6	40.5	47.3	12.7	40.0	61.2	16.5	22.3
Level II	63.3	30.8	5.9	45.3	6.7	48.0	54.9	2.3	42.8	53.8	16.4	29.8
Level III	49.4	36.9	13.7	36.6	4.2	59.2	41.7	0.0	58.3	44.1	22.1	33.8
Level IV	64.2	31.1	4.7	52.3	4.8	42.9	54.4	5.1	40.4	58.0	17.1	24.9
Level V	62.4	30.6	7.0	46.5	5.2	48.3	46.2	6.2	47.7	54.3	17.8	27.9
Level VI	64.6	30.4	5.0	45.6	4.2	50.2	52.2	3.6	44.2	55.1	16.5	28.4
Level VII	76.9	13.2	9.9	63.6	5.3	31.1	61.5	7.7	30.8	68.6	8.5	22.9
Level VIII	60.2	28.0	11.8	44.1	7.1	48.4	57.9	7.9	34.2	52.5	16.5	31.0
Level IX	73.7	22.3	4.0	58.5	5.9	35.7	60.2	5.8	33.9	66.4	14.3	19.3
Level X	57.6	33.8	8.6	56.8	5.8	37.3	65.7	4.3	30.0	58.0	18.0	24.0
Overall	64.1	29.7	6.2	49.7	5.3	45.0	53.9	4.8	41.3	56.9	16.8	26.3

*Crash configuration levels are as follows:

Level I. Head-on with vehicle
Level II. Rear end, striking
Level III. Rear end, struck
Level IV. Angle, striking
Level V. Angle, struck in left side
Level VI. Angle, struck in right side
Level VII. Rollover
Level VIII. Sideswipe
Level IX. Head-on with fixed object
Level X. Skidded sideways into fixed object

Table 22. AIS distribution of unrestrained occupants by crash configuration.

Crash Configuration* \ AIS	0 No Injury	1 Minor	2 Moderate	3,4,5 Severe	6 Fatal
Level I	40.6	42.4	11.6	3.6	1.8
Level II	41.3	46.9	9.1	1.9	0.8
Level III	48.6	42.9	7.3	1.1	0.1
Level IV	42.6	52.0	4.2	0.9	0.3

* See figure 7 for detailed description of crash configuration levels.

Table 23. Crash configuration distribution of <u>unrestrained</u> occupants by AIS.

AIS \ Crash Configuration*	Level I	Level II	Level III	Level IV
0 Not injured	27.2	23.8	41.0	8.1
1 Minor	28.0	26.6	35.6	9.7
2 Moderate	38.9	26.2	30.9	4.0
3,4,5 Severe	52.1	23.6	20.6	3.6
6 Fatal	68.8	25.0	3.1	3.1
Overall	29.4	25.3	37.0	8.3

* See figure 7 for detailed description of crash configuration levels.

Table 24. Belt effectiveness by damage severity.

AIS $\geq$ 2

		Effectiveness	Standard Error
MOD*	U vs L	19.7%	.0809
	U vs LS	62.1%	.0442
	L vs LS	52.8%	.0664
SEV	U vs L	26.0%	.0940
	U vs LS	49.0%	.0665
	L vs LS	31.2%	.1149

*MOD = (minor) + (moderate)

SEV = (moderately severe) + (severe)

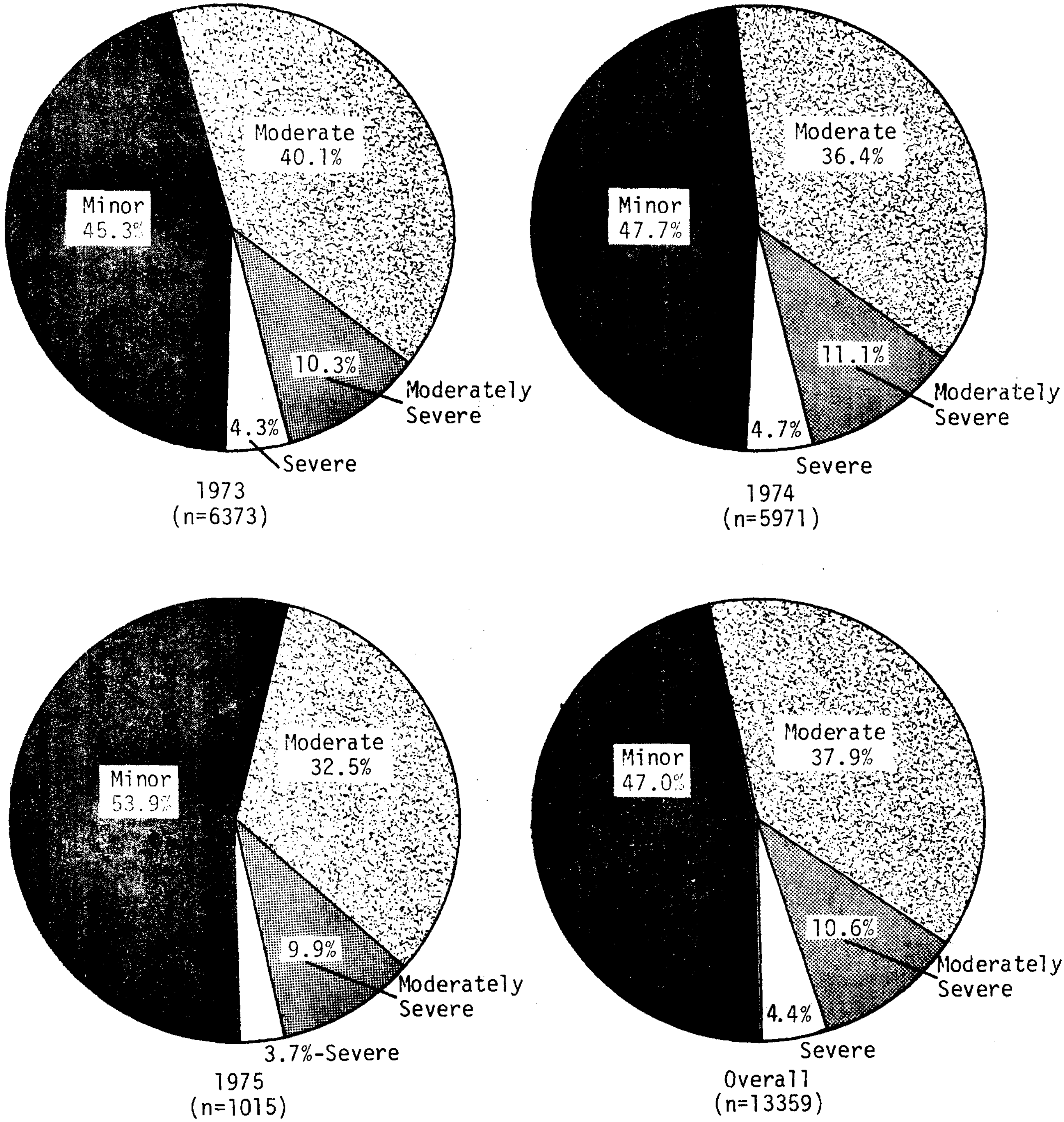

Figure 8. Damage severity distribution by model year.

Table 25. Percent restraint use by damage severity and model year.

Damage Severity* \ Model Year	1973			1974			1975			Overall		
	No Restraint	Lap Only	Lap & Shoulder	No Restraint	Lap Only	Lap & Shoulder	No Restraint	Lap Only	Lap & Shoulder	No Restraint	Lap Only	Lap & Shoulder
Minor	62.9	31.9	5.2	49.1	4.7	46.2	51.9	4.9	43.1	55.7	17.2	27.1
Moderate	64.8	28.0	7.2	49.7	5.2	45.1	52.7	6.7	40.6	57.5	16.8	25.6
Moderately Severe	70.6	24.8	4.6	56.0	4.7	39.3	54.0	3.0	43.0	62.6	13.9	23.5
Severe	67.0	25.6	7.3	51.9	3.5	44.5	68.4	7.9	23.7	59.9	14.0	26.1
Overall	64.6	29.3	6.0	50.2	4.9	44.9	53.0	5.4	41.6	57.3	16.6	26.1

* as described in Chapter I.

Table 26. AIS distribution of unrestrained occupants by damage severity.

Damage Severity \ AIS	0 No Injury	1 Minor	2 Moderate	3,4,5 Severe	6 Fatal
Minor	52.9	40.6	5.0	1.1	0.3
Moderate	38.5	50.0	9.6	1.5	0.5
Moderately Severe	23.2	52.0	18.1	4.9	1.8
Severe	21.2	39.7	19.7	11.6	7.8

Table 27. Damage severity distribution of <u>unrestrained</u> occupants by AIS.

AIS \ Damage Severity	Minor	Moderate	Moderately Severe	Severe
0 Not injured	57.0	34.4	6.3	2.3
1 Minor	40.9	41.8	13.2	4.1
2 Moderate	25.8	40.7	23.3	10.2
3,4,5 Severe	22.8	26.5	25.9	24.7
6 Fatal	14.1	20.3	23.4	42.2
Overall	45.8	38.0	11.5	4.7

Table 28. Seat position distribution by model year.

Seat Position \ Model Year	1973	1974	1975	Overall
Left front (driver)	74.0	73.5	76.3	74.0
Right front	26.0	26.5	23.7	26.0
N	7608	7130	1271	16009

Table 29. Percent restraint use by seat position and model year.

Seat Position \ Model Year	1973			1974			1975			Overall		
	No Restraint	Lap Only	Lap & Shoulder	No Restraint	Lap Only	Lap & Shoulder	No Restraint	Lap Only	Lap & Shoulder	No Restraint	Lap Only	Lap & Shoulder
Front left (usually driver)	60.7	32.4	6.9	46.4	5.5	48.1	51.9	4.6	43.5	53.6	18.2	28.1
Front right	70.9	25.3	3.8	55.9	5.6	38.5	57.8	4.7	37.5	63.1	14.9	22.0
Overall	63.3	30.6	6.1	48.9	5.5	45.5	53.3	4.6	42.1	56.1	17.4	26.5

Table 30. AIS distribution of unrestrained occupants by seat position.

Seat Position \ AIS	0 No Injury	1 Minor	2 Moderate	3,4,5 Severe	6 Fatal
Front left	45.5	42.8	8.8	2.0	0.8
Front right	43.0	46.2	8.1	2.0	0.6

Table 31. Seat position distribution of <u>unrestrained</u> occupants by AIS.

AIS \ Seat Position	Front Left	Front Right
0 Not injured	71.8	28.2
1 Minor	69.0	31.0
2 Moderate	72.4	27.6
3,4,5 Severe	70.5	29.5
6 Fatal	77.3	22.7
Overall	70.7	29.3

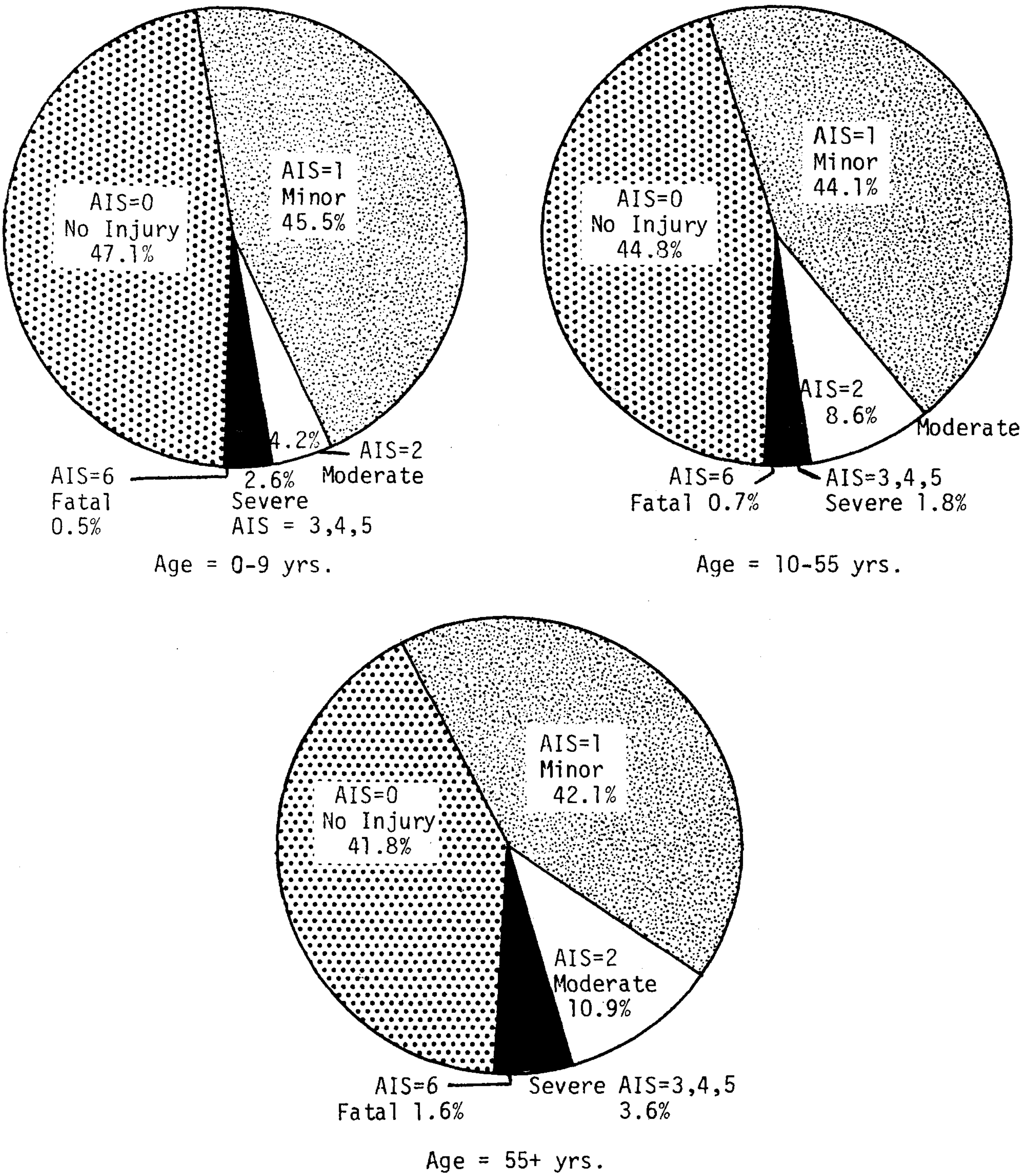

Figure 9. AIS distribution of <u>unrestrained</u> occupants by age.

Table 32. Percent restraint use by occupant age and model year.

Occupant Age \ Model Year	1973			1974			1975			Overall		
	No Restraint	Lap Only	Lap & Shoulder	No Restraint	Lap Only	Lap & Shoulder	No Restraint	Lap Only	Lap & Shoulder	No Restraint	Lap Only	Lap & Shoulder
< 5	82.9	17.1	0.0	70.0	15.0	15.0	28.6	28.6	42.9	74.5	16.8	8.8
5-14	78.3	20.4	1.4	62.3	5.4	32.3	63.2	53.3	31.6	71.0	13.5	15.5
15-24	64.7	29.4	5.9	52.2	4.5	43.3	50.9	3.2	45.9	57.9	15.9	26.2
25-34	61.2	31.2	7.6	45.7	6.6	47.6	55.1	4.6	40.3	53.7	17.8	28.5
35-44	62.8	31.9	5.3	42.4	6.2	51.4	60.2	8.0	31.8	53.9	18.8	27.3
45-54	59.8	36.0	4.3	46.8	6.2	47.0	46.6	1.8	51.5	53.4	20.8	25.9
≥ 55	59.6	32.3	8.1	44.0	5.2	50.8	56.6	7.0	36.4	53.1	19.3	27.6
Overall	63.2	30.7	6.1	48.7	5.5	45.7	53.5	4.7	41.8	56.0	17.4	26.6

Table 33. AIS distribution of unrestrained occupants by occupant age.

Occupant Age \ AIS	0 No Injury	1 Minor	2 Moderate	3,4,5 Severe	6 Fatal
< 5	54.7	43.2	1.1	1.1	0.0
5-14	46.2	44.4	6.5	2.5	0.4
15-24	46.1	43.4	8.1	1.8	0.6
25-34	42.9	46.0	8.4	1.9	0.7
35-44	41.1	46.2	10.6	1.6	0.5
45-54	44.7	42.7	8.7	2.4	1.4
≥ 55	43.9	40.2	10.9	3.4	1.6

Table 34. Occupant age distribution of unrestrained occupants by AIS.

AIS \ Occupant Age	<5	5-14	15-24	25-34	35-44	45-54	≥55
0 Not injured	1.4	3.4	43.9	22.6	10.2	9.2	9.4
1 Minor	1.1	3.3	41.8	24.5	11.6	8.9	8.7
2 Moderate	0.1	2.4	39.8	22.8	13.6	9.3	12.0
3,4,5 Severe	0.6	4.0	37.6	22.0	8.7	11.0	16.2
6 Fatal	0.0	1.5	33.3	21.2	7.6	16.7	19.7
Overall	1.1	3.2	42.4	23.4	11.1	9.2	9.5

VI. DISCUSSION; PLANS FOR FUTURE WORK

This interim report describes in considerable detail two different but parallel techniques for deriving injury rates and effectiveness measures for various belt usage categories. Both control for important variables which have a confounding effect on injury level. The limitations and/or advantages of each are pointed out and the results of various sensitivity analyses presented.

Table 35 summarizes comparable results for both procedures and presents the unadjusted rates, effectiveness measures and standard errors as a baseline. In addition to overall estimates, these measures are illustrated for the following subgroups: car weight (compact, full); crash configuration (A ("aggressive"), B ("passive"), C ("other")); damage severity (moderate, severe); occupant seat position (driver, right front seat occupant); and occupant age (young (10-55), old (over 55)).

The effect of assuming independence of the injury rates is shown in the last column of Table 35. Obviously there will be no effect on the effectiveness estimate; furthermore, there is but a slight effect on the estimate of the corresponding standard error. This assumption will reduce the collapsing requirements of GENCAT to 40 factor level combinations rather than the 16 used in Chapter IV. Current efforts are being directed at applying the procedures developed in Chapter IV using this assumption.

Table 35. Injury rates and effectiveness measures.

	Population	Estimate	Restraint System	Estimation Procedure: Unadjusted	I. Log-linear model	II. GENCAT and log-linear model	II. Assuming independence of the $\hat{R}$'s
	Overall	$\hat{R}$	U	.123 (.0042)[1]	.119	.120 (.0041)[2]	
			L	.085 (.0065)	.099	.093 (.0067)	
			LS	.050 (.0041)	.055	.051 (.0041)	
		$\hat{E}$	U vs L	.311 (.0581)	.171	.219 (.0618)	.219 (.0622)
			U vs LS	.595 (.0357)	.539	.574 (.0371)	.574 (.0374)
			L vs LS	.412 (.0659)	.443	.455 (.0586)	.455 (.0590)
CAR WEIGHT	COMP	$\hat{R}$	U	.126 (.0059)	.122	.122 (.0057)	
			L	.098 (.0096)	.111	.111 (.0098)	
			LS	.055 (.0054)	.059	.059 (.0056)	
		$\hat{E}$	U vs L	.222 (.0841)	.089	.093 (.0897)	.093 (.0904)
			U vs LS	.562 (.0472)	.514	.519 (.0505)	.519 (.0510)
			L vs LS	.438 (.0774)	.467	.469 (.0681)	.469 (.0687)
	FULL	$\hat{R}$	U	.120 (.0060)	.116	.116 (.0058)	
			L	.069 (.0087)	.084	.072 (.0090)	
			LS	.041 (.0060)	.050	.041 (.0061)	
			U vs L	.419 (.0786)	.278	.380 (.0828)	.380 (.0831)
			U vs LS	.658 (.0532)	.570	.645 (.0551)	.645 (.0553)
			L vs LS	.412 (.1139)	.405	.428 (.1103)	.428 (.1105)
CRASH CONFIGURATION	A	$\hat{R}$	U	.122 (.0055)	.118	.144 (.0078)	
			L	.098 (.0096)	.105	.115 (.0129)	
			S	.051 (.0058)	.056	.064 (.0092)	
		$\hat{E}$	U vs L	.191 (.0869)	.111	.199 (.0986)	.199 (.0999)
			U vs LS	.583 (.0509)	.524	.558 (.0676)	.557 (.0685)
			L vs LS	.484 (.0770)	.464	.448 (.1000)	.448 (.1013)
	B	$\hat{R}$	U	.112 (.0071)	.110	.109 (.0058)	
			L	.061 (.0091)	.075	.086 (.0099)	
			S	.046 (.0062)	.047	.048 (.0056)	
		$\hat{E}$	U vs L	.458 (.0880)	.317	.212 (.0995)	.212 (.0996)
			U vs LS	.588 (.0612)	.570	.565 (.0559)	.565 (.0560)
			L vs LS	.239 (.1529)	.370	.448 (.0903)	.448 (.0904)
	C	$\hat{R}$	U	.167 (.0155)	.160	.111 (.0085)	
			L	.105 (.0241)	.150	.082 (.0129)	
			S	.059 (.0139)	.076	.042 (.0072)	
		$\hat{E}$	U vs L	.372 (.1553)	.061	.266 (.1282)	.266 (.1288)
			U vs LS	.646 (.0895)	.524	.619 (.0705)	.619 (.0710)
			L vs LS	.436 (.1855)	.493	.481 (.1195)	.481 (.1203)

Table 35. Continued

	Population	Estimate	Restraint System	Estimation Procedure: Unadjusted	I. Log-linear model	II. GENCAT and log-linear model	II'. Assuming independence of the $\hat{R}$'s
DAMAGE SEVERITY	MOD	$\hat{R}$	U	.091 (.0040)	.089	.090 (.0040)	
			L	.071 (.0064)	.078	.073 (.0066)	
			LS	.035 (.0037)	.038	.034 (.0037)	
		$\hat{E}$	U vs L	.218 (.0790)	.122	.197 (.0809)	.197 (.0809)
			U vs LS	.617 (.0440)	,572	.621 (.0442)	.621 (.0442)
			L vs LS	.511 (.0682)	.512	.528 (.0664)	.528 (.0664)
	SEV	$\hat{R}$	U	.292 (.0146)	.290	.286 (.0144)	
			L	.185 (.0260)	.212	.212 (.0248)	
			LS	.142 (.0173)	.150	.146 (.0176)	
		$\hat{E}$	U vs L	.368 (.0946)	.256	.260 (.0940)	.260 (.0944)
			U vs LS	.513 (.0640)	.480	.490 (.0665)	.490 (.0668)
			L vs LS	.230 (.1433)	.302	.312 (.1149)	.312 (.1160)
SEAT POSITION	Driver	$\hat{R}$	U	.125 (.0050)	.120		
			L	.086 (.0073)	.102		
			LS	.047 (.0044)	.053		
		$\hat{E}$	U vs L	.316 (.0646)	.154		
			U vs LS	.622 (.0384)	.562		
			L vs LS	.448 (.0700)	.482		
	Passenger	$\hat{R}$	U	.117 (.0078)	.116		
			L	.081 (.0144)	.090		
			LS	.060 (.0098)	.062		
		$\hat{E}$	U vs L	.309 (.1314)	.225		
			U vs LS	.489 (.0903)	.465		
			L vs LS	.261 (.1787)	.309		
AGE	Young	$\hat{R}$	U	.118 (.0043)	.113		
			L	.087 (.0070)	.100		
			LS	.048 (.0042)	.051		
		$\hat{E}$	U vs L	.262 (.0651)	.115		
			U vs LS	.596 (.0385)	.547		
			L vs LS	.453 (.0651)	.488		
	Old		U	.176 (.0161)	.178		
			L	.066 (.0184)	.089		
			LS	.070 (.0148)	.091		
			U vs L	.624 (.1102)	.501		
			U vs LS	.599 (.0921)	.489		
			L vs LS	-.069 (.3735)	-.024		

[1]Standard error calculated using Taylor series expansion

[2]Standard error calculated using GENCAT program

It is of interest to note how relatively little the injury rates change for the various procedures. Even the baseline or unadjusted injury rates are not extensively modified by the different adjustment procedures. The effectiveness measures do change more, but that arises from slight changes in each of two injury rates becoming magnified by the relative change (or effectiveness measure).

A variety of analyses are currently underway. Revisions and refinements of the methodology described herein have been carried out and are about to be applied to the complete data set. As of mid-December, 1975, this represents a weighted sample in excess of 17,000 observations. To date, the dependent variable has been the proportion of occupants with AIS $\geq$ 2 which represents 9.7% of the towaway sample. The upcoming analyses will also investigate belt effectiveness for AIS $\geq$ 3 (or 2.3% of the sample) and for AIS = 6 = fatal (or 0.6% of the sample).

An alternative to using the categorical variable AIS to define an occupant's injury severity is to use the associated direct costs of medical bills, lost wages, etc., due to the injuries sustained. HSRC has secured certain physician, hospital, disability days, funeral, etc., costs from a variety of sources. From these combined sources, a direct cost will be estimated for the injuries of each Level 2 occupant. These direct costs will then become the dependent variable in a revised analysis of belt effectiveness.

Finally, there are clearly differences (e.g., completeness) in the data from the 5 investigation teams. The extent of these differences is being explored. In addition, the national representativeness of the Level 2 data is being examined by comparing such variables as crash configuration, vehicle size, age of occupant, etc., with 1974 accidents in North Carolina and in New York State. This data will be restricted to outboard front seat occupants in 1973-75 model cars involved in "towaway-type" crashes. To the extent possible, the Level 2 data will also be compared with national data (e.g., Accident Facts published by the National Safety Council).

Finally, a "Fact Book" about towaway accidents involving new cars is being prepared (see Chapter V for some initial tables). These tables will show differential belt usage as a function of vehicle type, occupant characteristics, accident type and location; association of the VDI with vehicle type and occupant injury; and a variety of other aspects considered to be of general interest, like examination of belt-caused injuries.

REFERENCES

Campbell, B.J. (1970) Driver injury in automobile accidents involving certain car models. Chapel Hill: University of North Carolina Highway Safety Research Center.

Forthofer, R.N. and Koch, G.G. (1973) An analysis for compounded functions of categorical data. Biometrics, 29, 143-157.

Goodman, L.A. (1970) The multivariate analysis of qualitative data: Interactions among multiple classifications. Journal of the American Statistical Association, 65, 226-256.

Goodman, L.A. (1971) The analysis of multi-dimensional contingency tables: Stepwise procedures and direct estimation methods for building models for multiple classification. Technometrics, 13, 33-61.

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

Hochberg, Y. and Reinfurt, D.W. (1974) An investigation of safety belt usage and effectiveness. Chapel Hill: University of North Carolina Highway Safety Research Center.

Kahane, C.J., Lee, S.N. and Smith, R.A. (1975) A program to evaluate active restraint system effectiveness. Paper presented at the Fourth International Congress on Automotive Safety, San Francisco, California.

Marsh, J.C.IV (1973) Vehicle occupant injury classification. Ann Arbor: University of Michigan HIT Lab Reports, Vol. 4, No. 1.

APPENDIX A

NHTSA Occupant Restraint System Summary Form.

Occupant Restraint System Summary Form

Part 1: — General Information —

Card 1

REPORT NO.					Prepared By:
Team	Accident Date yr	mo	day	Sequence No.	
1	2	3	4	5	6 7 8
2	3 4	5 6	7 8	9 10 11	12-13 14-15 16-17

Vehicle Category [9] 18
1. Control Vehicle
2. ACRS Vehicle

Hospitalized Sample Vehicle (Front Seat Occupants) [10] 19
1. Yes
2. No

Occupants
1 - 8 Actual Number
9 - Unknown

Front Seat [11] 20 Total [12] 21

Accident Location

State [1 | 3] 22 23 County [| 14 |] 24 25 26 Municipality [| 15 |] 27 28 29 30

Reporting Jurisdiction [16] 31
1. Local/Municipal
2. County
3. State
4. Federal

Check Items Submitted (1. Yes; 2. No)

Medical Form [17] 32 Vehicle Form [18] 33 Photographs [19] 34 Driver Interview [20] 35 Police Report [21] 36

Type of Accident (First Harmful Event) [22] 37

Collision With
1. Pedestrain
2. Pedalcycle
3. RR Train
4. Other Collision
5. Motor Vehicle
6. Fixed Object

Non-Collision
7. Overturn
8. Other Non-Coll.
9. Unknown

Area [23] 38
1. Urban
2. Rural

Type of Impact [24] 39
1. Head On
2. Rear end
3. Angle
4. Side Swipe
5. Rollover
6. Other
7. Not Applicable
9. Unknown

Number of Vehicles

Total [25] 40 Case [26] 41

1-7 Actual No.
8. Eight(8) or More
9. Unknown

Total Number

Killed [2 | 7] 42 43 Injured [2 | 8] 44 45

00-98 Actual No.
99 Unknown

To include the Total Killed/Injured in the Accident.

Occupant Ejected [29] 46
1. Yes
2. No
3. Not Stated
9. Unknown

Number of Lanes

Trafficway [30] 47 48

01-98 Actual No.
99- Unknown/Not Stated

Limited Access [31] 49
1. Yes
2. No
3. Not Applicable
9. Unknown

Road Surface [32] 50
1. Paved
2. Unpaved
3. Not Applicable
9. Unknown

Surface Condition [33] 51
1. Dry
2. Wet
3. Snow
4. Ice
5. N/A
6. Other:
9. Unknown

Day of Week [34] 52
1. Mon
2. Tue
3. Wed
4. Thu
5. Fri
6. Sat
7. Sun
9. Unk

Time of Accident [| 35 |] 53 54 55 56

0000-Midnight
0615-6:15 am
1200-Noon
1815-6:15 pm
2359-11:59pm
9999-Unknown

Light Condition [36] 57
1. Daylight
2. Dawn
3. Dusk
4. Dark
5. Dark-Lighted
6. Dark-Not Lighted
7. Not Stated
9. Unknown

ATTACHMENT

DEPARTMENT OF TRANSPORTATION
NATIONAL HIGHWAY TRAFFIC SAFETY ADMINISTRATION

Occupant Restraint System Summary Form

Part 2: —Vehicle Information—

Card 2/1

Prepared By:

Date: Yr. 7__ Mo.____

REPORT NO.

Team	Accident Date yr	mo	day	Sequence No.
2	3 4	5 6	7 8	9 10 11

Vehicle Identification Number (Left Justify) Unit Production Nos. — 37 — 12 13 14 15 16 17 18

Odometer Reading — 38 — 19 20 21 22 23 24

Vehicle Make/Model Code (Make: ______________) — 39 (25 26) 40 (27 28 29)

The five digit Make/Model code to be extrapolated from SECTION 5-1973 Editing Manual and Reference Information-MDAI Rpt. Automation & Utilization.

Body Style — 41 (30)
1. 2dr. Hardtop
2. 2dr. Sedan/Coupe
3. 4dr. Hardtop
4. 4dr. Sedan
5. Stationwagon
8. Other
9. Unknown

Number of Cylinders — 42 (31)
1. Rotary
2. 2-cyl.
4. 4-cyl.
6. 6-cyl.
8. 8-cyl.
9. Unknown.

Model Year — 7 43 (32)
3. 1973
4. 1974
5. 1975
9. Unknown

Transmission — 44 (33)
1. Automatic
2. Standard
9. Unknown

Air/Conditioned — 45 (34)
1. Yes
2. No
9. Unknown

Test Buzzer/Warning Light-Ignition Interlock (Explain Details on Vehicle Field Form) — 46 (35)
1. System Tested NORMAL
2. System Tested ABNORMAL
3. Unable to Test-Damage
4. Not Applicable
9. Unknown

Vehicle Towing Another Vehicle — 47 (36)
1. Yes
2. No
9. Unknown

Type of Front Seat — 48 (37)
1. Bench
2. Bucket
9. Unknown

Evidence of Restraint System (Details on Veh. Field Form)

	Malfunction LF	Malfunction CF	Malfunction RF	Defeat LF	Defeat CF	Defeat RF	Maladjustment LF	Maladjustment CF	Maladjustment RF
Box	49	50	51	52	53	54	55	56	57
Col.	38	39	40	41	42	43	44	45	46

CODES: (1. Yes; 2. No; 3. N/A; 9. Unk:) - Code Each of the Nine Blocks.

OBJECTS CONTACTED & C.D.C.

Chronological Order	Object Contacted	Collision Deformation Classification			Inches Crush
1	58 (47 48)	59 (49 50)	60 61 62 63 (51 52 53 54)	64 (55)	65 (56 57)
2	66 (58 59)	67 (60 61)	68 69 70 71 (62 63 64 65)	72 (66)	73 (67 68)
3	74 (69 70)	75 (71 72)	76 77 78 79 (73 74 75 76)	80 (77)	81 (78 79)

Number of Event Producing Most Severe Injury — 82 (80)

0- No Injury
1-7 Event No.
8- Pre-Crash Event
9- Unknown

Occupant Restraint System Summary Form

Part 3: — Restraint Usage Information — Card 3 / T

Prepared By:	REPORT NO.		
	Team	Accident Date (yr, mo, day)	Sequence No.
Date: Yr. 7_ Mo. ___	2	3 4 5 6 7 8	9 10 11

Active Restraint System Usage

Left Front	Center Front	Right Front
83 (12)	84 (13)	85 (14)

1. No Occupant
2. None Used
3. Lap & Torso
4. Lap Only
5. Torso Only
6. Child Seat
7. Other (Desc. on Veh. Form)
8. Not Known if Occupied
9. Unknown if Restraint Used

FACTORS DETERMINING USAGE CLASSIFICATION

	Belt/Fittings Damaged by Occupant Load	Location or Condition of Belts	System Defeated	Ext. Veh. Damage or Occupant Contact Points	Police Report	Police or Witness Observation	Subject Interview	Other Interview	Occupant Injury Pattern	Belt Caused Injury	Occupant Ejected
Left Front	86 (15)	87 (16)	88 (17)	89 (18)	90 (19)	91 (20)	92 (21)	93 (22)	94 (23)	95 (24)	96 (25)
Center Front	97 (26)	98 (27)	99 (28)	100 (29)	101 (30)	102 (31)	103 (32)	104 (33)	105 (34)	106 (35)	107 (36)
Right Front	108 (37)	109 (38)	110 (39)	111 (40)	112 (41)	113 (42)	114 (43)	115 (44)	116 (45)	117 (46)	118 (47)

1. Supported Evaluation
2. Contrary to Evaluation
3. Neither Supported or Contradicted Evaluation
4. Position Not Occupied
5. Not Applicable
9. Unknown

Reliability of Information

Left Front	Center Front	Right Front
119 (48)	120 (49)	121 (50)

1. Certain
2. Reliable
3. Unreliable
4. Unable to Estimate

Occupant Restraint System Summary Form

Part 4: Occupant Information

Card $\frac{K}{1}$

Complete one card for each Front Seat Occupant

(N=34 Occupant No.)

REPORT NO.

Team	Accident Date						Sequence no.		
	yr		mo		day				
2	3	4	5	6	7	8	9	10	11

Occupant Role — 122 (12)

1. Driver
2. Passenger
9. Unknown

Seat Position — 123 (13)

1. Left Front
2. Center Front
3. Right Front
4. Other
9. Unknown

Ejection or Entrapment — 124 (14)

1. Not Ejected/ Not Trapped
2. Ejected (Degree Not Stated)
3. Partial Eject and Trapped
4. Partial Ejection
5. Total Ejection
6. Trapped
9. Unknown

Sex — 125 (15)

1. Male
2. Female
9. Unknown

Age (Years) — 126 (16 17)

00-97 Actual
98 - 98yrs.or Over
99 - Unknown

Height (Inches) — 127 (18 19)

01-98 Inches
99- Unknown

Weight (Pounds) — 128 (20 21 22)

001-998 Pounds
999- Unknown

Police Injury Code — 129 (23)

1. (K) (Fatal)
2. (A) (Incapacitating)
3. (B) (Non-incapt'g)
4. (C) Possible)
5. (O) No-Injury)
9. (U) Unknown

Treatment — Mortality — 130 (24)

0. Not Injured
1. First Aid at Scene
2. Stated-would consult MD
3. Directed to Consult MD
4. Did consult MD
5. Emer.Rm. Treatment-Rel.
6. Admitted to Hosp.-Non Fatal
7. Fatal
8. Other
9. Unknown

OCCUPANT INJURY CLASSIFICATION — Injury Detail — Use O.I.C. Code

Coding for Belt Caused Category ONLY:

0. No
1. Possible
2. Probable
3. Definite
9. Unknown

Injury Number	Body Region	Aspect	Lesion	Sys/Organ	Severity	Belt Caused
1	131 (25)	132 (26)	133 (27)	134 (28)	135 (29)	136 (30)
2	137 (31)	138 (32)	139 (33)	140 (34)	141 (35)	142 (36)
3	143 (37)	144 (38)	145 (39)	146 (40)	147 (41)	148 (42)
4	149 (43)	150 (44)	151 (45)	152 (46)	153 (47)	154 (48)
5	155 (49)	156 (50)	157 (51)	158 (52)	159 (53)	160 (54)
6	161 (55)	162 (56)	163 (57)	164 (58)	165 (59)	166 (60)

More Than Six Injuries ? — 167 (61)

1. Yes -Note Details on Med. form
2. No
9. Unknown

Occupant Pregnant ? — 168 (62)

1. Yes
2. No
3. N/A (Male)
9. Unknown

APPENDIX B

Codebook for Extract File A.

<u>Var. 1</u>: Team (Var.1 on Occupant Restraint System Summary Form)

1. CALSPAN (W. New York)
2. U. of Miami
3. HSRI (S.E. Michigan)
4. SWRI (S. Texas)
5. USC (Los Angeles)

<u>Var. 2</u>: Accident year (2)

4. 1974
5. 1975

<u>Var. 3</u>: Accident month (3)

1. January
2. February
3. March
4. April
5. May
6. June
7. July
8. August
9. September
10. October
11. November
12. December

<u>Var. 4</u>: Sequential number (5)

(3 digit numeric)

<u>Var. 5</u>: Case weight factor (Function of 1, 2, 3, 10)

1. Sampled at 100%
2. Sampled at 50%
3. Sampled at 33%

<u>Var. 6</u>: Restraint system usage (83, 85)

2. No restraints used
3. Lap and shoulder belts
4. Lap belt only
9. Unknown

<u>Var. 7</u>: AIS injury (129, 130, 135)

0. Not injured
1. Minor
2. Moderate
3. Severe
4. Serious nonfatal
5. Critical nonfatal
6. Fatal
9. Unknown

Var. 8: Crash configuration (22, 24, 58-63)

0. Unknown
1. Head-on with veh
2. Rear end, striking
3. Rear end, struck
4. Angle, striking
5. Angle, struck in left side
6. Angle, struck in right side
7. Rollover
8. Other noncollision
9. Sideswipe
10. Head-on with fixed object
11. Side of vehicle into fixed object

Var. 9: Case vehicle weight (37, 39, 40)

0. Unknown
1. Subcompact
2. Compact
3. Intermediate
4. Full-sized

Var. 10: Damage severity (24, 58-64)

0. Unknown
1. Minor (e.g., 12-FDEW-1, 12-FYEW-1, 12-FLEW-1, 12-FLEE-1, 12-FLEE-2)
2. Moderate (e.g., 12-FDEW-2, 12-FYEW-2, 12-FLEW-2, 12-FLEW-3, 12-FLEE-3, 12-FLEE-4)
3. Moderately severe (e.g., 12-FDEW-3, 12-FYEW-3, 12-FLEW-4, 12-FLEE-5)
4. Severe (e.g., 12-FDEW-4, 12-FYEW-4, 12-FLEW-5, 12-FLEE-6)

Var. 11: Occupant age group (126)

0. Unknown
1. Under 10
2. 10-55
3. Over 55

Var. 12: Occupant position (122, 123)

1. Driver
2. Passenger

Var. 13: Occupant sex (125)

1. Male
2. Female
3. Unknown

Var. 14: Vehicle model year (43)

3. 1973
4. 1974
5. 1975

Var. 15: Exact occupant age (126)

0. Less than 1 year
1 - 97. Exact age in years
98. 98 years or more
99. Unknown

APPENDIX C

Log-Linear Model (Goodman)

Consider a 2-way contingency table with variables A (levels $i = 1,\ldots,I$) and B (levels $j = 1,\ldots,J$). If P_{ij} (>0) denotes the probability that an observation will fall in cell (i,j) and $\nu_{ij} = \ln p_{ij}$, then it is possible to write

$$\nu_{ij} = \mu + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB} \tag{C.1}$$

where the λ_i^A and λ_j^B represent "main effects," λ_{ij}^{AB} represents an "interaction effect," and

$$\sum_i \lambda_i^A = \sum_j \lambda_j^B = \sum_i \lambda_{ij}^{AB} = \sum_j \lambda_{ij}^{AB} = 0 \tag{C.2}$$

Because $\sum_{i,j} p_{ij} = 1$, μ satisfies the condition

$$(\exp \mu)\left(\sum_{i,j} \exp(\lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB})\right) = 1$$

Equivalently, considering a sample of n observations, define

f_{ij} = number of observations in cell (i,j),

$F_{ij} = E[f_{ij}]$,

$y_{ij} = \ln f_{ij}$ and

$\xi_{ij} = \ln F_{ij}$.

Then

$$\xi_{ij} = \theta + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$$

with the same conditions for the λ*'s and* θ *satisfying*

$$(\exp\ \theta)\Big(\sum_{i,j} \exp(\lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB})\Big) = n$$

Denoting $\nu_{i\cdot} = \sum_j \nu_{ij}/J$, $\nu_{\cdot j} = \sum_i \nu_{ij}/I$ and $\nu_{\cdot\cdot} = \sum_{i,j} \nu_{ij}/IJ$,

we have, as in the usual analysis of variance,

$$\lambda_i^A = \nu_{i\cdot} - \nu_{\cdot\cdot}\ ,\ \lambda_j^B = \nu_{\cdot j} - \nu_{\cdot\cdot}\ ,\ \lambda_{ij}^{AB} = \nu_{ij} - \nu_{i\cdot} - \nu_{\cdot j} + \nu_{\cdot\cdot}$$

and maximum likelihood estimates of the λ's are given by (c.4) with the ν's replaced by the corresponding y's. It follows that $\hat{\lambda}$ will have the form $\hat{\lambda} = \sum_{i,j} a_{ij}y_{ij}$ where the a_{ij}'s are constants with $\sum_{i,j} a_{ij} = 0$. For example, if I = J = 3,

$$\hat{\lambda}_{11}^{AB} = y_{11} - y_{1\cdot} - y_{\cdot 1} + y_{\cdot\cdot} = y_{11} - \frac{1}{3}\sum_j y_{1j} - \frac{1}{3}\sum_i y_{i1} + \frac{1}{9}\sum_{i,j} y_{ij}$$

$$= \frac{1}{9}\Big\{4y_{11} - 2(y_{12} + y_{13} + y_{21} + y_{31}) + y_{22} + y_{23} + y_{32} + y_{33}\Big\}\ .$$

The variance of $\hat{\lambda}$ can be estimated by

$$s_{\hat{\lambda}}^2 = \sum_{i,j} a_{ij}^2/f_{ij}\ .$$

The standardized value of $\hat{\lambda}$, i.e., $\hat{\lambda}/s_{\hat{\lambda}}$ is asymptotically distributed as a standard normal variate and can be used to test H_0: $\lambda = 0$. The values of each $\hat{\lambda}$, $s_{\hat{\lambda}}$ and $\hat{\lambda}/s_{\hat{\lambda}}$ can be obtained using the ECTA program.

Because of condition (C.2) there is a basic set of I-1 parameters associated with the effect A, a basic set of J-1 parameters associated with the effect B and a basic set of (I-1)(J-1) parameters associated with the interaction AB. This implies that interest can be placed on the I-1 parameters, λ_i^A or on any other I-1 independent linear combinations of them. This was very useful in the marginal collapsing discussed in Chapter IV.

In addition to the estimation and hypothesis testing under the saturated model, the log-linear approach makes it possible to study unsaturated models (i.e., assuming some λ's equal to zero). Goodman's work in this area has usually been restricted to hierarchical models (i.e., if a certain effect is assumed nil, any interaction parameter involving such an effect will be set to zero). For example, if the basic set $\lambda^B = \{\lambda_1^B, \ldots, \lambda_J^B\}$ is set to zero, also $\lambda^{AB} = \{\lambda_{11}^{AB}, \ldots, \lambda_{IJ}^{AB}\}$ will be set to zero, which is equivalent to saying that "the levels of the variable B are equiprobable, given the level of variable A." The maximum likelihood estimates $\hat{F}_{ij}$ will be calculated as the product of effects corresponding to the non-zero λ's by fitting the marginal distribution of A.

The goodness-of-fit of such a model (and the validity of the underlying null hypothesis) can be judged using the Pearson χ^2 statistic

$$\chi_p^2 = \sum_{i,j} [(f_{ij} - \hat{F}_{ij})^2/\hat{F}_{ij}]$$

or (recommended) the likelihood-ratio criterion

$$\chi^2_{LR} = 2 \sum_{i,j} f_{ij} \ln(f_{ij}/\hat{F}_{ij})$$

The degrees of freedom for these statistics are given by the number of λ's set to zero. In our example, df = (J-1) + (I-1)(J-1) = I(J-1).

APPENDIX D

Contingency Table Analysis for Compounded Logarithmic - Exponential - Linear Functions

Grizzle, Starmer and Koch (1969) describe how linear regression models and weighted least squares can be used to either test hypotheses or fit simplified models to multi-dimensional contingency tables which arise when frequency counts are obtained for respective cross-classifications of specific qualitative variables. Briefly, assuming an underlying product multinomial model for the cell frequencies and certain regularity conditions on $\underline{F}(\underline{p}) = (F_1(\underline{p}),\ldots,F_u(\underline{p}))$, a set of functions of the cell proportions, attention is directed at fitting a linear model

$$E(\underline{F}(\underline{p})) = \underline{X}\underline{\beta}$$

where $\underline{X}$ is a known $(u\times t)$ coefficient matrix of full rank $t \leq u$ and $\underline{\beta}$ is an unknown $(t\times 1)$ parameter vector. Weighted least squares provides the BAN estimator

$$\underline{b} = \underline{\hat{\beta}} = (\underline{X}'\underline{V}_F^{-1}\underline{X})^{-1}\underline{X}'\underline{V}_F^{-1}\underline{F}$$

where

$$\underline{V}_F = \underline{H}\underline{V}(\underline{p})\underline{H}'$$

with

$$\underline{H} = [d\underline{F}(\underline{x})/d\underline{x}\,|\,\underline{x}=\underline{p}]$$

$\underline{V}(\underline{p})$ is block-diagonal with matrices

$\underline{V}_i(\underline{p}_i) = (\underline{D}_{\underline{p}_i} - \underline{p}_i\underline{p}_i')/n_i$ on the main diagonal with $\underline{D}_{\underline{p}_i}$ a diagonal matrix with

$\underset{\sim}{p}_i$ on the main diagonal, $i = 1, 2, \ldots, s$ = number of populations.

Also

$$\underset{\sim}{V}_b = \text{var}(\underset{\sim}{b}) = (\underset{\sim}{X}'\underset{\sim}{V}_F^{-1}\underset{\sim}{X})^{-1}$$

A goodness of fit test statistic is given by

$$X_F^2 = SS(\mathcal{E}(\underset{\sim}{F}) = \underset{\sim}{X}\underset{\sim}{\beta}) = \underset{\sim}{F}'\underset{\sim}{V}_F^{-1}\underset{\sim}{F} - \underset{\sim}{b}'(\underset{\sim}{X}'\underset{\sim}{V}_F^{-1}\underset{\sim}{X})\underset{\sim}{b}$$

which, under the rule hypothesis that the model fits, is approximately χ^2 (df = u-t). Given an adequate fit, general linear hypotheses H_c: $\underset{\sim}{C}\underset{\sim}{\beta} = \underset{\sim}{0}$, where $\underset{\sim}{C}$ is a known (d×t) matrix of full rank $d \leq t$, can be tested using

$$X_c^2 = SS(\underset{\sim}{C}\underset{\sim}{\beta} = \underset{\sim}{0}) = \underset{\sim}{b}'\underset{\sim}{C}'[\underset{\sim}{C}(\underset{\sim}{X}'\underset{\sim}{V}_F^{-1}\underset{\sim}{X})^{-1}\underset{\sim}{C}']^{-1}\underset{\sim}{C}\underset{\sim}{b}$$

which, under H_c, is approximately χ^2 (df = d).

Grizzle, Starmer, and Koch (1969) restrict attention to linear functions $\underset{\sim}{F}(\underset{\sim}{p}) = \underset{\sim}{A}\underset{\sim}{p} = \underset{\sim}{a}$ and log-linear functions

$$\underset{\sim}{F}(\underset{\sim}{p}) = \underset{\sim}{K}[\underset{\sim}{\ln}(\underset{\sim}{A}\underset{\sim}{p})] = \underset{\sim}{f}$$

where $\underset{\sim}{A}$ and $\underset{\sim}{K}$ are known matrices and $\underset{\sim}{\ln}$ transforms a vector to the corresponding vector of natural logarithms.

Forthofer and Koch (1973) extend the previous work to exponential functions of the type

$$\underset{\sim}{F}(\underset{\sim}{p}) = \underset{\sim}{Q}(\underset{\sim}{\exp}\{\underset{\sim}{K}[\underset{\sim}{\ln}(\underset{\sim}{A}\underset{\sim}{p})]\}) = \underset{\sim}{g}$$

and compounded logarithmic functions of the type

$$\underset{\sim}{F}(\underset{\sim}{p}) = \underset{\sim}{L}\{\underset{\sim}{\ln}[\underset{\sim}{Q}(\underset{\sim}{\exp}\{\underset{\sim}{K}[\underset{\sim}{\ln}(\underset{\sim}{A}\underset{\sim}{p})]\})]\} = \underset{\sim}{h}$$

where $\underset{\sim}{Q}$ and $\underset{\sim}{L}$ are known matrices and $\underset{\sim}{\exp}$ transforms a vector to the corresponding vector of exponential functions (i.e., of anti-logarithms).

Forthofer and Koch (1973) illustrate this extension with four examples, two of which deal with problems in highway safety - relationship between car size and accident injuries for accompanied and for un-accompanied drivers

The present study further extends Forthofer and Koch (1973) to handle functions of the form

$$\underline{F}(\underline{p}) = \underline{\exp}\ (\underline{L}\{\underline{\ln}[\underline{Q}(\underline{\exp}\{\underline{K}[\underline{\ln}(\underline{A}\underline{p})]\})]\}) = \underline{k} = \frac{R_L}{R_u} \qquad (D.1)$$

the ratio of standardized injury rates for lap belted and unrestrained occupants respectively, for example. A consistent estimate for the covariance matrix of $\underline{F}(\underline{p})$ is given by

$$\mathrm{var}(\underline{F}(\underline{p})) = \underline{D}_{\underline{z}}\underline{L}\underline{D}_{\underline{g}}^{-1}\underline{Q}\underline{D}_{\underline{y}}\underline{K}\underline{D}_{\underline{a}}^{-1}\underline{A}[\underline{V}(\underline{p})]\underline{A}'\underline{D}_{\underline{a}}^{-1}\underline{K}'\underline{D}_{\underline{y}}\underline{Q}'\underline{D}_{\underline{g}}^{-1}\underline{L}'\underline{D}_{\underline{z}} \qquad (D.2)$$

where

$$\underline{y} = \underline{\exp}(\underline{f})$$

$$\underline{z} = \underline{\exp}(\underline{h})$$

Hypothesis testing and model fitting for this complex situation is carried out using a computer program for generalized categorical data models called GENCAT, which is an extension of the previous LINCAT and MODCAT programs developed by the Department of Biostatistics, University of North Carolina at Chapel Hill.