REPORT

CEM Report 4286-751

DEVELOPMENT OF A METHODOLOGY FOR ACCIDENT CAUSATION RESEARCH

Hans C. Joksch
Jim C. Knoop

Contract No.: DTNH22-80-C-07284
Contract Amount: $247,555

The Center for the Environment and Man, Inc.
Hartford, Connecticut 06120

June 1983

FINAL REPORT

Prepared for

U.S. DEPARTMENT OF TRANSPORTATION
National Highway Traffic Safety Administration
Washington, D. C. 20590

NOTICE

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle DEVELOPMENT OF A METHODOLOGY FOR ACCIDENT CAUSATION RESEARCH		5. Report Date June 1983
		6. Performing Organization Code
7. Author(s) Hans C. Joksch, Jim C. Knoop		8. Performing Organization Report No. CEM Report No. 4286-751
9. Performing Organization Name and Address THE CENTER FOR THE ENVIRONMENT AND MAN, INC. 275 Windsor Street Hartford, Connecticut 06120		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. DTNH22-80-C-07284
12. Sponsoring Agency Name and Address National Highway Traffic Safety Administration 400 Seventh Street, S.W. Washington, DC 20590		13. Type of Report and Period Covered Final Report
		14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

The objective of this study was to fully develop and apply a methodology to study accident causation, which was outlined in a previous study. "Causal" factors are those pre-crash factors, which are statistically related to the accident rate for specific types of accidents, in specific pre-crash situations. To calculate accident rates, exposure measures matched to the pre-crash situations are needed.

Data collection techniques to collect certain basic data by roadside observations were developed. A data collection design was developed, indicating when and where to collect data in Ulster and Schenectady Counties, New York. Data were collected, and exposure estimates for several pre-crash situations derived. Using police accident reports, accident rates were calculated for these pre-crash situations. The relations of these rates to various pre-crash factors and their interactions were studied.

The general methodology is described in a separate "Manual for Accident Causation Research."

17. Key Words		18. Distribution Statement Document is available to the U.S. public through the National Technical Information Service, Springfield, Virginia 22161.	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 183	22. Price

Form DOT F 1700.7 (8-72)

ACKNOWLEDGMENTS

Jim C. Knoop played a major role in planning and conducting much of this work. Other participants were John T. Ball, Stephen Thoren, Kayla Costenoble, Marjorie Northrop and Howard Slason. The field data were collected by KLD Associates, Inc. The New York Department of Motor Vehicles provided vehicle registration information. Ernst Meyer was NHTSA's contract technical manager.

METRIC CONVERSION FACTORS

Approximate Conversions to Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
in	inches	*2.5	centimeters	cm
ft	feet	30	centimeters	cm
yd	yards	0.9	meters	m
mi	miles	1.6	kilometers	km
		AREA		
in^2	square inches	6.5	square centimeters	cm^2
ft^2	square feet	0.09	square meters	m^2
yd^2	square yards	0.8	square meters	m^2
mi^2	square miles	2.6	square kilometers	km^2
	acres	0.4	hectares	ha
		MASS (weight)		
oz	ounces	28	grams	g
lb	pounds	0.45	kilograms	kg
	short tons (2000 lb)	0.9	tonnes	t
		VOLUME		
tsp	teaspoons	5	milliliters	ml
Tbsp	tablespoons	15	milliliters	ml
fl oz	fluid ounces	30	milliliters	ml
c	cups	0.24	liters	l
pt	pints	0.47	liters	l
qt	quarts	0.95	liters	l
gal	gallons	3.8	liters	l
ft^3	cubic feet	0.03	cubic meters	m^3
yd^3	cubic yards	0.76	cubic meters	m^3
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C

*1 in = 2.54 (exactly). For other exact conversions and more detailed tables, see NBS Misc. Publ. 286, Units of Weights and Measures, Price $2.25, SD Catalog No. C13.10:286.

Approximate Conversions from Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
mm	millimeters	0.04	inches	in
cm	centimeters	0.4	inches	in
m	meters	3.3	feet	ft
m	meters	1.1	yards	yd
km	kilometers	0.6	miles	mi
		AREA		
cm^2	square centimeters	0.16	square inches	in^2
m^2	square meters	1.2	square yards	yd^2
km^2	square kilometers	0.4	square miles	mi^2
ha	hectares (10,000 m^2)	2.5	acres	
		MASS (weight)		
g	grams	0.035	ounces	oz
kg	kilograms	2.2	pounds	lb
t	tonnes (1000 kg)	1.1	short tons	
		VOLUME		
ml	milliliters	0.03	fluid ounces	fl oz
l	liters	2.1	pints	pt
l	liters	1.06	quarts	qt
l	liters	0.26	gallons	gal
m^3	cubic meters	35	cubic feet	ft^3
m^3	cubic meters	1.3	cubic yards	yd^3
		TEMPERATURE (exact)		
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F

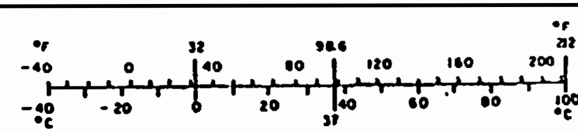


TABLE OF CONTENTS

Section	Title	Page
1.	INTRODUCTION	1
1.1	Objectives	1
1.2	The Demonstration Study	1
2.	EXPOSURE INFORMATION	3
2.1	Introduction	3
2.2	Exposure Measures	3
2.3	Exposure Data Collection Techniques	7
2.4	Exposure Data Collection Design	8
2.4.1	Overall Approach	8
2.4.2	Sampling of Locations	9
2.4.3	Sampling Over Time	11
2.4.4	Combining Time and Highway Strata	12
2.5	The Exposure Data Collection Plan	14
2.5.1	Selection of Observation Sites	14
2.5.2	Selection of Observation Dates	15
2.5.3	Combining Observation Times and Locations	16
2.6	The Exposure Data Collection Plan for Schenectady County	16
2.7	The Actual Data Collection	18
3.	ACCIDENT DATA	19
3.1	Accident Selection and Classification	19
3.2	Accident Data Collection and Coding	21
4.	ANALYSIS OF ULSTER COUNTY DATA	22
4.1	Overview	22
4.2	Estimating Exposure	23
4.2.1	Overview	23
4.2.2	Unbiased Estimators	23
4.2.3	Biased Estimators	25
4.2.4	Estimating the Variance	26
4.3	Overall Exposure and Accident Rates	29
4.3.1	Overall Exposure Estimates	29
4.3.2	Errors of Exposure Estimates	29
4.3.3	Overall Accident Rates	34
4.3.4	Disaggregation by Time of Day	35
4.4	Driver Characteristics	37
4.4.1	Driver Sex	37
4.4.2	Driver Age	40
4.5	Highway and Traffic Factors	44
4.5.1	Introduction	44
4.5.2	Studying the Highway Factors Separately	47
4.5.3	Interactions of Factors	54
4.6	Vehicle Factors	74
4.6.1	Introduction	74
4.6.2	Vehicle Age and Weight	75
4.6.3	Interactions with Car Age	79
4.6.4	Interactions with Car Weight	83
4.7	Exploratory Regression Analyses	87
4.8	Summary of Findings in Ulster County	91

Section	Title	Page
5.	VALIDATION WITH SCHENECTADY COUNTY DATA	93
5.1	The Approach	93
5.2	Overall Exposure Estimates and Accident Rates	94
5.3	Driver Characteristics	97
5.4	Highway Factors	98
5.5	Vehicle Characteristics	102
5.6	Summary of Validation	105
6.	SUMMARY AND CONCLUSIONS	107
6.1	Overview	107
6.2	The Methodology	107
6.3	Effects of Pre-crash Factors	109
APPENDIX A.	COLLECTION AND PROCESSING OF EXPOSURE DATA	A-1
APPENDIX B.	ACCIDENT DATA CODING	B-1

1. INTRODUCTION

1.1 Objectives

This study had two objectives:

1. To further develop and refine a methodology to study accident causation; and to write a manual for such causation studies.

2. To perform a causation study to demonstrate the feasibility of the methodology and to revise it as necessary.

The concept of causation used in this study was described in the Request for Proposal as follows:

> "...there is no known cause or factor or 'pre-crash' condition whose presence makes the occurrence or non-occurrence of an accident a certainty. It follows from this that for any combination of factors the occurrence of an accident is a matter of probability.
>
> Instead of primarily seeking to identify factors or clusters of factors that are somehow shown to be 'causally' related to accident occurrence, the main thrust of the investigation is directed towards determining the effect that a change in level of the various human, environmental and vehicular factors has on accident probability."

The overall methodology for studying accident causation as defined here was developed in a previous study.* Starting from that basis, the methodology was developed to a level of detail allowing its application. The experience from a small pilot study in the previous study and the demonstration study was used in developing the methodology. Also, studies reporting application of similar methods were used.

This report describes the demonstration study applying the methodology. The methodology itself is described in a separate "Manual for Accident Causation Studies."

1.2 The Demonstration Study

The demonstration study applied the methodology for studying accident causation described in the manual. Objectives of this application were to test the workability of the methodology, to modify it where necessary, and

---

*H.C. Joksch, "Development of an Accident Causation Methodology for NASS--Conceptual Approach," Appendix B of [M.L. Squires, R.D. Hume, Y. Hochberg, H.C. Joksch, J. Reidy, D. Zaidel, D. Shinar, and J. Treat], Accident Causation Methodology for the National Accident Sampling System. Institute for Research in Public Safety, Indiana University, Bloomington. July 1979.

to provide a basis of actual experience for future applications, as well as to obtain substantive information on accident causation.

The scope was defined by various factors. One was that the study originated in the context of the National Accident Sampling System (NASS) and was therefore conducted in NASS Primary Sampling Units (PSU)--the study itself was conducted in one PSU, the results were validated in another PSU. In NASS, detailed information ("level II") is collected for a sample of accidents in each PSU. Their number in a PSU is still too small for statistical studies. Therefore, police-investigated accidents with less detailed information (level I") had to be used.

Another limitation was the manpower available for field observations. Two alternative levels of effort had been proposed to NHTSA: one sufficient to cover highway and time strata adequately, the other allowing double coverage. The alternative selected allowed to make the necessary estimates, but the lack of double coverage precluded rigorous estimates of sampling errors.

The study was limited to accidents involving passenger cars, excluding those involving pedestrians and bicyclists.

Because exposure observations on Interstate Highways are much more difficult than on other highways, Interstate Highways were excluded from this study.

The overall approach was the following:

(1) An exposure data collection technique was developed and tested;

(2) An exposure data collection plan was developed;

(3) Exposure data were collected in Ulster County and Schenectady County, also accident data;

(4) For Ulster County, accident rates per exposure unit were derived from various pre-crash conditions, it was analyzed how they related to various pre-crash factors, and tentative relations re-established.

(5) For Schenectady County, accident rates were calculated for those pre-crash conditions characterized by the factors tentatively selected in Ulster County. It was determined which relations agreed in both counties.

(6) The experience from conducting the study was reviewed and incorporated into the methodology.

2. EXPOSURE INFORMATION

2.1 Introduction

The objective of the accident causation methodology developed is to determine how accident risk depends on pre-crash factors. The empirical estimates of accident risks being studied are accident involvement rates. These rates are for specific pre-crash conditions, characterized by certain pre-crash factors. This means that accidents must be categorized by pre-crash factors, and also that exposure must be measured separately for the various pre-crash conditions characterized by the selected factors.

To measure exposure under specific conditions, which also could be identified in accident reports as pre-crash conditions, vehicles, their drivers and their maneuvers, and highway and environmental factors were to be observed at sampled times and locations. From these observations, exposure estimates were to be derived.

The study area was a part of Ulster County, New York; including its central city of Kingston. Part of Ulster County is a NASS Primary Sampling Unit (PSU). Figures 2.1-1 and 2.1-2 show Ulster County and Kingston.

To validate the findings, accidents and exposure in Schenectady County, New York, excluding the city of Schenectady itself, were used. This is also a NASS PSU. Figure 2.1-3 illustrates Schenectady County.

2.2 Exposure Measures

The most commonly used exposure measure is vehicle miles of travel (VMT). Though it measures, in a very gross sense, the "quantity" of exposure to accident risk, it is not a suitable measure for many specific pre-crash situations. Entering an intersection, e.g., and performing a certain maneuver is a pre-crash situation carrying a certain accident risk. A natural measure of exposure to such pre-crash situations is a count of such maneuvers. VMT has, at best, an indirect relation to such an exposure if the number of intersection maneuvers is on the average, proportional to VMT. Similarly, one can find specific exposure measure for many types of pre-crash situations.

In our case, the level of effort available for collecting exposure data was limited. Therefore, only exposure measures requiring limited and relatively simple data could be used. Two measures were used: (1) VMT, and (2) intersection maneuvers. Four types of intersection maneuvers were considered:

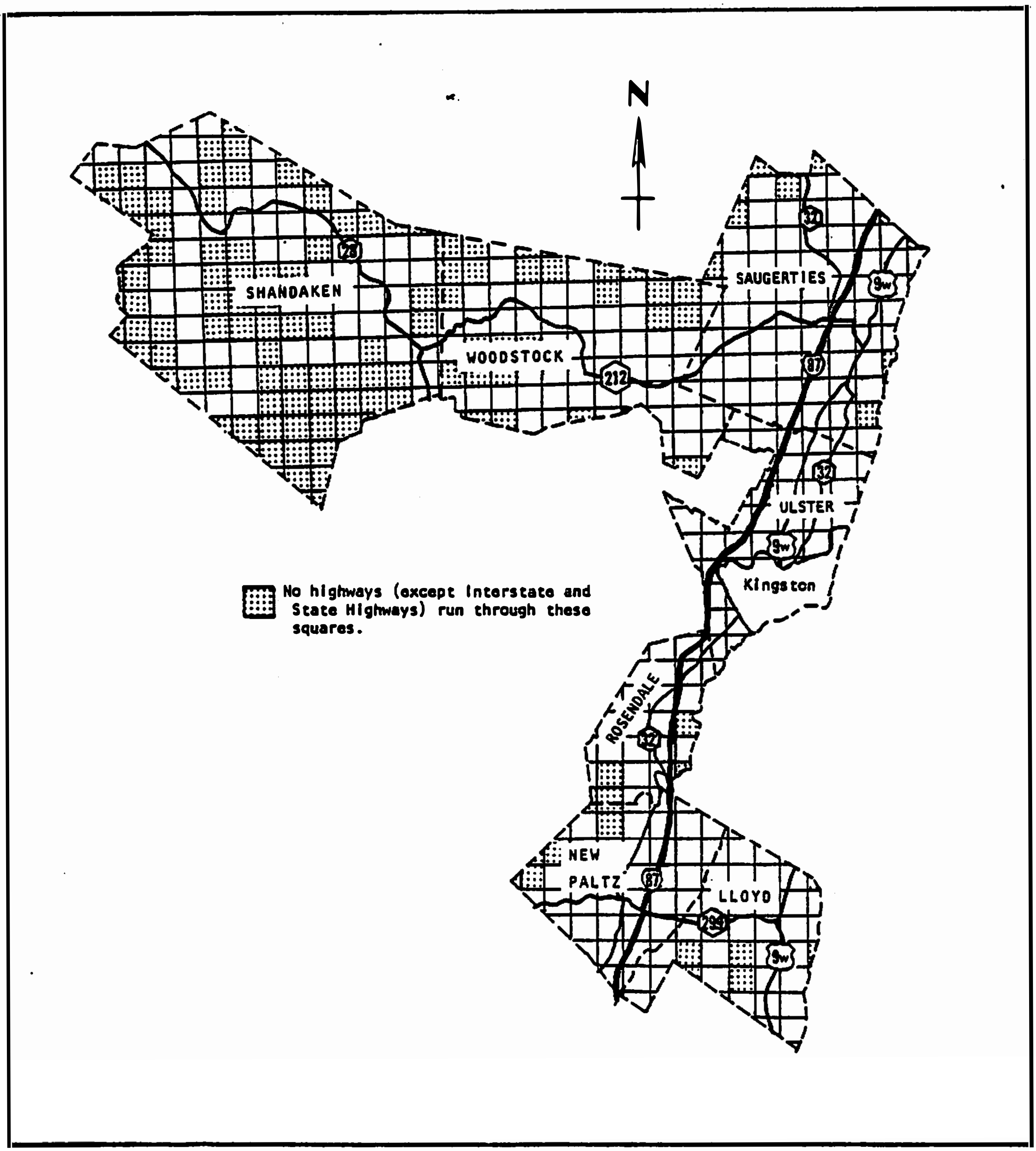


Figure 2.1-1. Study area in Ulster County. Shandaken was included in the original plan but not in the final plan. The squares comprise the grid used for sampling on non-state highways.

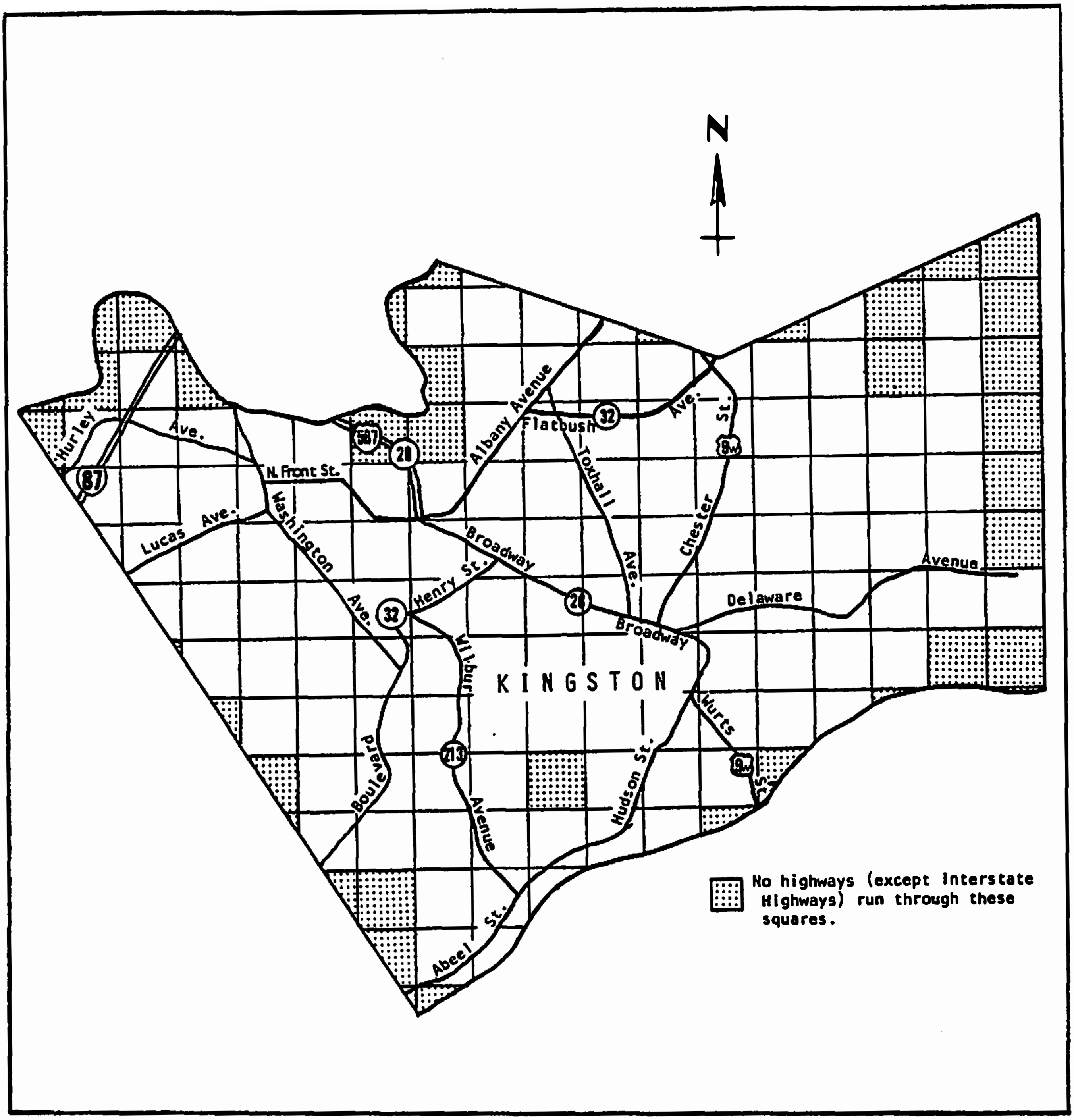


Figure 2.1-2. Kingston. The streets shown were considered "arteries." The squares represent the grid used for sampling other streets.

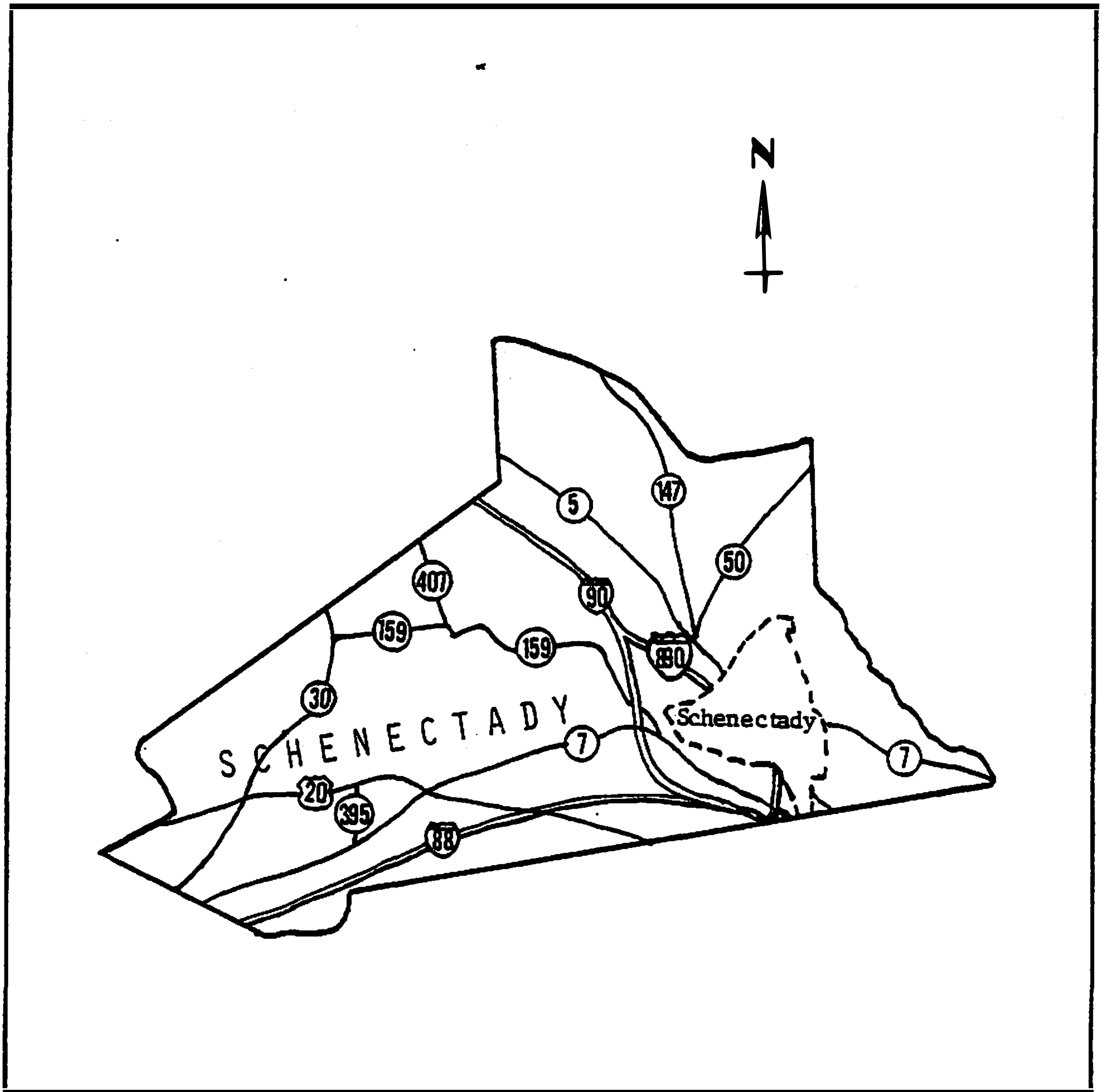


Figure 2.1-3. Schenectady County. The highways shown, with the exception of Interstates 88, 90 and 890 represent the study universe.

(1) going straight, (2) turning left, (3) turning right, and (4) "other" (mainly u-turns). In most cases we will use only the total number of intersection maneuvers as an exposure measure, in some we will use separate counts for the four types of maneuvers as measures of exposure to four different pre-crash situations.

VMT were used as exposure measure for all "segment" accidents (see Section 3.1 for definition), and also to calculate overall accident rates per VMT where needed for comparison with rates from other sources. VMT are not the preferred exposure measure for most types of accidents, but were used because data for more specific measures could not be collected.

2.3 Exposure Data Collection Techniques

It was decided to collect exposure information for pre-crash situations described by the following factors:

- driver:
  - age
  - sex
  - vehicle occupancy (not a driver characteristic, but similar with regard to data collection)
  - vehicle maneuver (at intersections)
  - compliance with "stop" signs (at intersections)
  - speed (at segment sites and uncontrolled intersection approaches)
- vehicle:
  - weight (as an indicator of size)
  - age
- highway:
  - number of lanes
  - grade
  - alignment
  - surface condition
  - traffic control (at intersections)
  - segment or intersection site, intersection configuration
- ambience:
  - light level
  - weather
- traffic
  - traffic volume
  - traffic mix

The following technique for data collection was developed. A team of two observers drove to the pre-determined observation site, parked the car on the roadside and set up equipment so that it was not visible to the traffic to be observed. A camera for photographing vehicles and their license plates, a radar for measuring speed, and tape recorders for recording spoken descriptions of visual observations were used. When the traffic was too dense, only a sample of all passing cars was observed. Appendix A describes the data collection procedure in detail.

Later in the study a third observer was used at intersections of a major and a minor road. He visually observed traffic from the minor road while the main team observed traffic at the major road.

At night, driver characteristics and vehicle license number were usually not obtainable. The "chase car" technique was tried: the observers followed the target vehicle to gain more time and pass perhaps a better illuminated (by roadside sources or other vehicles) area to observe the driver and read the license plate. This technique proved not useful: sometimes travel speeds were too high (above the speed limit), often the road was so curved or otherwise bad that one could not follow closely enough to observe the chased cars.

It was also considered to observe most factors at the sample site, and driver and license number farther downstream at an illuminated site. This was not practicable, because illuminated sites outside the city were usually so far away that it was not possible to reliably identify a vehicle at both sites.

2.4 Exposure Data Collection Design

2.4.1 Overall Approach

The purpose of the exposure data collection was not to obtain one overall measure of exposure for cars in Ulster County, or overall exposure measures for a few well-defined pre-crash conditions, but to obtain a representative sample of "all," or at least most pre-crash conditions in order to permit study of a wide range of combination of pre-crash factors. Therefore, the problem was not to design a sampling plan which yielded one or a few quantities with the least error, given a number of observations, but to design a plan which represented as many different pre-crash conditions as possible, even if that meant reducing the precision of aggregate measures of exposure.

Given the person-hours available for field data collection, travel time within Ulster County, setup times, it was estimated that observations could be taken at 140 "sites." Each site was either a location on a highway segment, or an intersection. At a segment site, one "setup" was made, at an intersection one for each approach. It was decided to use 70 segments, and 70 intersection sites. This resulted finally in 283 setups in Ulster County (most intersections have 3 or 4 approaches; some have only 1 or 2, due to one-way streets), due to the nature of the actually sampled intersection.

The following steps were performed:

(1) Select 140 combinations of locations and time periods (in generic terms) which represent most pre-crash conditions in Ulster County.

(2) Select specific sites, and specific dates for data collection.

(3) Combine sites and dates into a practicable observation plan.

2.4.2 Sampling of Locations

Pre-crash conditions, in terms of physical environment as well as in traffic characteristics, vary greatly among highway locations. Ideally, one would take a large random sample of locations, using a complete highway inventory as a sampling frame. In our case, the sample size was limited. Therefore, locations were stratified. Also, an inventory was only available for state highways.

For the stratification, at a first level Kingston and the rest of the area were distinguished, because Kingston is the only urban area. Within Kingston, arteries and local streets were distinguished. Outside of Kingston, state highways, collectors and local roads were distinguished. State highways outside of Kingston were further stratified according to volume (ADT < 5000, $5000 \leq$ ADT $< 10,000$, ADT $\leq 10,000$). Considering the extent of the five highway strata, initially the following sample sizes were selected:

- State highways 20 (high volume 6, medium 8, low 6)
- Collectors 30
- Local roads 60
- Kingston arteries 10
- Kingston local roads 20

When planning the study, it had been estimated that, on the average, 35 cars could be observed during a half-hour observation session. After one month of field observation it became clear that this average could not be achieved. On local rural roads, the average was only 8, and there were 12 sessions when no vehicle at all was observed. The average for all other highways was 25. Extrapolating from this basis, we estimated that, at most, 5,000 exposure observations could be collected under the original plan. This was less than half of what was expected. Therefore, the plan was radically modified. No further observations were made on local sites outside Kingston, and the observations rescheduled to other highway sites.

Table 2.4.2-1 shows the number of sites on the various highway strata in the actually realized sample. Also shown are the total highway miles, and number of intersections in the study area, for each stratum. Highway miles for state highways were obtained from the inventory, and the number of intersections on state highways by a complete count on maps. The number of highway miles and intersections for the other strata were estimated from the sample (see Section 2.5).

TABLE 2.4.2-1
DISTRIBUTION OF HIGHWAY MILES AND INTERSECTIONS IN THE STUDY AREA OF ULSTER COUNTY AND IN THE SAMPLE

Highway Class	Universe		Sample	
	Highway Miles	Intersections	Segment Sites	Intersection Sites
State highways outside Kingston				
AADT <5000	51.7	101	6	6
5000-10,000	29.9	90	8	10
AADT >10,000	14.5	51	10	8
Collectors	109	103	20	19
Local roads	343	646	11	10
Kingston				
Arteries	18	253	7	6
Other streets	88	343	10	10
Total	654	1,587	72	69

Intersection maneuvers have to be observed directly at the intersection. Vehicle miles of travel can be estimated from observations made at intersections, or somewhere on highway segments between intersections. However, speed was among the pre-crash factors to be observed. Speed at intersections, especially on controlled approaches, is not representative of all travel speeds, therefore, VMT were estimated from observations at sampled locations on segments, where representative speeds could be measured (note that some of these observations were close to intersections; this is proper if the measured speeds are to be representative for all VMT).

2.4.3 Sampling Over Time

Exposure observations in Ulster County were conducted from May 7 through October 2, 1980. Sampling frame was the entire time period May through September. Sampling unit was the "shift": the "early" shift was from 7:00 to 15:00 hours (7 a.m. to 3 p.m.), the "late" shift was from 15:00 to 23:00 hours (3 p.m. to 11 p.m.). Originally, 6:00 to 15:00 and 15:00 to 24:00 hours were planned, but traffic volume during the first and last hour were extremely low.

These months contain three periods with different traffic patterns: the pre-vacation period May to June 19, the vacation period June 20 through August, and the post-vacation period after Labor Day. These three periods (for brevity we will call them "seasons") were treated separately. To separate the vacation period is important, because the State University of New York in New Paltz has a large student population, the Catskill Park is in Ulster County, and vacation and summer weekend traffic to upstate New York passes through Ulster County.

Traffic and accidents have daily and weekly patterns. Monday through Thursday have essentially the same daily pattern; Friday morning is similar to Monday through Thursday morning. Friday during the afternoon traffic is somewhat higher than on other weekdays, and it is considerably higher during the evening and at night. The Saturday and Sunday patterns differ from each other and from that of weekdays.

Therefore, we stratified the time into three seasons, and within each season into five stratas:

- Monday through Friday early (7:00 - 15:00)
- Monday through Thursday late (15:00 - 23:00)
- Saturday, early
- Sunday, early
- Friday, late
- Saturday, late
- Sunday, late

Thus, a total of 21 time strata were used.

Initially, the following plan was made. Each of the seven periods of the week should be represented by the same number of observation sites. The first two seasons should have twice the number of observation sites of the third season, because observations in Schenectady County should begin in the third season, and also because it was desirable to make many observations early, to gain experience and modify the procedures, if necessary.

The result was that each period of the week should be represented by two observation periods during the first two seasons, and by one during the last season.

Several practical restrictions did not allow implementation of the schedule exactly as planned (see Section 2.7). Table 2.4.3-1 shows the distribution of the sample as actually implemented.

TABLE 2.4.3-1
DISTRIBUTION OF TIME PERIODS OVER STUDY PERIOD
IN ULSTER COUNTY, AND IN SAMPLE

Time Period		Days in† Study Period	Sample	
			Segments	Intersection
Pre-vacation	Mo-Fr	56	5	6
early shift	Sa	11	1	3
	Su	11	5	3
late shift	Mo-Th	45	7	6
	Fr	11	5	3
	Sa	11	4	3
	Su	11	5	3
Vacation	Mo-Fr	55	5	4
early shift	Sa	12	4	4
	Su	12	2	5
late shift	Mo-Th	44	4	4
	Fr	11	6	3
	Sa	12	5	3
	Su	12	3	4
Post-vacation	Mo-Fr	48	*	3
early shift	Sa	10	2	2
	Su	10	1	2
late shift	Mo-Th	38	3	1
	Fr	10	2	2
	Sa	10	2	3
	Su	10	1	2

*There was no segment observation in this time stratum. Therefore it was combined with Mo-Fr, day of the vacation period.
†Because accident data for some areas were not available for the first or last days of the study period, the number of days was correspondingly adjusted.

2.4.4 Combining Time and Highway Strata

With 7 highway and 21 time strata, 147 "cells" had to be covered with samples. Splitting the 140 observation sites evenly between segment and intersection samples (sites) gave only less than half the number necessary to represent all combinations of time and location factors. However, since 70 is greater than 7, and greater than 21, an approach developed by Bryant, Hartley and Jessen* can be used to stratify the sample. Because this approach was designed to estimate a population average (or total) but not averages (or totals) for strata or subpopulations, we modified it somewhat. We imposed the requirement that highway strata should be similarly represented in all three seasons, and also the periods of the week. We required also that within each season, early and late shift, weekday and weekend, Kingston and outside Kingston, state highway (or artery in Kingston) and other highways, were balanced as far as possible. This was not completely possible. A fairly complex ad hoc procedure of random sampling was developed which attempted to incorporate all requirements.

The actually implemented plan, however, had to be modified because of various practical restrictions described in Sections 2.5 and 2.6. It is shown in Table 2.4.4-1.

TABLE 2.4.4-1
EXPOSURE DATA COLLECTION DESIGN AS ACTUALLY IMPLEMENTED IN ULSTER COUNTY
(Upper left figures give the number of segment sites, lower right figures the number of intersections.)

Highway Stratum	Pre-vacation Period							Vacation Period							Post-vacation Period							Total
	Early			Late				Early			Late				Early			Late				
	M-F	Sa	Su	M-Th	F	Sa	Su	M-F	Sa	Su	M-Th	F	Sa	Su	M-F	Sa	Su	M-Th	F	Sa	Su	
1	/2						1/		1/	1/	1/	1/2					/1			1/	/1	6/6
2				/1	1/			2/1	/1	/2	1/		/1	1/2				2/	/1	1/1		8/10
3	2/		/1	2/1	1/	1/			1/1			2/		/1			/1	/1	1/1	/1		10/8
4	2/3	/1	2/	1/	1/1	/2	1/1	1/1	1/1	/2	1/1	2/	4/2	/1	/2	2/1		1/	1/			20/19
5		/1	2/1	4/2	1/1	2/1	2/1	/1		/1	/1											11/10
6					1/1		1/1	1/	/1	1/	/1		1/	2/	/1	/1						7/6
7	1/1	1/1	1/1	/2		1/		1/1	1/		1/1	1/1					1/			/1	1/1	10/10
Total	5/6	1/3	5/3	7/6	5/3	4/3	5/3	5/4	4/4	2/5	4/4	6/3	5/3	3/4	0/3	2/2	1/2	3/1	2/2	2/3	1/2	72/69

---
*Design and Estimation in Two-way Stratification, Journal of the American Statistical Association, 55, 1960, 105-124. A brief description can also be found in Section 5A.5 of W.G. Cochran, Sampling Techniques, 3rd Edition, Wiley 1977.

2.5 The Exposure Data Collection Plan

2.5.1 Selection of Observation Sites

In the previous section it was described how observation sites in terms of strata were selected. The next step was to select the actual locations.

To estimate intersection maneuvers, it is natural to sample intersections, and to expand from the sample of intersections to all intersections. To estimate VMT, two different approaches are obvious: (1) to sample "locations" on highways, or (2) to sample sections between intersections. In the first approach, one has to define a "location"--e.g., a 0.1- or 0.01-mile segment of highway. Thus, a highway system with a total length of L miles has 10L or 100L locations. From the number x of vehicles observed (in one direction) during an hour at a sampled location, one estimates 20Lx (or 200Lx) VMT for the entire highway system during this hour.

In the second approach, the highway system is considered as N segments between intersections. One segment is sampled, it has the length ℓ. If x vehicles are observed anywhere on this segment (in one direction) during one hour, the VMT on this segment are $2\ell x$, and on the entire highway system, $2N\ell x$. Both approaches have the same expected value for the VMT estimate. The second, however, is likely to have the greater standard error, because the estimate is affected by the variability among the segment lengths ℓ. Also, if segments are sampled, a short segment has the same probability of being selected as a long segment. This tends to concentrate the sample on areas with short segments: typically more densely populated areas, with higher traffic volumes and lower travel speed. Therefore, sampling of locations instead of sampling of segments was chosen.

For state highways, an inventory was obtained. Sections were stratified according to ADT, and each stratum sampled separately. For the selection of intersection sites, each intersection was numbered, and the required number of intersections selected by generating random numbers. For the selection of segment sites, cumulative mileage was defined for the sections of each, between, and locations selected by random numbers to one tenth of a mile.

For other than state highways, no inventory was available. Therefore, sampling was done in two steps. First, the entire study area was covered by a square grid with 1-mile distance between the lines (Kingston was treated separately; here a finer grid was used), and squares randomly selected without replacement. Then, within each selected square an inventory of collectors and local roads was produced.* Within each stratum, an intersection or a segment location was randomly selected by number or mileage (intersections were assigned to strata according to the highest classified road).

In addition to the required number of sites, additional sites were selected as a reserve.

An initial field inspection of the sites was performed to determine whether roadside observations were safely feasible; some had to be eliminated and replaced. Also, some roads turned out to be essentially access to one or only a few houses, or one commercial facility. It was decided to omit local roads which provided only access to one facility or fewer than three houses. In a few cases, locations identifiable on a map could not be found in nature.

2.5.2 Selection of Observation Dates

For the observation period May through September (for Ulster County) an inventory of available days was made, each day divided into the early and late shift, and each shift was assigned to the appropriate stratum. Holidays with long weekends were treated differently because traffic patterns differ strongly from those of ordinary weekends (though total accidents do not always differ much from other comparable periods). To schedule observations for such weekends would have allowed to estimate exposure for them, but reduced the possible number of observations under more common conditions. Not to observe on such days and also to exclude accidents on those days from the study would have reduced the already low number of accidents further. Therefore, the following compromise was made: the Monday or Friday of a long weekend was excluded from the sampling frames, because traffic is obviously not typical for a Monday or Friday, but Saturdays and Sundays were retained.

An observation schedule had to satisfy the following requirements: no more than 5 shifts per calendar week; no more than one shift on one day, and not a late shift one day and an early shift the following day. A desirable

*The following maps were used:

- State of New York, Office of Planning Coordination, Ulster County, November 1, 1976.
- New York State Department of Transportation, Federal-Aid Highway System 1980, Highway Classification, 1978 (based on quadrangle maps).
- Ulster County Highway Department, Map of Ulster County, New York 1981 (Copyright, The National Survey, 1981).
- Visual Encyclopedia, Ulster County, N.Y. (Copyright: Marshall Penn-York Co., Inc., 1979).

The last two, commercially produced maps, proved to be most up-to-date in details.

restraint was not to have two shifts on one weekend. For the post-vacation season, the schedule had also to be compatible with that for observations in Schenectady County.

Within these constraints, observation dates were randomly selected. The actual plan had again to deviate somewhat because of unforeseeable events.

2.5.3 Combining Observation Times and Locations

The result of the preceding steps was a list of dates and shifts, a list (and map) of observation sites, and a design which showed which dates could be combined with which location. Ideally, sites and locations would be combined randomly. In practice, a random allocation would have resulted in unacceptable travel times between sites. Therefore, the random selection was restricted to sites within an area which required only "acceptable" travel times between sites. As a consequence, frequently nearby sites on the same highway, or on intersecting highways were selected. Because traffic on such sites is to a large extent the same, the observations would not have been independent, and in effect, have reduced the sample size. Therefore, such nearby sites were excluded.

When assigning sites to dates in this manner, the sites remaining for the later dates became more scattered, and the problem of excessive travel times reappeared. Therefore, an *ad hoc* procedure was developed to reassign sites, until an overall satisfactory plan was developed.

As discussed in Section 2.4.1, the sample design had to be changed during the course of the study. The entire procedure had to be repeated with the remaining dates and locations. Because the initial observations showed that driver characteristics and license number were usually not recognizable at night, the following was done. Where for night observations lighted and unlighted sites were eligible, lighted sites were selected. This might have introduced a bias, but that appeared preferable to losing much information.

2.6 The Exposure Data Collection Plan for Schenectady County

The exposure data collection plan for Schenectady County was developed similarly as for Ulster County. However, it was slightly simpler. First, because of the short duration of two months in the post-vacation period, only seven strata for the seven periods of the week were needed. Because here the study was restricted to state highways, an inventory could be used to select the observation sites. However, because of the very great variation in traffic volume, five strata were used.

Table 2.6-1 shows the highway stratification, Table 2.6-2 the sampling design and the time stratification, as actually implemented in Schenectady County.

TABLE 2.6-1
STRATIFICATION OF STATE HIGHWAY IN SCHENECTADY COUNTY

				Sample	
Highway Stratum	ADT	Highway Miles	Number of Intersections	Segment Sites	Inter-sections
1	160-800	14.1	16	2	2
2	1050-4450	46.0	80	7	5
3	5150-9750	30.5	110	8	8
4	10700-14800	8.8	61	3	3
5	16500-29700	9.0	68	2	1
Total		108.4	335	22	19

TABLE 2.6-2
ACTUAL EXPOSURE SAMPLING DESIGN FOR SCHENECTADY COUNTY
(Upper left figures are segment sites, lower right figures, intersection sites)

Highway Stratum	Early Shift			Late Shift				Total
	Mo - Fr	Sa	Su	Mo - Th	Fr	Sa	Su	
1			2					2
			1	1				2
2	1	2	1	2	1			7
	2	2	1					5
3	1		1	1	3	2		8
	2		1	1	3		1	8
4				1	1		1	3
				1		1	1	3
5				1			1	2
						1		1
Total	2	2	4	5	5	2	2	22
	4	2	3	3	3	2	2	19
Actual Sept/Oct	44	9	8	35	9	9	8	

2.7 The Actual Data Collection

The actual implementation of the plans showed, that during an 8-hour shift, observations could be made at 7-8 "setups," 1/2-hour at each setup. A segment site requires one setup, an intersection 3 to 4, depending on the number of approaches. The rate at which passing passenger cars were sampled varied between 1 and 0.18. It depends on the traffic volume, but also on "platooning" of traffic. The upper limit of cars which can be observed during a 1/2 observation period is 70 to 80; the maximum was 98.

In Ulster County, 6331 exposure observations were made, in Schenectady 2008.

The main problem encountered in daytime observation was the very low traffic volume on local roads, which resulted in very unproductive observer hours. The change of the sampling design towards higher volume roads is not the ideal solution, because the number of accidents on local roads is not small. It is prefereable to use more efficient data collection techniques, or to exclude such accidents from the analysis.

Relatively minor problems in daytime observations were observing the driver on high-speed roads, to photograph license plates in dense traffic, especially with two lanes in the observed direction. Heavy rain can affect visual, photographic and radar observations.

At night, traffic volumes tend to become lower than during the day. License plates can, generally, not be photographed, though often visually read. Drivers and occupants can be observed only under very favorable conditions.

3. ACCIDENT DATA

3.1 Accident Selection and Classification

The origin of the study was to develop an accident causation methodology for NASS. Therefore, it was considered to use the detailed accident data collected by NASS. However, the number of NASS-investigated cases, including those a few years old, in Ulster and Schenectady Counties was too low to guarantee a successful analysis. Therefore, it was decided to use police investigated accidents which are much more numerous. The police accident reports contain information on the most obvious pre-crash factors.

Ideally, the accident and exposure data should be for the same time period and areas. However, the number of police investigated accidents during the exposure data collection period in the study area was considered too small. After reviewing some accident and exposure information, it was decided to use for Ulster County accident data from April 1980 through November 1980, and from April 1981 through October 1981. For one combination of pre-crash conditions, the adequacy of this procedure was tested (Section 4.5.1). Because the exposure observations in Schenectady County covered only a limited time period, and were to be less thoroughly analyzed, only concurrent accident data for September and October 1981 were used.

The study was limited to accidents involving passenger cars, excluding those involving pedestrians. Also, only cars registered in New York State were to be considered (the same as in the case of exposure).

For the analyses, accidents have to be classified on the basis of pre-crash factors. Some of the relevant pre-crash factors are vehicle maneuvers immediately preceding the crash. Because this information is often not or only sketchily available in police reports (though some are quite detailed about the pre-crash events), and because the number of accidents was too small to allow a fine classification, only a few classes of accidents were distinguished.

The main criterion was to distinguish accidents where the interaction of two or more vehicles was essential. Such accidents occur typically at intersections, when the paths of two vehicles cross at the same time. The same holds if one vehicle enters or leaves traffic from a driveway, etc. or makes a similar maneuver. However, whereas it was possible with the given level of effort to estimate the number of pre-crash maneuvers at intersections, it was not possible to estimate it at other, more dispersed sites. Therefore, only interaction accidents at intersections were considered as such (and are called "intersection accidents"). Accidents at intersections where the interaction of two vehicles was not essential, such as resulting from loss of control when turning, were not considered "intersection" accidents. Also, if an accident could have occurred also on a segment, such as a head-on collision when passing, or a rear-end collision, it was treated as a segment accident, not as an intersection accident.

In addition to classifying accidents as intersection accidents, we used a fine classification of accident involvements by maneuvers: going straight, turning left, turning right and "other" (primarily u-turns).

With this definition it is clear that "intersection" accidents result from an unsuccessful interaction of two vehicles' maneuvers. The remaining accidents which we call "segment" accidents are a less homogeneous category. First, there are the single-vehicle accidents, where vehicle driver and environmental characteristics are clearly the relevant pre-crash factors (and only rarely another "phantom" vehicle). Collisions between two (or more) vehicles, are, however, less clear. First, there are those cases where the interaction of two vehicles was definitely essential, such as turning into or off the road. This is usually clearly described in accident reports. However, because we could not estimate the specific exposure to these pre-crash conditions, we did not define a separate class of accidents, In other collision cases, the process can be so that the interaction of two vehicles is essential, or that the involvement of the second vehicle is just incidental. For instance, if a car continues to go straight in a right curve (for whatever reasons, which might be in the driver, the vehicle, or the highway) an accident will result with near certainty. Whether it will be a single-vehicle accident, or involve another vehicle, will depend on the traffic density. The same holds to various degrees for other collisions between cars. However, if passing another vehicle is involved, or in the case of rear-end collisions, one can expect with a high degree of confidence that the interaction of the vehicles was essential. In many cases, the police reports allow quite reliable conclusions, whether the accident was due to an interaction, or whether the other vehicle was only incidentally involved. However, because this is not always the case, and because we could not collect exposure to such pre-crash conditions as passing another vehicle, or following another vehicle with a short headway, we could not make this classification.

The consequence is that the class of "segment" accidents is too aggregated to be quite satisfactory. Even the further classification into single vehicle, head-on, rear-end, and other is crude and not quite satisfactory.

Though accidents are reported, the unit of the analysis is the accident involved vehicle. The reason is that no promising exposure measure is known which corresponds directly to the pre-accident condition (considering this as a unit), whereas various plausible measures can quantify the exposure of a vehicle to certain pre-crash conditions.

3.2 Accident Data Collection and Coding

From the police agencies in the study area (Section 2.1) copies of accident reports for the periods April 1980 through November 1980, and April 1981 through October 1981 were obtained. In some cases, data for a few days at the beginning or end of the period were missing. All reports were on the form MV-104A (Appendix B).

Accidents were eligible if they occurred in the study period between 7:00 and 23:00 hours, if at least one vehicle was a passenger car registered in New York, and if no pedestrian or bicyclist was involved. Information from the accident form was coded as described in Appendix B. In addition to the information given in structured form, the information from the narrative description of the accident, and from the diagram was coded.

For Ulster County, 1639 accidents involving 2383 New York cars were eligible; for Schenectady, 184 with 284 New York cars.

4. ANALYSIS OF ULSTER COUNTY DATA

4.1 Overview

In this section, accident and exposure data for Ulster County are studied to estimate accident involvement rates, and to identify pre-crash factors and their combinations which influence accident rates.

Two measures of exposure are used: Vehicle miles of travel (VMT) and intersection maneuvers. In some cases the following maneuvers were distinguished and counted separately: going straight through an intersection, turning left, turning right, and "other" (mainly u-turns). Accident involvements were classified accordingly: "segment" accident for which VMT were used as exposure measure, and "intersection"* accidents, where maneuvers were used as exposure measure. In some cases segment accidents were further classified into "single vehicle," "head-on," "rear-end," and "others." Intersection accident involvements were classified by "going straight," "turning left," "turning right," and "other" maneuvers.

First, overall exposure is estimated, the errors of the estimates assessed, and the balance of the actually implemented sampling--which differed somewhat from the original plan--examined.

Then, accident involvement rates for specific pre-crash conditions were determined and analyzed. Factors considered were driver age, driver sex, car age, car weight, highway alignment, grade, and surface (dry/wet). Since they were not known for accident cases, speed and traffic volume were used only as "covariates." These analyses used rates for discrete classes of pre-crash conditions, calculated from tabulations of accidents, and of exposure.

Finally, some exploratory regression analyses were performed. These analyses do not require that accident involvement rates be explicitly calculated, and they allow to use continuous pre-crash factors, without converting them into discrete categories. Only exploratory analyses were performed, because the sparse distribution of the observation sites over time and space did not match the fine "resolution" of which this approach is capable.

---

*The definition differs somewhat from the conventional one; see Section 3.1.

4.2 Estimating Exposure

4.2.1 Overview

In this section, the formulas used to estimate exposure from the original observations are derived. These formulas are based on the sampling design as described in Section 2. First, unbiased estimators are given. These estimators could have a larger variance than certain biased estimators. Therefore, also biased estimators are derived (actually, the differences between the resulting estimates are completely negligible). Finally, a formula for the error of the estimates is derived.

4.2.2 Unbiased Estimators

The observations were stratified by time, and by highway class. Two sampling frames were used for sampling locations: for the three strata on state highways an inventory was used, for the other four strata grids on a map in a first stage, and in a second stage an inventory which was developed for each square which was selected from the grid.

Let the index i denote the highway stratum, and j the time stratum ($i=1...7$; $j=1...21$). For each time stratum, the universe of the study period contains T_j hours, for highway stratum, $i=1...d$, S_i miles (counting each highway mile twice, because traffic moves in both directions, but only one direction was observed; one-way streets were neglected). For each stratum $i=4...7$ the universe consisted of N_i grid squares containing highways of this stratum,* and in each grid-square h, ℓ_h highway miles. When intersections are studied, S_i and ℓ_h have to be interpreted as the number of intersections in stratum i and grid square h, respectively.

Within a "cell" (i,j), several observation sites may have been selected. If so, they are numbered $k=1,2...$ within each cell. We assume that at site (i,j,k) during a time period f_{ijk}, x_{ijk} cars were observed, which were sampled at a rate r_{ijk} from all cars passing the site. To extrapolate from the cars observed to all cars passing the site during one hour, one calculates

$$z_{ijk} = \frac{x_{ijk}}{f_{ijk} r_{ijk}} \qquad (4.2.2\text{-}1)$$

*The selection of squares was modified so as to exclude those which did not contain highways of the stratum.

If we are in one of the strata $i=4...7$, we extrapolate to exposure in the entire square by multiplying the number of cars per hour with the total highway miles ℓ_h in that square. Because the square is completely identified by i,j,k we replace the index h by i,j,k.

$$\frac{\ell_{ijk}}{f_{ijk}a_{ijk}} \; x_{ijk} \qquad (4.2.2\text{-}2)$$

is the estimate of exposure for the square i,j,k.

For a given highway class i and time stratum j, the "cell" of the sampling plan contains S_iT_j highway miles, or N_iT_j squares.

Therefore, an observation site gives the estimates

$$S_iT_jz_{ijk} \qquad (4.2.2\text{-}3)$$

or

$$N_iT_j\ell_{ijk}z_{ijk} \qquad (4.2.2\text{-}4)$$

for total exposure in this cell. Since locations and times were selected independently (at least in principle), the probability that with n_i observation sites on highway class i and t_j in time stratum j, one out of n sites is in i,j is

$$p_{ij} = \frac{n_i}{n} \frac{t_j}{n} . \qquad (4.2.2\text{-}5)$$

Therefore, $1/p_{ij} = n^2/n_it_j$ is the expansion factor from one observation site to the total universe. The resulting exposure estimates are

$$n^2 \frac{S_iT_j}{n_it_j} z_{ijk} \qquad (4.2.2\text{-}6)$$

and

$$n^2 \frac{N_iT_j}{n_it_j} \ell_{ijk} z_{ijk} . \qquad (4.2.2\text{-}7)$$

Since there were n observation sites, one combines the estimates from all sites, and averages them, dividing the sum by n:

$$Z = n \sum_{\substack{i=1,2,3 \\ j,k}} \frac{S_iT_j}{n_it_j} z_{ijk} + n \sum_{\substack{j=4-7 \\ j,k}} \frac{N_iT_j}{n_it_j} \ell_{ijk}z_{ijk} . \qquad (4.2.2\text{-}8)$$

If one is dealing with intersections, one has to note that z_{ijk} counts the maneuvers on all approaches to the intersection i,j,k, because intersections, and not approaches were sampled. Thus, if an intersection has four approaches, the observations from four 1/2-hour observation periods are combined, if it has three approaches, only those from three 1/2-hour periods.

Equation (4.2.2-8) gives an estimate for total exposure in the study universe. Our problems require estimates of disaggregated exposure. For instance, we may be interested in the exposure of male drivers in cars of a certain weight class on wet road surfaces. In that case, we have to calculate the z_{ijk} only from those observed vehicles which fall into the category of interest. The easiest way to do this is to attach to each individual observation of a car a weight (expansion factor)

$$n \frac{S_i T_j}{n_i t_j} \frac{1}{f_{ijk} a_{ijk}} \qquad (4.2.2\text{-}9)$$

or

$$n \frac{N_i T_j}{n_i t_j} \frac{\ell_{ijk}}{f_{ijk} a_{ijk}} , \qquad (4.2.2\text{-}10)$$

and count the cars falling into the class of interest, weighting each car with the expansion factor.

These unbiased expansion factors have some disadvantages. Precise error estimates are very cumbersome. Also, analysis for simple cases suggest that the standard errors are very likely larger than those of the biased estimator described in Section 4.2.3, because the ℓ_{ijk} vary among the squares of the grid which contributes to the total variance of Z.

4.2.3 Biased Estimators

The disadvantages of the unbiased estimator are due to the factor ℓ_{ijk} in (4.2.2-7). Actually, the product $N_i \ell_{ijk}$ is an estimate of total mileage S_i in highway stratum i. In the unbiased estimator, each observation site uses its own mileage estimate $S_{ijk} = N_i \ell_{ijk}$. If one uses one common estimate

$$S_i = \frac{1}{n_i} \sum_{j,k} \ell_{ijk} \qquad (4.2.3\text{-}1)$$

one obtains a biased estimator, but the bias may have any sign, and its variance is likely to be lower than that of the unbiased estimator. In our case, there was an additional advantage that in most strata more than n_i grid squares were inventoried (to obtain alternate sites). Therefore, a more precise estimate of S_i could be used than given by (4.2.3-1).

Using the estimated S_i for strata 4...7, (4.2.2-8) simplifies to

$$Z = n \sum_{i,j,k} \frac{S_i T_j}{n_i t_j} z_{ijk} \quad (4.2.3\text{-}2)$$

With this estimator, we can establish a relation with the method for two-way stratification developed by Bryant, Hartley and Jessen (see Section 2.4.4) and use their variance estimator. With our notation and considering that we are interested in total exposure in the study universe, not an average per hour, etc., their estimator takes exactly the form (4.2.3-2).

If one uses observations of individual cars for estimating exposure to selected pre-crash situations, the expansion factor for each car becomes

$$n \frac{S_i T_j}{n_i t_j} \frac{1}{f_{ijk} r_{ijk}} \quad (4.2.3\text{-}3)$$

4.2.4 Estimating the Variance

A rigorous estimate of the variance of total exposure requires that each cell (i,j) contains at least two observation sites. In our case, the number of observation sites is insufficient. However, Bryant, Hartley and Jessen (op. cit.) have developed a method to estimate the variance with a smaller sample, as long as there are at least two observations in each highway stratum, and in each time stratum, and one can assume that the within-cell variances are equal.

Using our notation, their formula (33) becomes

$$\mathrm{var}(Z) = \frac{n-1}{n}\left[\sum_i \frac{n_i}{n_i-1} \sum_j n_{ij}(V_{ij} - \overline{V}_{i.})^2\right. \tag{4.2.4-2}$$

$$+ \sum_j \frac{t_j}{t_j-1} \sum_i n_{ij} (V_{ij} - \overline{V}_{.j})^2$$

$$\left. - \frac{n}{n-1} \sum_{ij} n_{ij}(V_{ij} - \overline{V})^2 \right]$$

$$+ \frac{s^2}{n} \sum_{ij} F_{ij}^2(n_i t_j - f_{ij}) \quad ,$$

where

$$F_{ij} = n \frac{s_i T_j}{n_i t_j} \tag{4.2.4-3}$$

$$V_{ij} = \frac{F_{ij}}{n_{ij}} \sum_k z_{ijk} \tag{4.2.4-4}$$

$$\overline{V}_{i.} = \frac{1}{n_i} \sum_j n_{ij} V_{ij} \tag{4.2.4-5}$$

$$\overline{V}_{.j} = \frac{1}{t_j} \sum_i n_{ij} V_{ij} \tag{4.2.4-6}$$

$$\overline{V} = \frac{1}{n} \sum_{i,j} n_{ij} V_{ij} \tag{4.2.4-7}$$

$$f_{i,j} = (n-1)\left[1 + \frac{n-n_i t_j}{n}\left(\frac{1}{n_i-1} + \frac{1}{t_j-1} - \frac{1}{n-1}\right)\right] \tag{4.2.4-8}$$

and n_{ij} is the number of observation sites in cell (i,j).

$$s^2 = \frac{\sum_{ijk}^1 (z_{ijk} - \overline{z}_{ij})^2}{\sum_{ij}^1 n_{ij} - n^1} \quad , \tag{4.2.4-9}$$

where sums are to be taken only over cells with two or more sites, and n^1 is the number of cells with two or more sites.

The first sum in (4.2.4-2) is essentially the variance of the estimates within rows, controlling for differences among rows. This is the variance due to variability over time if one stratifies the observation sites by highway class. The second sum is the variance within columns, controlling for differences among columns. It is the variance one would obtain if one stratified by time, but not by highway class. The third sum is the variance among the cell estimates without stratification. Together, the three sums give the variance of the estimate using space and time stratification, as it is due to the between-cell variability.

The within-cell variability is considered in the last term. Bryant, _et al._ state that it is likely to be small in comparison to the first term; computation of s^2/n would indicate whether this is the case.

A closer look at the term

$$n_i t_i - (n-1)\left[1 + \frac{n-n_i t_j}{n}\left(\frac{1}{n_j-1} + \frac{1}{t_j-1} - \frac{1}{n-1}\right)\right] \qquad (4.2.4\text{-}10)$$

shows the following. If one has a table with m rows and columns, exactly 2 observations in each row and column, resulting in a total of 2m observations, each of the terms (4.2.4-10) becomes

$$16 - 6(m-1/m) \ . \qquad (4.2.4\text{-}11)$$

For $m \geq 3$, this is negative. For a tableau of m rows and columns, where each is covered with m observations, resulting in a total of m^2 observations, each of these terms becomes 1. With the n, n_i and t_j of our sampling design, some of these terms are positive, some are negative. If the sum with which s^2/n is multiplied becomes negative, the product can no longer be the contribution of the within-cell variance to the total variance of the estimate. Whether this is due to an approximation used by Bryant, et al. (omission of the finite-sample correction) or whether more serious problems are behind it was not further pursued.

4.3 Overall Exposure and Accident Rates

4.3.1 Overall Exposure Estimates

Using the unbiased expansion factors, we obtain a total of 413.4 x 10^6 VMT for the study period. With the biased expansion factors, one obtains 413.6 x 10^6. An independent estimate of VMT can be obtained from intersection observations*: 458 x 10^6 VMT. This is 11% more.

The estimate of intersection maneuvers is 1748 x 10^6, using unbiased, and 1738 x 10^6 using the biased expansion factors.

These figures show that there are 4.2 intersection maneuvers for each VMT. We estimated 1587 intersections, and 654 highway miles in the study area, that gives 2.4 intersections per highway mile. The difference between 4.2 and 2.4 is explained as follows. The number of intersections per highway mile varies widely: from a low of 1.9 on local roads outside Kingston to 14 for arteries within Kingston (Table 2.4.2-1). Since most travel occurs on highways with relatively many intersections per mile, the average number of intersection maneuvers per VMT is greater than the average number of intersections per highway mile.

4.3.2 Errors of Exposure Estimates

In order to make a rigorous estimate of the error of the exposure estimate, one needs a sampling plan with at least two observation sites within each time-stratum x highway-stratum cell. The limited number of observation sites did not allow such a plan. In Section 4.2.4 an approach is described which allows to estimate the error as long as at least two observation sites are in each time stratum, and in each highway stratum, if certain assumptions can be made. In our case, not all time strata contain at least two observation sites. One contained none, and was therefore combined with another, similar stratum, three contained only one observation site. They could be used for estimating exposure without aggregating them with other strata. For estimating the error of the exposure, we used a heuristic device.

Table 4.3.2-1 shows the passenger cars per hour at the observation sites which were used for estimating the error of exposure. These figures are rounded,

*A very simple estimation procedure was used: average cars per hour for the entire study universe were estimated and multiplied with total highway miles.

and the calculations were performed with limited precision. Therefore, the resulting estimate of 435×10^6 VMT is 5% higher than that derived from the individual observations, using exact expansion factors.

TABLE 4.3.2-1
PASSENGER CARS PER HOUR (IN THE OBSERVED DIRECTION)
AT THE SEGMENT OBSERVATION SITES, BY TIME AND HIGHWAY STRATUM

Highway Stratum	Pre-vacation Period							Vacation Period							Post-vacation Period						
	Early			Late				Early			Late				Early			Late			
	M-F	Sa	Su	M-Th	F	Sa	Su	M-F	Sa	Su	M-Th	F	Sa	Su	M-F	Sa	Su	M-Th	F	Sa	Su
1							90		58	42	154	296								34	
2					186			120 225			222			96				336 155		224	
3	403 290			269 168	560	225			622			392 444							534		
4	32 4		6 17	24	32		14	20	12		14	200 120	12 26 88 12			8 8		37	34		
5			26 4	2 10 18 17	64	16 6	6 23														
6					88		96	200		29			56	60 140							
7	64	2	0			74		2	12		32	48					0				18

The first term in equation (4.2.4-2) for the variance of the exposure estimate has three parts: one expressing the variance of the estimates among the columns within rows, the other the variance among the rows within columns, and the third the variance of the estimates among the cells. The corresponding standard errors are

within columns	54×10^6 VMT
within rows	57×10^6 VMT
among cells	69×10^6 VMT.

Comparing these three standard errors shows that stratification by time and by space stratum separately reduced the variance of the estimates about equally (a close inspection of the figures in Table 4.3.2-1 confirms that they vary less within rows, and within columns than over the entire array.) The effect of using time and space stratifications simultaneously is expressed by the first term in (4.2.2-4).

We have to consider that when estimating the variance among the rows from within the column variations, three observation sites in three columns which contain only one site each could not be used. We argue heuristically that each of them would have contributed similarly to the total variance as each of the 69 sites used, if they had been in columns with more than one site. Therefore, we increased the within-column variance by a factor of 72/69. The resulting standard error of the total VMT estimate is

$$39 \times 10^6 = \sqrt{\left(\frac{72}{69} \times 54^2 + 57^2 - 69^2\right)} \times 10^6 .$$

This is an error of 9%.

The second term in (4.2.4-2) considers the effect of the within-cell variance. The first factor--s^2/n--corresponds to a standard error of only 0.6 x 10^6 VMT. The sum with which it is multiplied, however, is negative. The implications of this are not clear, but it should caution against accepting the error estimate at face value.

Table 4.3.2-2 shows the numbers of cars approaching intersections per hour.

TABLE 4.3.2-2
PASSENGER CARS PER HOUR AT THE INTERSECTION SITES
(FROM ALL APPROACHES) BY TIME AND HIGHWAY STRATUM

Highway Stratum	Pre-vacation Period							Vacation Period							Post-vacation Period						
	Early			Late				Early			Late				Early			Late			
	M-F	Sa	Su	M-Th	F	Sa	Su	M-F	Sa	Su	M-Th	F	Sa	Su	M-F	Sa	Su	M-Th	F	Sa	Su
1	381 108											166 273					141				150
2				359				546	256	231 468			393	351 739					550		425
3			114	218					594					599			562	426	1152	539	
4	99 137 111	26			84	157 90	16	33	178	96 108	54		62 99	122	130 92	20					
5		8	85	15 60	36	34	120	11		18	26										
6					1104		490		982		180				172	297					
7	284	46	106					106			116	30								10	51

The total number of maneuvers estimated from these rounded figures is 1840×10^6 maneuvers, again, 5% higher than obtained from the individual observations. The standard errors corresponding to the three components of variance are

within columns	319×10^6 maneuvers
within rows	203×10^6 maneuvers
among cells	322×10^6 manuevers

Comparing these values shows that the variability within columns is essentially the same as among the cells, indicating that stratifying by time reduces the error of the estimate little, whereas stratifying by highway class reduces it considerably. The estimate of the error considering the effect of the two-way stratification is

$$202 \times 10^6 = \sqrt{\left(\frac{69}{68} \times 319^2 + 203^2 - 322^2\right)} \times 10^6 \quad .$$

This is 11% of the estimated maneuvers. This standard error is essentially the same one would have obtained if one had stratified by highway class only, and not by time and highway class.

The first factor--s^2/n--of the second term of (4.2.4-2) corresponds to a standard error of 1.6×10^6 maneuvers, which is small compared with the first term.

These two estimates indicate a standard error of about 10% for the overall exposure measures. The comparison of VMT estimates obtained from segment observations with those obtained from intersection observations (Section 4.3.1) confirms this. However, one should be cautious, because the observation plan was not strictly random, and therefore did not satisfy the assumptions under which Bryant, et al. derived the estimator. Also, comparing the traffic volume figures within cells with more than one suggests that the assumption of constant within-cell variance may not be satisfied.

The observation that stratifying by time and space reduced the variance compared with that using only space stratification for the VMT estimate, but not for the intersection maneuver estimate is noteworthy. It may be just due to imperfections in the sampling plan, but it could also indicate that travel in areas with many intersections varies less with time than travel in areas with relatively few intersections.

The standard error of 10% applies to overall exposure estimates. If one estimates exposure for pre-crash situations which vary only among observation sites, and/or times, but not within each observation site time "cell," e.g., curvature, the error will be larger. On the other hand, if one estimates exposure disaggregated by pre-crash factors which can vary within "cells," such as driver and vehicle characteristics (and also maneuvers) the error can be much smaller. We will show this by estimating the error of the proportion of male and female drivers. The proportion of male and female drivers was not readily available in the detail of Table 4.3.2-1. However, Table 4.3.2-3 shows the VMT by male and female drivers in a similar but more aggregated manner.

TABLE 4.3.2-3
DISTRIBUTION OF VMT BY DRIVER SEX, BY HIGHWAY CLASS AND TIME STRATUM
(The upper figures are VMT in 1000, the lower figures are percent of those VMT with known driver sex)

Time of Day:	7:00 - 19:00 Hours						19:00 - 23:00 Hours					
Weekday:	Mo-Fr			Sa,Su			Mo-Th			Fr-Su		
Highway Stratum	Male	Female	Unkn.	Male	Female	Unkn.	Male	Female	Unkn.	Male	Female	Unkn.
KINGSTON												
Arteries	15,910	5,920	740	516	516	0	--	--	--	656	164	19,812
%	73	27		50	50					80	20	
Streets	8,618	6,672	0	2,918	2,340	0	0	0	7,124	94	0	5,328
%	56	44		55	45					100		
OUTSIDE KINGSTON												
State Highways	123,014	74,396	11,984	20,182	9,030	4,968	56	112	8,948	1,400	476	6,316
%	62	38		69	31		33	67		75	25	
OTHER HIGHWAYS	18,410	15,896	2,952	14,456	4,206	0	0	0	8,028	0	0	11,636
%	54	46		77	23							

Since driver sex was either not at all observable at night, or missing in a very high percentage of cases, day and night exposure will be treated separately. Using the eight values for the percentage of male drivers during the day, ranging from 50% to 77% and weighting them with the VMT for which sex was given, one obtains a standard error of 5.9% for each of these eight values, and $2.2\% = 5.9\%/\sqrt{7}$ for the average of 63%. From the eight values available at night, one obtains a standard error of 11.3% for each individual value, or $6.5\% = 11.3\%/\sqrt{3}$ for the average of 75%. Thus, though the absolute mileage of male and female drivers will have a similar standard error as total VMT in the study universe, about 10%, the relative frequencies of male and female drivers is known with much higher accuracy, namely 63 ± 2 for the percentage of male drivers during the day, 37 ± 2 for that of female drivers. These are errors of 3% and 5%, respectively. The same holds for other pre-crash factors which can vary within the observation period at one site.

4.3.3 Overall Accident Rates

Table 4.3.3-1 shows the numbers of involvements for the classes of accidents distinguished in this study, the estimates of the corresponding exposure measures, and the resulting accident rates.

TABLE 4.3.3-1
ACCIDENT INVOLVEMENTS, EXPOSURE AND ACCIDENT INVOLVEMENT RATES

	Accident involvements	Exposure	Accident Involvement rates (per 10^6)
Segment accidents		VMT(10^6)	
Single	362		0.9
Head-on	374		0.9
Rear-end	202		0.5
Other	769		1.8
Total	1,707	413.4	4.2
Intersection accidents		Maneuvers (10^6)	
Going straight	404	1267	0.32
Turning left	205	165	1.25
Turning right	54	272	0.20
Other	13	44	0.30
Total	676	1748	0.38

To compare this with conventional rates of accidents per VMT, we have to combine the involvements for the two classes of accidents to a total of 2383. With overall VMT of 414 x 10^6, the rate is 6 accident involvements per 10^6 VMT. NHTSA's Report on Traffic Accidents and Injuries for 1979-80, reports an annual average of 11.6 million vehicles involved in accidents and 1.4 x 10^{12} VMT for all vehicles; this gives 8 involvements per 10^6 VMT. The order of magnitude is the same but the Ulster County rate is only 75% of the national rate. One can not expect exact agreement, because one is for passenger cars, the other for all motor vehicles. Also, there are regional differences in accident reporting and in driving conditions.

Detailed VMT and accident involvement data are available for North Carolina.*

*J.R. Stewart, C. Lederhaus Carroll, Annual Mileage Comparisons and Accident and Injury Rates by Make, Model. Highway Safety Research Institute, University of North Carolina, October 1980.

They give an accident involvement rate for passenger cars of 5.1 per 10^6 VMT, which is comparable to the Ulster County rate. However, the North Carolina rate includes pedestrian accidents. The North Carolina rate is composed of 0.75 for single-vehicle accidents, and 4.4 for multi-vehicle accident involvements. In Ulster County, there are 362 single-car accident involvements per 414 VMT, which is 0.9 per 10^6 VMT, and 2021 multi-vehicle accident involvements, which gives an involvement rate of 4.9. The agreement between North Carolina and Ulster County is good.

Comparing the overall rates for the various types of segment accidents yields relatively little information because all are based on the same exposure measure, and differences in rates reflect only differences in the accident numbers. It is more informative to compare the rates for the various types of intersection accidents. Left turns have the highest risk, right turns the lowest, going straight through an intersection and "other" maneuvers (e.g., u-turns) have about the same risk.

4.3.4 Disaggregation by Time of Day

The experimental design stratified time by "season," day of the week, and two shifts in a day. Within each shift, observations were scheduled sequentially; their actual spacing was the result of travel and setup times. To each shift, strata of locations were assigned according to the data collection design. Actual locations were randomly selected as far as practical; however, often locations were rejected if travel time would have been excessive, and another random selection made.

During the course of the observations we found that at night, drivers could be observed only under very favorable conditions; also license plates could often not be read. In order to reduce potential biases due to cases with missing data being concentrated at night, we treated, in some analyses, day and night observations separately.

This raises the question, whether expansion factors which were designed to estimate exposure for the time period 7-23 hours do correctly estimate separate exposures for the periods 7-19, and 19-23 hours (approximately the periods of daylight and of darkness) when applied separately to the observations during these time periods. To test this, we divided the day into shorter periods. An examination of the observation times showed that the finest practical breakdown was into the periods 7-11, 11-15, 15-19, and 19-23 hours.

First, we simply applied the expansion factors separately to the observations during each of these time periods. The results are shown in Table 4.3.4-1.

TABLE 4.3.4-1
ESTIMATES OF SEGMENT AND INTERSECTION EXPOSURE FOR
FOUR PARTS OF THE DAY, USING DIFFERENT EXPANSION FACTORS

	7-11	11-15	15-19	19-23	Total
Overall expansion factors - Unbiased					
VMT (10^6)	84	62	194	74	414
Intersection maneuver (10^6)	150	800	350	448	1748
Maneuver/VMT	1.8	12.9	1.8	6.1	4.2
	6.5		3.0		
Biased					
VMT (10^6)	88	60	194	70	414
Intersection maneuver (10^6)	166	646	300	628	1738
Maneuver/VMT	1.9	10.8	1.5	8.9	4.2
	5.5		3.5		
Specific expansion factors - Unbiased					
VMT (10^6)	80	52	162	76	370
Intersection maneuver (10^6)	132	558	350	452	1490
Maneuver/VMT	1.6	10.6	1.8	6.0	4.1
	5.2		3.4		
Biased					
VMT (10^6)	86	52	144	70	352
Intersection maneuver (10^6)	148	496	304	560	1506
Maneuver/VMT	1.7	9.6	1.8	8.0	4.3
	4.7		4.0		

First, we simply applied the expansion factors separately to the observations during each of these time periods. The results are shown in Table 4.3.4-1. The ratio between intersection maneuvers and VMT varies greatly among the four time periods. Part of this variation may be random; indeed, the ratios vary somewhat less for the biased estimates where one component of the random variation, that due to the between squares variation of the expansion factors, has been eliminated. A closer look at the observation schedule suggests that most of the differences are due to an imbalance in the observation schedule: there are relatively more observations in Kingston during the periods 11-15 and 19-23 hours than during the others.

To compensate for this, the following was done. Each of the four periods was treated separately, and separate expansion factors were developed. There were not enough setups to cover each location stratum at least once, and each time stratum at least once, therefore similar strata had to be combined. This will increase the errors of the resulting estimates. The estimates using these time-specific expansion factors are also shown in Table 4.3.4-1. The variation in the ratio of intersection maneuvers to VMT has been noticeably reduced, but not completely eliminated. This was surprising, because the time-specific extrapolation factors should automatically reduce the effect of the over-represented Kingston observations, and increase that of the under-represented observations outside of Kingston. Inspection of the observation schedule showed that the observation schedule for Kingston arteries was relatively unbalanced over time and that interactions between time and location gave a few observation sessions relatively high weights, thereby preventing the "averaging" of random fluctuations. We concluded that using time of day as factor or stratifier, even as grossly as in four-hour periods, is not advisable, because it strongly interacts with the "urban" (Kingston) and "rural" environment. Where a distinction day/night is necessary, because of missing information at night, one has to keep in mind that night observations over-represent Kingston, day observations under-represent it.

If one aggregates the data over the early shift and the late shift, differences between the intersection maneuvers per VMT remain but they are reduced, even more so when using biased estimators, and further using time-specific expansion factors. This is not surprising, because the design balanced the observations over shifts as much as possible.

4.4 Driver Characteristics

4.4.1 Driver Sex

If a driver can be observed, his or her sex can be assessed with a high degree of confidence. At night, drivers could be observed only under favorable circumstances. Because the proportions of male and female drivers differ probably between day and night, they were treated separately. "Night" was defined to begin at 19 hours, because this allowed easiest separation of the observations.

During the day, the driver's sex was undeterminable for 6% of the VMT, but for 13% of the intersection maneuvers. At night, it was undeterminable for 95% of the VMT, but for only 85% of the intersection maneuvers. One reason for these discrepancies may be that conditions for observing drivers vary greatly between sites, especially at night.

Table 4.4.1-1 presents the percentages of exposure and accidents for male drivers. Men have a higher proportion of VMT at night than during the day. The

TABLE 4.4.1-1
RELATIVE EXPOSURE AND ACCIDENT INVOLVEMENT BY DRIVER SEX
(Note that exposure estimates for the hours 19-23 are based on a small fraction of total exposure.)

	Accident involvement		Exposure	
	7-19 hrs	19-23 hrs	7-19 hrs	19-23 hrs
Segment accidents	Percent of Male Drivers		Percent of Male Drivers	
Single	60	72		
Head-on	62	80		
Rear-end	57	62		
Other	54	66		
Total	57	70	63	76
Intersection accidents				
Straight	58	66	64	60
Left turn	53	48	62	67
Right turn	60	79	51	70
Other	75	*	68	58
Total	57	62	61	62

*Only 1 accident.

average (using the distribution of VMT over day/night from Table 4.3.4-1) is 65%. This is less than the nationwide proportion of 71% of total VMT driven by male drivers (derived from Table 15 of Report No. 1: Characteristics of 1977 Licensed Drivers and Their Travel, 1977 Nationwide Personal Transportation Study, U.S. Dept. of Transportation, Federal Highway Administration, October 1982).

Total VMT, however, are for all motor vehicles. Tabulations of VMT for passenger cars and station wagons were provided by Ms. S. Smith of NHTSA from the NPTS data.* Here, the percentage of male drivers is 66%, which agrees well with our observations.

*Also, these VMT figures are based on the estimates for individual trips and therefore possibly more reliable than estimates of total annual mileage by a driver.

The proportion of male driver involvements in segment accidents is lower than their proportion of exposure. The relation is the same for all segment accident types, with one exception: head-on collisions at night, where the male involvement is much higher.

If men have a proportion p of exposure and q of accidents, women have 1-p and 1-q. Consequently, the risks of accident involvements for men and women relate as q/p to (1-q)/(1-p); the ratio of the risk for women to that for men is

$$r = \frac{(1-q)p}{(1-p)q} \qquad (4.4.1\text{-}1)$$

It is 1.28 for the day, and 1.36 for the night. This means that women have a 28% to 36% higher segment accident risk per VMT than men. This contradicts all previous findings which, however, are based on aggregate VMT figures, usually relying on drivers' subjective estimates of their VMT.

Though our estimate of the overall percentage of male exposure agrees with the findings from the NPTS, one cannot necessarily conclude that our figures are correct. It is conceivable that there are regional differences in the distribution of exposure by driver sex, and that the actual percentage of female exposure in Ulster County is higher than the nationwide average. If one makes the extreme assumption that the 6% of VMT by unknown drivers during the day are driven by women, then the percentage of male VMT decreases to 59% which is close to the male accident involvement. Such an extreme bias, however, is not very likely. For the night, the actual percentages could differ drastically from the estimates, because of the overwhelming percentage of unknown cases.

During the day, the proportion of intersection accident involvement for men is the same as for segment accidents; at night it equals their proportion of exposure. Therefore, at night, the intersection accident risk for women is equal to that for men; during the day it is 18% higher.

Comparing the different types of intersection accidents suggests no pattern.

That women have higher (up to one-third higher) accident risks per exposure than men is unexpected and surprising. To explore this finding further, the covariates speed and traffic volume* were studied in Table 4.4.1-2. Traffic

*Traffic valumes are averaged over VMT or intersection maneuvers. Therefore high volume corresponding to high exposure were weighted heavily. This is the reason for the apparently high volume. The same holds for speed.

TABLE 4.4.1-2
SPEED AND TRAFFIC VOLUME BY DRIVER SEX AND TIME OF DAY
(Averages were calculated using exposure for weighting.)

	Day (7-19)			Night (19-23)		
	Male	Female	Unkn.	Male	Female	Unkn.
Segments						
Average individual speed (mph)	38	38	40	31	38	32
Average traffic speed (mph)	38	38	40	29	37	32
Average traffic volume (vph)	236	249	378	382	389	233
Intersections						
Average traffic volume (vph)	191	188	250	157	158	337

volumes do not differ between the exposures for men and women. Speeds during the day do not differ but they do at night. However, the average speed of the "unknowns" is so much lower that the actual average speed of female drivers may not differ from that of male drivers.

It is noteworthy that the average individual speeds differ only little, or not at all from the average traffic speeds.

4.4.2 Driver Age

In addition to observing a driver's sex, his or her age was estimated. There were a few cases where driver age could be recognized as "young," but the sex not determined. The percentages of missing data are essentially the same as for driver sex. Again, "day" and "night" was distinguished in the analysis.

Table 4.4.2-1 shows the percentages of intersection and segment accidents, involvements by driver age group, and the percentages of VMT, of the total for which driver age estimates were made.

During the day, the accident rate for young drivers is only about 10% higher than that for middle age drivers; that for old drivers (over 50) is two to four times as high as that for middle age drivers. At night, the rate for young drivers is much lower than for middle age or old drivers.

This contradicts the current knowledge about the relation between driver age and accident risk. Table 4.4.2-2 shows data similar to those in 4.4.2-1 derived from an NHTSA study of national accident data and exposure. There is practically no difference between rates for middle age and old drivers,

TABLE 4.4.2-1
ACCIDENT INVOLVEMENTS AND VMT BY DRIVER AGE,
AS PERCENTAGE OF THE TOTAL WITH DRIVER AGE INFORMATION

	Driver Age		
	Young	Middle	Old
7 to 19 hours			
% involvements	36	39	25
% VMT from segment observation	42	50	8
% involvements/ % VMT	0.86	0.78	3.0
% VMT from inter-section observation	38	47	15
% involvements/ % VMT	0.95	0.85	1.6
19 to 23 hours			
% involvements	45	42	13
% VMT from segment observations	57	39	4
% involvements/ % VMT	0.79	1.08	3.2
% VMT from inter-section observations	65	25	10
% involvements/ % VMT	0.69	1.7	1.3

TABLE 4.4.2-2
AGE DISTRIBUTION OF ACCIDENT INVOLVED DRIVERS AND OF VMT
(1979-80 Average from the National Accident Sampling System*

	Driver Age		
	Under 26	26 - 50	Over 50
% accident involved drivers	42	41	17
% VMT	22.8	54	23
% involvements/ % VMT	1.8	0.76	0.74

*Derived from Table 37 of, U.S. Department of Transportation, National Highway Traffic Safety Administration, Report on Traffic Accidents and Injuries for 1979-80. DOT HS 806 176, February 1982.

and the rate for young drivers is about twice as high as that for others. These data include VMT for all motor vehicles. Data for passenger cars and station wagons were obtained from the NHTSA tabulations mentioned in Section 4.4.1. They give 20% of VMT for drivers under 25, 55% for drivers 25 to 49, and 25% for drivers of 50 years or older. They differ only little from those in Table 4.4.2-2.

One plausible explanation for these discrepancies is that the ages estimated by the observers are biased downwards (the ages in the accident reports are taken from the drivers licenses and presumably correct). Underestimating the drivers' ages increases the number of VMT for "young" drivers, and reduces the estimated risk. VMT for old drivers are reduced, and their accident risk increased. VMT for middle age drivers may be increased or decreased, depending on the magnitude of the bias in relation to age, and the distribution of driver ages.

To test this, Fig. 4.4.2-1 shows the distribution of VMT by driver age, derived from the NHTSA tabulations in 5-year intervals. If we assume that

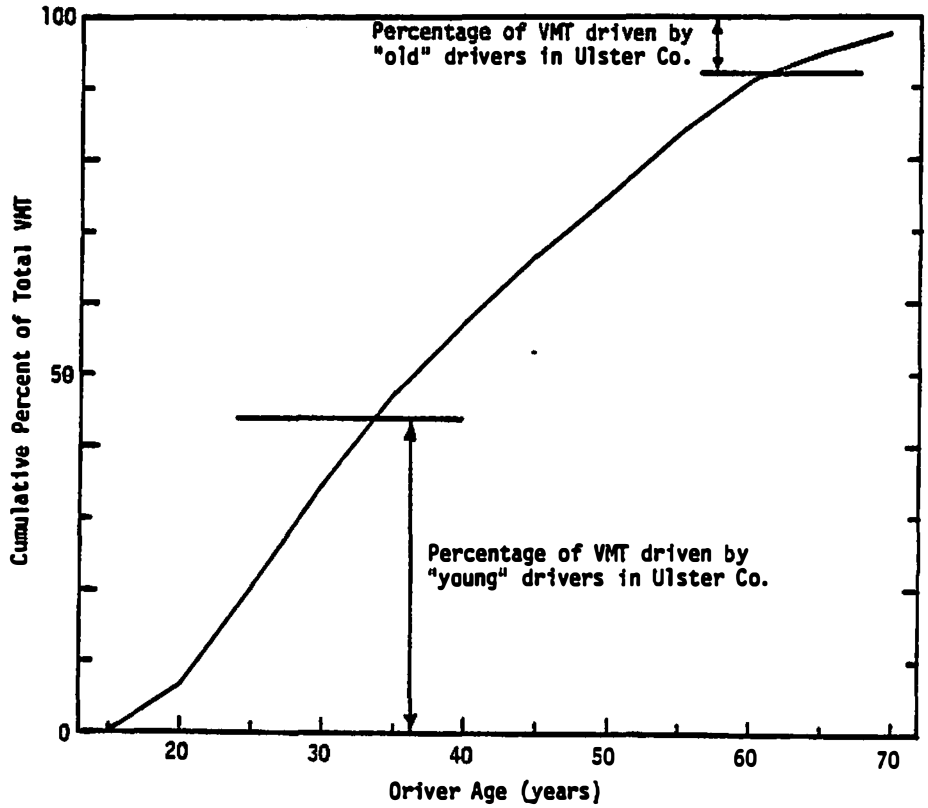


Figure 4.4.2-1. Cumulative distribution of VMT in passenger cars and station wagons, by driver age. Source: 1977 Nationwide Personal Transportation Study Data, analyzed by NHTSA. Also shown are the estimated percentages of total VMT (7-23 hours) in Ulster County driven by "young" and "old" drivers.

the distribution of VMT by age in Ulster County is the same as nationwide, then the 44% of VMT driven in Ulster County by the youngest age group would correspond to a cut-off age of 33-34 years. The 8% of VMT driven by the oldest age group corresponds to a cut-off point of 61 years. This would mean that the observers' estimates of ages were biased by 8-9 years for younger, and 11 years for older drivers. Though it is quite plausible that in an individual case an estimated age can be wrong by 10 years, it appears implausible for the average of thousands of observations.

Table 4.4.2-3 allows a different look at the same question. It shows within each of the age groups the percentage of VMT by male drivers. In Ulster

TABLE 4.4.2-3
PERCENTAGE OF VMT BY MALE DRIVERS BY AGE GROUPS
NPTS(A) uses 25 years for "young," and 50 for "old";
NPTS(B) uses 35 for "young," 50 for "old."

Driver Age Group	Percent of Male Drivers		
	Ulster County	NPTS	
		(A)	(8)
Young	65	61	63
Middle Age	67	65	66
Old	64	71	76

County, the percentage of male drivers is higher in the younger groups, and lower in the oldest group, when compared with the national figures using the same age breakdown (A). If one uses cut-off points of 35 and 60 years for the national figures (B), the percentage of young drivers is closer to that observed in Ulster County, but that for old drivers differs more. These data do not support the hypothesis of an age bias of about 10 years, but do not contradict it either.

To determine whether there were obvious differences in driving environment among the age groups, we looked at travel speeds. Table 4.4.2-4 shows travel speeds, and their standard deviations. During the day, there is practically no difference among the age groups; at night, young drivers drive in faster traffic.

TABLE 4.4.2-4
AVERAGE TRAFFIC SPEED, AND STANDARD DEVIATION OF TRAFFIC SPEEDS BY DRIVER AGE AND SEX

	Young		Middle		Old	
	Male	Female	Male	Female	Male	Female
7-19 average						
Traffic speed	32	31	32	31	30	30
Standard Deviation	5.7	5.3	5.0	5.6	5.2	5.8
19-23 Average						
Traffic speed	40	40	24	28	--	--
Standard Deviation	6.0	6.0	4.0	4.3	--	--

Though it is possible that driver age estimates are biased, it is implausible that the differences in the involvement rates among the age groups are entirely due to bias. Because the findings contradict the current state-of-the-knowledge, further, more thorough studies are worthwhile.

For the current study, we will not use age as a pre-crash factor, because it is likely to be unreliable. However, we may use it as a "stratifier," and compare rates within driver age groups, but not among driver age.

4.5 Highway and Traffic Factors

4.5.1 Introduction

The following highway characteristics were selected for study because we expected them to have, separately and in interaction with each other and other factors, strong effects on accident risk:

- alignment
- grade
- surface conditions.

Because the information on police accident records was of limited detail, only similarly limited information was collected for exposure. For instance, there is no indication of the degree of curvature. Not even the direction of the curve is shown in the structured data, though it can often be seen on the accident diagram. Neither the degree, nor the direction of the grade are given, and they are not shown on the diagram. Because the limited number of cases would not allow disaggregation into too many categories, only two levels were distinguished for each factor:

- straight/curved
- grade/level
- dry/wet.

In the accident records, this information depends on the subjective assessment by the police officer investigating the accident, in the exposure data on that of the observers. It is possible that there are systematic differences between the judgment of the police officers—and also among the many officers investigating accidents--and that of the observers. However, this could not be tested within the scope of the study.

A conceptual shortcoming is that curvature was used only as a stratifier for the exposure measure. Thus, for segment accidents, rates per "straight" VMT, and per "curved" VMT were compared. If the probability of an accident in a curve would increase proportional to its length, "straight" and "curved" VMT would indeed be the appropriate exposure measure to assess the effect of curvature on risk. However, if the risk in a curve is greatest when entering it, and much lower thereafter, so that each curve is a unit of exposure, essentially independent of its length, then the number of curves passed would be the preferred exposure measure. This exposure measure, however, was not available; it would have required the collection of additional data. Because of the use of a possibly less suitable exposure measure, findings on the effects of curvature should be interpreted with great caution.

The situation with grades appears different. First, one would expect that the risk on a downgrade is much greater than on an upgrade. Since we cannot distinguish upgrades and downgrades in the accident data, we can estimate only an "average" of the effects of both types on risk: it could be small, though each direction of grade could have a fairly large effect on the risk. Second, one would expect that the effect of grade is proportional, if not increasing, with its length. Therefore, VMT on grades may be an adequate measure of exposure.

The most obvious traffic factors are traffic speed, and traffic volume. They could not be rigorously studied, because they were not available for accidents, only for exposure data. Therefore, they could be considered only as "covariates."

The use of exposure observations for the period May through October 1981, and of accident data for April through November 1980 and October 1981, raises the question whether pre-crash conditions for the two periods were comparable. This holds especially for weather conditions which vary greatly over the year, and differ also from year to year. Such variations could affect the conclusions regarding the effect of a wet highway surface on accident risk.

Therefore, the distribution of accidents over highway characteristics and surface conditions was compared for the entire study period April through November 1980 and October 1981, and the period May 7 through October 1981, covered by the exposure observations. Tables 4.5.1-1 and 4.5.1-2 show the results for segment accidents, and for intersection accidents, respectively.

TABLE 4.5.1-1
DISTRIBUTION OF SEGMENT ACCIDENTS BY HIGHWAY CHARACTERISTICS AND SURFACE CONDITION FOR THE PERIOD 5/81-10/81, AND THE ENTIRE STUDY PERIOD APRIL THROUGH NOVEMBER 1980 AND APRIL THROUGH OCTOBER 1981
(Upper figures are numbers of accident involvements; lower figures are percentages of all accidents.)

	Level		Grade	
	3/81-10/81	1980,81	5/81-10/81	1980,81
Dry				
Straight	327 51.0	881 51.6	72 11.2	190 11.1
Curved	54 8.4	164 9.6	50 7.8	144 8.4
Wet				
Straight	69 10.8	164 9.6	20 3.1	53 3.1
Curved	24 3.8	54 3.2	25 3.9	56 3.3

TABLE 4.5.1-2
DISTRIBUTION OF INTERSECTION ACCIDENTS BY HIGHWAY CHARACTERISTICS AND SURFACE CONDITIONS FOR THE PERIOD 5/81-10/81, AND THE ENTIRE STUDY PERIOD
(Upper figures are numbers of accident involvements; lower figures are percentages of all accidents.)

	Level		Grade	
	3/81-10/81	1980,81	5/81-10/81	1980,81
Dry				
Straight	154 58.8	406 60.0	39 14.9	96 14.2
Curved	8 3.1	30 4.4	14 5.3	28 4.1
Wet				
Straight	32 12.2	85 12.6	8 3.0	17 2.5
Curved	3 1.1	7 1.0	4 1.5	7 1.0

Of segment accidents in the study period, 80% occurred on dry roads, of those during the exposure observation period, 78%. Of the intersection accidents during the study period, 83% occurred on dry roads, of those during the exposure observation period, 82%. These differences are so small that they cannot affect the conclusions reached.

The distribution of accidents over the combinations of highway and urban characteristics are also very similar for the study period, and the exposure observation period. There are some differences, especially for the combinations with few accidents. However, comparing the distribution of accidents over the eight factor combinations during the exposure observation period, and the remaining part of the study periods gives Chi-squares of 5 for the intersection accidents, and 6 for the segment accidents. Both values are far from being significant with seven degrees of freedom.

Therefore, it is acceptable to use the accident data for the entire study period. However, one has to be aware that doing this means trading off a relatively lower variance of the accident numbers against a possible bias.

4.5.2 Studying the Highway Factors Separately

Table 4.5.2-1 shows VMT, segment accident involvements, and involvement rates, separately for straight and curved highway locations. Three-quarters of VMT are on straight segments of road, one quarter in curves. The overall rates for straight and curved road sections are nearly equal, but the rate for single-car accidents is much higher in curves than on straight segments, whereas rates for collisions between vehicles are lower. The first corresponds to the intuitive expectation. However, one would not expect that the rates for multi-vehicle accidents would be lower in curves than on straight segments. While speeds on straight and curved sections differ only little, traffic volume differs much. On curved segments, it is only 0.7 of that on straight segments; the rate for multi-vehicle accidents on curved sections is 0.6 of that on straight sections. This suggests that traffic volume has a strong influence on the occurrence of multi-vehicle accidents, which is very plausible.

Table 4.5.2-2 shows VMT and accidents by highway grade. Fifty-seven percent of the VMT is on level sections. The overall accident rate on grades is only about half as high as on level sections. This is surprising. However, the differences are not uniform. For single-vehicle accidents, the rate

TABLE 4.5.2-1
SEGMENT ACCIDENTS AND EXPOSURE BY ROAD ALIGNMENT

	Straight		Curved		Ratio of Rates
VMT (10^6)	304		106		
Accident involvements	Number	Rate	Number	Rate	
Single car	187	0.6	175	1.6	2.8
Head-on	305	1.0	69	1.2	0.6
Rear-end	182	0.6	20	0.19	0.3
Others	614	2.0	154	1.4	0.7
Total	1,288	4.2	418	4.0	0.9
Average speed (mph)	36		39		
Average volume (vph)	267		195		0.7

TABLE 4.5.2-2
SEGMENT ACCIDENTS AND EXPOSURE BY HIGHWAY GRADE

	Level		Grade		Ratio of Rates
VMT (10^6)	232		180		
Accident involvements	Number	Rate	Number	Rate	
Single car	243	1.0	119	0.6	0.6
Head-on	276	1.2	98	0.5	0.4
Rear-end	146	1.2	56	0.3	0.4
Others	598	2.6	170	0.9	0.4
Total	1,263	5.5	443	2.4	0.4
Average speed (mph)	38		36		0.8
Average volume (vph)	275		213		0.7

on grades is nearly two-thirds of that on level sections, but for all types of collisions it is less than half. Speeds are slightly lower on grades, as is volume. However, the 20% reduction in volume is much less than the 60% reduction in accident rates.

Table 4.5.2-3 shows accidents and exposure by surface condition. Only 8% of VMT were travelled on wet roads. Accident rates for wet roads are consistently higher than for dry roads; it is surprising that the ratio varies relatively little among the different types of accidents.

Speed on wet roads is lower than on dry roads. However, one should be cautious in interpreting this as an effect of the surface condition, because traffic volume is twice as high on wet surfaces as on dry surfaces! It is unlikely that a wet surface increases travel; it is more plausible that some observations when the road was wet happened by chance to occur at high volume locations.

TABLE 4.5.2-3
SEGMENT ACCIDENTS AND EXPOSURE BY SURFACE CONDITION

	Dry		Wet		Ratio of Rates
VMT (10^6)	376		34		
Accident involvements	Number	Rate	Number	Rate	
Single car	294	0.8	68	2.0	2.5
Head-on	301	0.8	73	2.2	2.7
Rear-end	165	0.4	37	1.1	2.4
Other	620	1.6	149	4.4	2.6
Total	1,380	3.6	327	9.6	2.6
Average speed (mph)	38		32		
Average volume (vph)	228		466		2.0

Since the ratio of accident rates for the two factor levels has some similarity with the ratio of traffic volumes, Fig. 4.5.2-1 shows the rates versus the traffic volume. For multi-vehicle accidents the rates for curved and straight, and dry and wet sections are practically on a straight line, and those for level and grades do not differ much. One could have the data represented by a straight line. This would mean that the accident rate is only a function of traffic volume, and that alignment, grade, and surface conditions have no or only small effects on the multi-vehicle segment accident rates. A line fitting the points best would give rates which increase faster than proportional with volume. However, a line representing a proportional increase with volume (shown in the figure) would still represent the points, with the exception of that for wet surfaces.

For single-vehicle accidents, the pattern is similar, with the exception of straight and curved segments. This is surprising, because one would not expect the single-vehicle accident rate to increase with traffic volume; if at all, one would expect a decrease because some incidents which would result in a single-vehicle accident, if no other vehicles are present, can result in a collision if other vehicles are present.

A clear deviation from this pattern is that the single-vehicle accident rate is much higher on curved sections than on straight sections.

To explore the effects of highway characteristics further, intersection accidents were also studied. One would expect that in intersection accidents where the interaction of vehicles is important, highway characteristics have a lesser impact than in segment accidents, especially single-vehicle accidents.

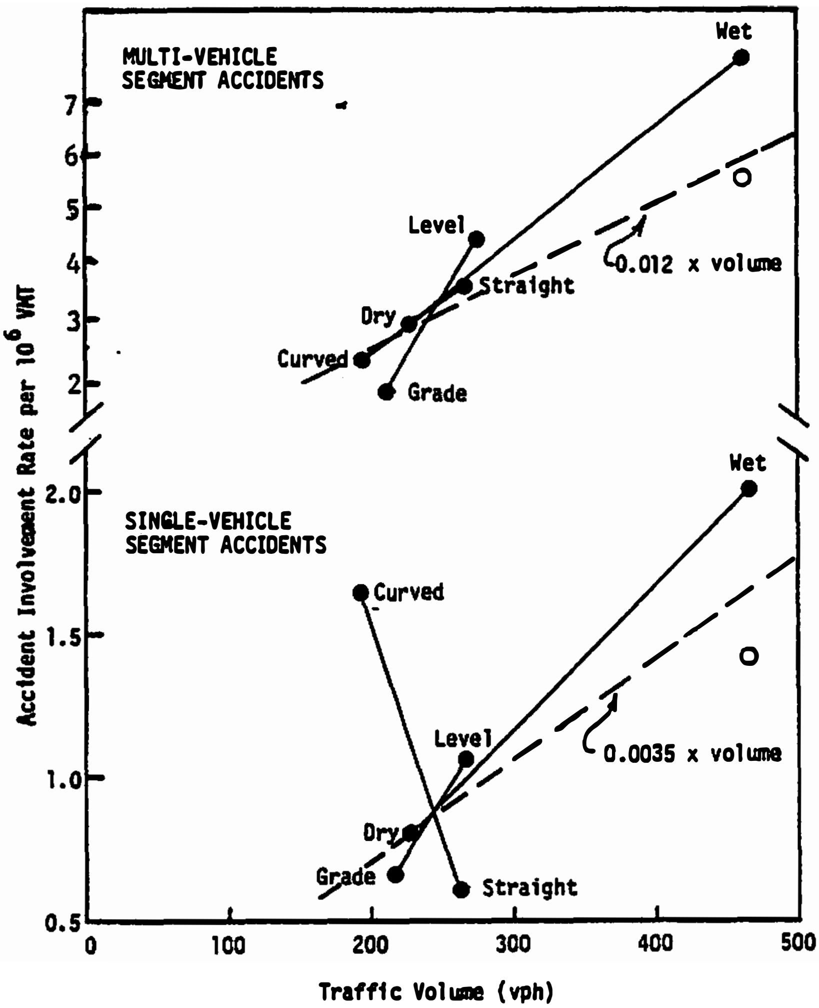


Figure 4.5.2-1. Involvement rates for segment accidents versus traffic volume. Note that the pairs of points represent different dichotomies of the same accident and exposure data. The circles show the rates for wet pavements under differing assumptions on its frequencies.

Tables 4.5.2-4 through 6 show the comparisons of straight/curved, level/grade, and dry/wet intersection accidents. Figure 4.5.2-2 summarizes some of the results, similar to Fig. 4.5.2-1.

Straight/curve, and level/grade rates in Fig. 4.5.2-2 show a similar pattern as in Fig. 4.5.2-1 for single-vehicle accidents; dry/wet rates differ, but wet rates are 20% higher than dry rates, which were obtained in traffic with slightly lower volume. With regard to specific maneuvers, right turns have an increased absolute—and even more increased relative risk—in curves, compared with straight intersections. There is no plausible explanation for this. On wet surfaces, the risk in left turns is greatly increased. Again, there is no obvious and no convincing explanation for this, but one might speculate that wet surfaces increase stopping distances and thereby the risk of a collision if one vehicle turns into the path of another.

TABLE 4.5.2-4
INTERSECTION ACCIDENTS AND EXPOSURE BY ROAD ALIGNMENT

Maneuver	Alignment						Ratio of Rates
	Straight			Curved			
	Accident Involvements	Maneuver (10^6)	Rate	Accident Involvements	Maneuver (10^6)	Rate	
Straight	370	1092	0.34	34	180	0.19	0.6
Left turn	182	110	1.6	23	32	0.75	0.4
Right turn	41	198	0.20	13	32	0.40	2.0
Other	11	44	0.25	2	5	0.40	1.6
All	604	1442	0.42	72	248	0.29	0.7
Traffic volume (vph)	257			153			0.6

TABLE 4.5.2-5
INTERSECTION ACCIDENTS AND EXPOSURE BY HIGHWAY GRADE

Maneuver	Level			Grade			Ratio of Rates
	Accident Involvements	Maneuver (10^6)	Rate	Accident Involvements	Maneuver (10^6)	Rate	
Straight	318	814	0.39	86	456	0.19	0.5
Left turn	161	98	1.65	44	44	1.0	0.6
Right turn	39	160	0.24	15	68	0.22	0.9
Other	10	44	0.23	3	5	0.65	2.8
All	528	974	0.54	148	574	0.28	0.5
Traffic volume (vph)	279			200			0.7

TABLE 4.5.2-6
INTERSECTION ACCIDENTS AND EXPOSURE BY ROAD SURFACE

Maneuver	Surface Condition						Ratio of Rates
	Dry			Wet			
	Accident Involvements	Maneuver (10^6)	Rate	Accident Involvements	Maneuver (10^6)	Rate	
Straight	327	1050	0.31	77	222	0.34	1.1
Left turn	172	130	1.3	33	9.8	3.4	2.6
Right turn	51	212	0.24	3	15.4	0.20	0.8
Other	10	46	0.22	3	2.6	1.2	5.4
All	560	1440	0.39	116	250	0.47	1.2
Traffic volume (vph)	244			231			0.9

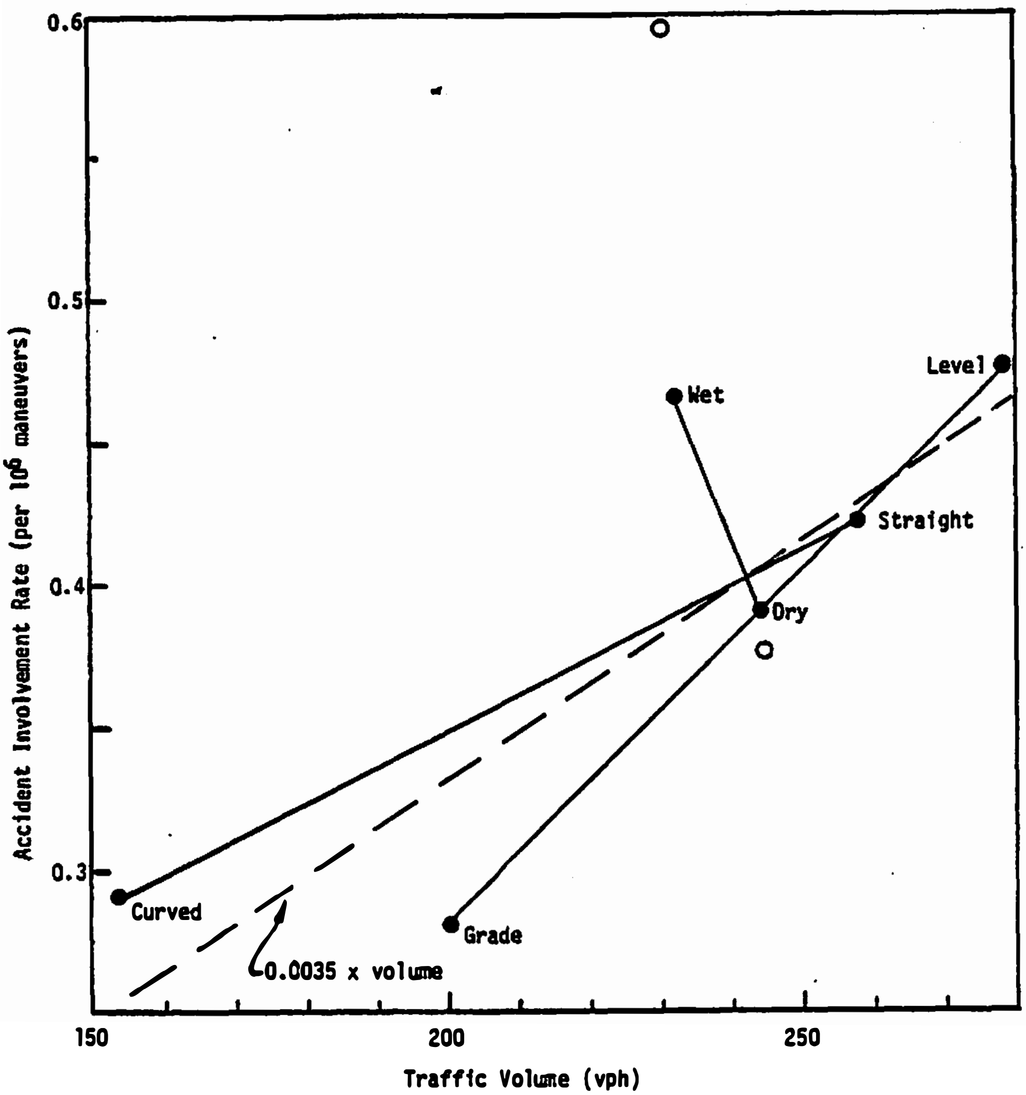


Figure 4.5.2-2. Involvement rates for intersection accidents versus traffic volume. Note that the pairs of points represent different dichotomies of the same accident and exposure data. The circles show the rates for wet and dry pavements under differing assumptions on their frequency.

Comparing the distribution of straight/curved, and of grade/level exposure between intersections and segments allows no conclusions, because intersections are concentrated in Kingston, segments outside of Kingston. However, comparing dry/wet exposure is of interest. Of segment exposure, 8% occurred on wet surfaces of intersection exposure 15%. Various factors can contribute to such differences. For instance, there can be great local differences in precipitation. It could be possible that travel in rural areas is reduced by rain (to determine this, one would need to know the exact time distribution of rain and also the travel patterns to estimate which exposure would have occurred at the times of the rain, but without it) more than in urban areas. The most plausible explanation is that because wet pavement was encountered only in about one-tenth of the exposure observations, the estimates of its frequency are not very reliable--the difference between 8% and 15% could be a chance variation.

To explore the consequences of this, we hypothesize that the relative exposure on wet surfaces is the same for intersections as for segments--11.5%, the average of 8% and 15%.

The consequence is that the segment accident rates would be only 1.8 times as high on wet surfaces as on dry surfaces. For intersection accidents, the ratio of wet/dry rates would increase to 1.6. In Figs. 4.5.2-1 and 4.5.2-2, the changed points for "dry" and "wet" surfaces are shown.

Summarized, the findings are that accident rates on wet surfaces are always higher than on dry surfaces, as one would expect. Contrary to expectation, however, they are lower on grades than on level sections, and in curves than on straight sections, with the exception of single-vehicle accidents, for which the rate is higher in curves.

However, a relation between traffic volumes and accident involvement rates is apparent which could hide the actual effects of grades, curves and wet surfaces. Only the effects of wet surfaces on intersection accidents and curves on single-vehicle accidents could not be affected by such an effect.

Without knowing the traffic volume for the times and locations of accidents, it is not possible to eliminate an effect of traffic volume convincingly. If the data were more finely stratified, an analysis of traffic volume as a covariate could shed more light on its potential effect. This will be done in the later sections dealing with interactions.

4.5.3 Interactions of Factors

4.5.3.1 Interactions of Driver Sex and Highway Factors

Differences among drivers in driving experience and risk-acceptance might result in differences in accident involvements in curves and on wet highways. We did not study interactions with grade, because we believe that they are less perceived as risky situations, though they might in fact be.

Table 4.5.3.1-1 shows segment exposure and accident involvement by driver sex and highway alignment. In this Table, total VMT on straight and curved segments, during the periods of the day indicated, were allocated according to the percentages of male and female drivers which were observed. This was necessary to make the rates in the four "cells" of the Tables comparable, because the proportions of missing data varied.

The overall pattern is the same as found in Sections 4.4.1 and 4.5.2: rates for females are higher than for males, and higher for straight than for curved segments. Figure 4.5.3.1-1 shows the rates in relation to average traffic volumes: there is no consistent relation with traffic volume. Neither is there an interaction of sex and highway alignment: the differences between male and female rates for straight segments and curves are not consistent between the two parts of the day studied.

Table 4.5.3.1-2 and Fig. 4.5.3.1-2 show the corresponding rates for intersection accidents. The pattern is essentially the same, except that for the 7-15 hour period the accident rate for females in curves is higher than on straight segments.

Table 4.5.3.1-3 and Fig. 4.5.3.1-3 show the segment accident involvement by driver sex and highway surface. Total VMT for dry and wet surfaces are distributed according to the observed proportions of male and female drivers--the sex could not be ascertained for 3% of the VMT on dry surfaces during the 7-19 hour period, 1% for the 7-15 hour period; for wet surfaces, however, these percentages were 60% and 15%. This suggests that rain impeded the driver observations.

Table 4.5.3.1-4 and Fig. 4.5.3.1-4 show the data for intersection accidents. Sex was missing for 2% of the dry surface maneuvers 7-19 hours; 1% for 7-15 hours, but for 66% and 70% of the wet surface maneuvers.

TABLE 4.5.3.1-1
SEGMENT ACCIDENT INVOLVEMENT BY HIGHWAY ALIGNMENT AND DRIVER SEX

	Straight				Curved			
	Accident Involve-ments	VMT (10^6)	Rate (per 10^6)	Volume (vph)	Accident Involve-ments	VMT (10^6)	Rate (per 10^6)	Volume (vph)
Male								
7-19	576	152	3.8	248	164	62	2.6	214
7-15	303	68	4.4	245	90	26	3.4	301
Female								
7-19	441	98	4.5	261	111	28	4.0	212
7-15	245	40	6.0	33	52	14	3.6	291

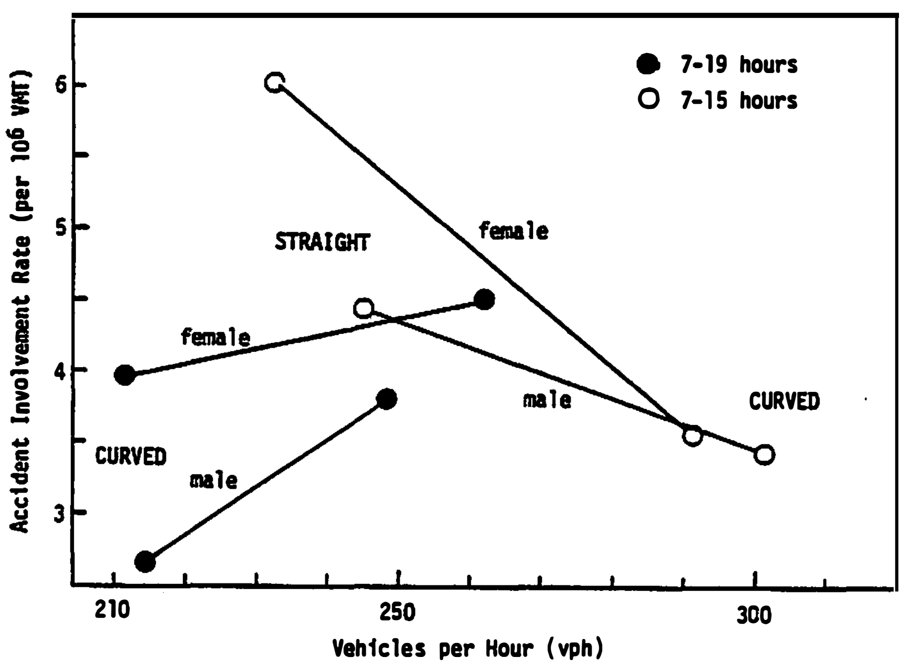


Figure 4.5.3.1-1. Segment accident involvement rates versus traffic volume by highway alignment and driver sex.

TABLE 4.5.3.1-2
INTERSECTION ACCIDENT INVOLVEMENTS BY HIGHWAY ALIGNMENT AND DRIVER SEX

	Straight				Curved			
	Accident Involvements	Maneuver (10^6)	Rate (per 10^6)	Volume (vph)	Accident Involvements	Maneuver (10^6)	Rate (per 10^6)	Volume (vph)
Male								
7-19	285	588	0.48	207	30	92	0.32	99
7-15	147	424	0.34	220	13	58	0.22	75
Female								
7-19	212	370	0.55	208	26	60	0.44	77
7-15	128	300	0.42	216	16	30	0.55	71

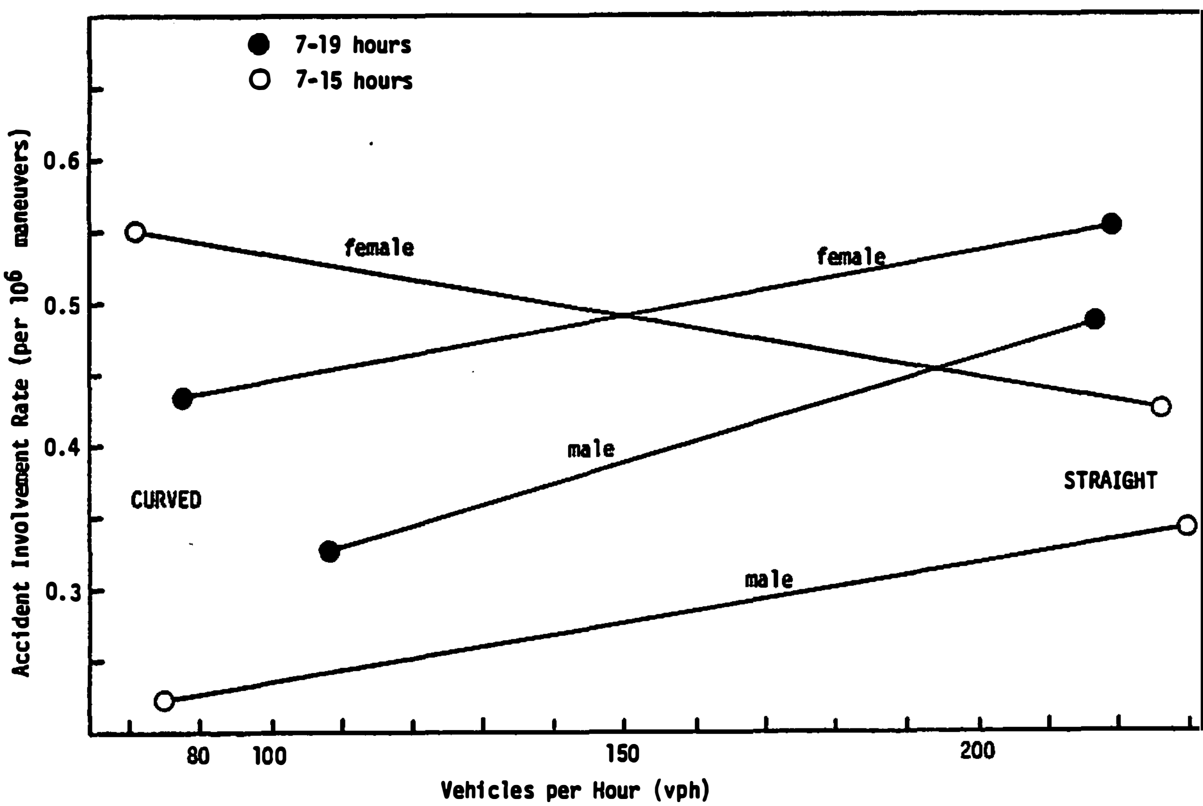


Figure 4.5.3.1-2. Intersection accident involvement rates versus traffic volume by highway alignment and driver sex.

TABLE 4.5.3.1-3
SEGMENT ACCIDENT INVOLVEMENTS BY DRIVER SEX AND HIGHWAY SURFACE

Highway Surface	Driver							
	Male				Female			
	Acci-dents	VMT (10^6)	Rate (per 10^6)	Avg. Volume (vph)	Acci-dents	VMT (10^6)	Rate (per 10^6)	Avg. Volume (vph)
Dry								
7-19 hrs	594	202	3.0	243	447	118	3.8	253
7-15 hrs	313	80	3.4	268	244	52	4.7	252
Wet								
7-19 hrs	146	14.4	10	53	105	5.6	19	57
7-15 hrs	80	4.0	20	52	53	1.4	38	60

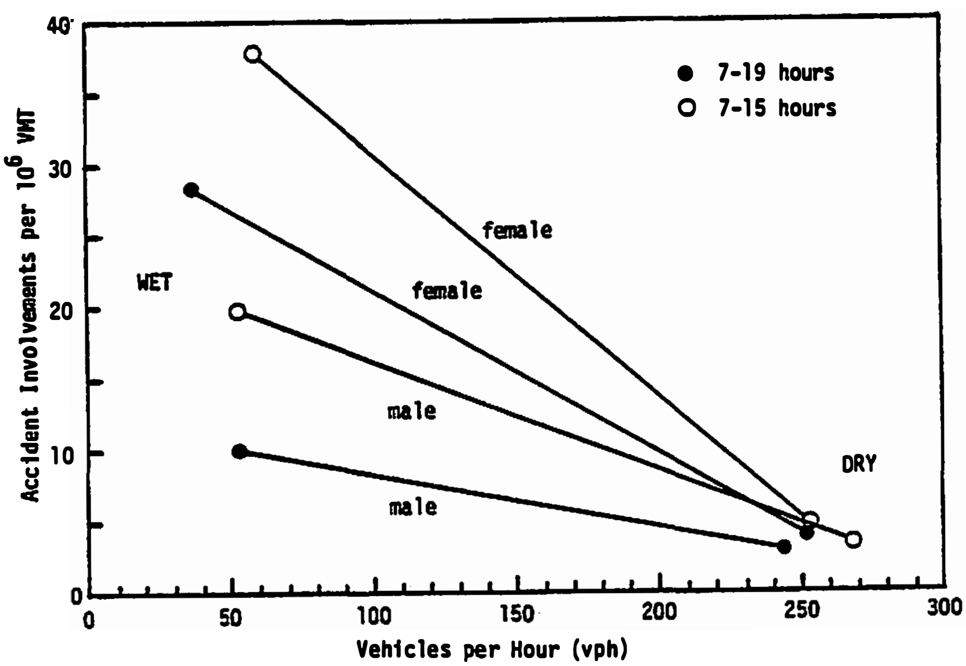


Figure 4.5.3.1-3. Segment accident involvement rates versus traffic volume by highway surface and driver sex.

TABLE 4.5.3.1-4
INTERSECTION ACCIDENT INVOLVEMENTS BY DRIVER SEX AND HIGHWAY SURFACE

Highway Surface	Driver							
	Male				Female			
	Acci-dents	VMT (10^6)	Rate (per 10^6)	Avg. Volume (vph)	Acci-dents	VMT (10^6)	Rate (per 10^6)	Avg. Volume (vph)
Dry								
7-19 hrs	211	528	0.40	193	188	332	0.55	190
7-15 hrs	135	382	0.36	212	108	260	0.42	211
Wet								
7-19 hrs	44	152	0.29	173	64	98	0.65	173
7-15 hrs	25	98	0.26	67	36	74	0.50	85

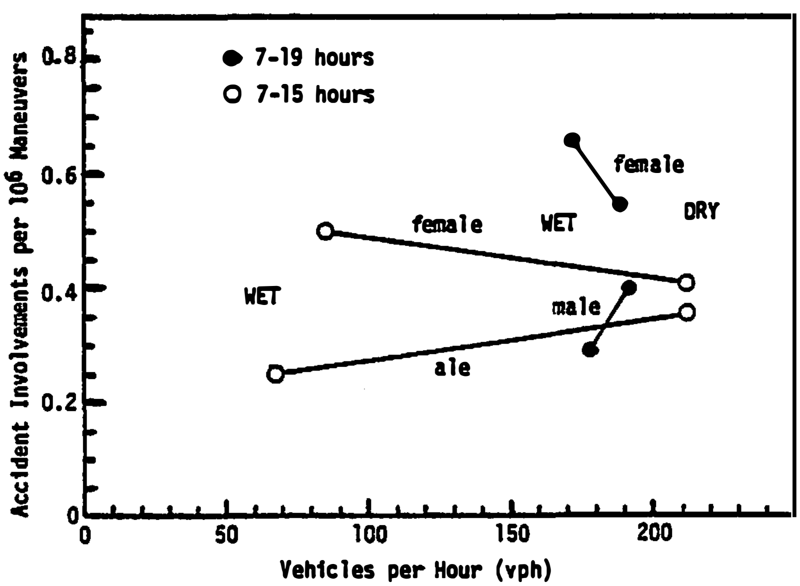


Figure 4.5.3.1-4. Intersection accident involvement rates versus traffic volume by highway surface and driver sex.

The accident rates for female drivers are higher than for male drivers, for wet as well as for dry surfaces. Rates for wet and dry surfaces should be compared only with caution, as discussed in Section 4.5.2. However, even if the relative frequencies of dry and wet surfaces are affected by sampling variations, one can still examine the interaction between driver sex and highway surfaces. Figures 4.5.3.1-3 and 4 suggest such an interaction: the accident rates for male and female drivers differ more in wet weather than in dry weather. We make the following comparisons:

Segment Accidents		Ratio of Male to Female Rates	Double Ratio
7-15	dry	1:1.39	1.4
	wet	1:1.9	
15-19	dry	1:1.21	1.6
	wet	1:1.95	
Intersection Accidents			
7-15	dry	1:1.17	1.7
	wet	1:1.96	
15-19	dry	1:2.1	1.6
	wet	1:3.3	

This shows that, although the ratios of risks for men and women and for dry and wet surfaces vary, the double ratio (odds-ratio) expressing the interactions between the two factors varies relatively little; it is about 1.6. This means that female drivers have a 60% higher accident risk on wet roads than one would expect, when combining their accident risk relative to that of male drivers with the relative accident risk for wet versus dry roads. However, this conclusion should be used with caution because of the high percentage of exposure with unknown drivers' sex for wet roads.

4.5.3.2 Interactions of Highway Factors

Table 4.5.3.2-1 and Fig. 4.5.3.2-1 show segment accident involvements, exposure and accident rates by highway alignment and grade. It is surprising that accident rates on grades are lower than on level segments. The rate in level curves is nearly three times as high as on straight level sections--and for single-vehicle accidents it is 8 times as high--but the rate in curves on grades is about 20% lower than on straight sections on grades--but for single-vehicle accidents it is still twice as high. Average speeds do not explain this pattern. However, the low average volume for curved level sections cautions against accepting the figures at face value. If by chance locations with extremely low volume had been selected on curved level segments, the total VMT on these sections would be underestimated, and rates overestimated.

Table 4.5.3.2-2 and Fig. 4.5.3.2-2 show intersection accident involvement and exposure. The pattern of rates is the same as in Table 4.5.3.2-1 but the differences among the rates tend to be smaller. However, again the average volume for intersections on level curves is quite low. The percentage of intersection maneuvers on level curves is 4% of all, that of VMT on curved level segments is 3%. This agreement suggests that the high accident rate on level curves is not entirely due to chance, though it might be affected by the low volume, or other factors related to the volume.

Table 4.5.3.2-3 presents the rates separate by type of accident. First, one can notice that the distributions of accident types for level and grade are similar. This suggests that grade has only a small effect on accident risk, because otherwise one would expect it to affect different accident types differently. Comparing straight and curved segments the pattern for single-vehicle accidents is plausible: their rates are higher in curves than on straight segments. Otherwise, it is surprising that the multi-vehicle accident rate is higher on curved than on straight level segments, though the traffic volume is much lower. However, the multi-vehicle accident rate for curves is lower than for straight segments on grades, though the traffic volumes are comparable.

TABLE 4.5.3.2-1
SEGMENT ACCIDENT INVOLVEMENTS AND
EXPOSURE GY HIGHWAY ALIGNMENT AND GRADE

	Straight					Curved				
	Accident Involvements	VMT (10^6)	Rate (per 10^6)	Avg. Speed (mph)	Avg. Volume (vph)	Accident Involvements	VMT (10^6)	Rate (per 10^6)	Avg. Speed (mph)	Avg. Volume (vph)
Level										
All accidents	1045	214	4.8	38	293	218	16.4	14	36	31
Single vehicle	147		0.7			96		6		
Multi-vehicle	898		4.2			122		7		
Grade										
All accidents	243	90	2.7	32	202	200	90	2.2	40	225
Single vehicle	40		0.44			79		0.9		
Multi-vehicle	203		2.2			121		1.4		

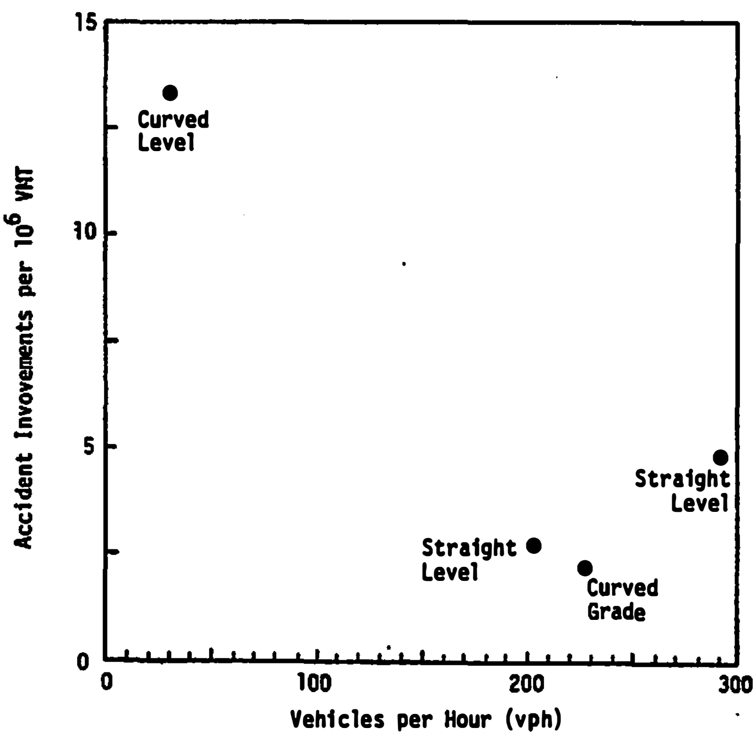


Figure 4.5.3.2-1. Segment accident involvement rates versus traffic volume, by highway alignment and grade.

TABLE 4.5.3.2-2
INTERSECTION ACCIDENT INVOLVEMENTS AND EXPOSURE BY HIGHWAY ALIGNMENT AND GRADE

	Straight				Curved			
	Accident Involvements	Maneuver (10^6)	Rate (per 10^6)	Avg. Volume (vph)	Accident Involvements	Maneuver (10^6)	Rate (per 10^6)	Avg. Volume (vph)
Level	491	1062	0.46	273	37	52	0.7	83
Grade	113	378	0.30	214	35	196	0.18	171

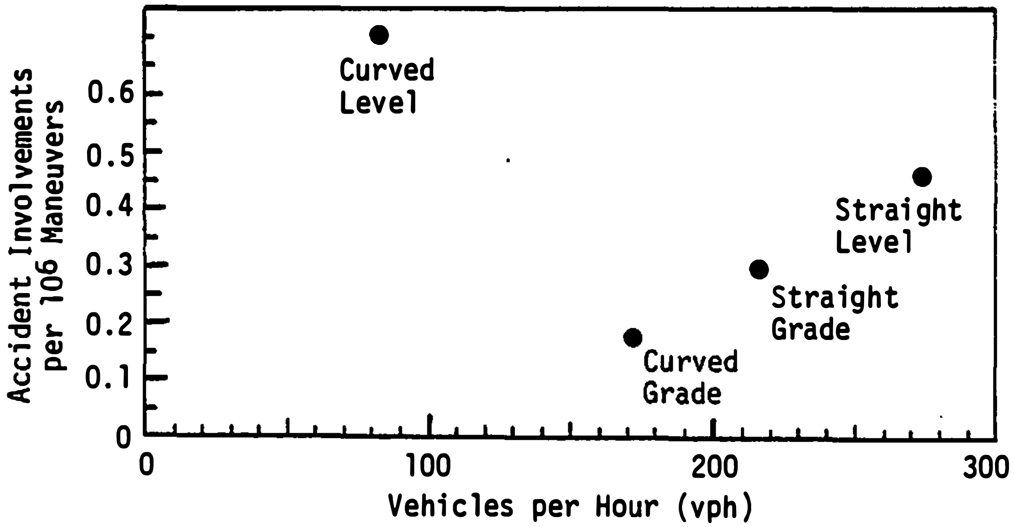


Figure 4.5.3.2-2. Intersection accident involvement rates by highway alignment and grade.

TABLE 4.5.3.2-3
DISTRIBUTION OF SEGMENT ACCIDENT INVOLVEMENTS BY TYPE, AND ACCIDENT INVOLVEMENT RATES, BY HIGHWAY ALIGNMENT AND GRADE

	Straight			Curved		
	Accident Involvements	Percent	Rate (per 10^6 VMT)	Accident Involvements	Percent	Rate (per 10^6 VMT)
Level						
Single-vehicle	147	14	0.7	96	44	6
Head-on	245	23	1.2	31	14	1.9
Rear-end	140	14	0.65	6	3	0.4
Other	513	49	2.4	85	39	5
		100			100	
Grade						
Single-vehicle	40	16	0.45	79	40	0.9
Head-on	60	25	0.65	38	19	0.4
Read-end	42	17	0.45	14	7	0.2
Other	101	42	1.1	69	34	0.8
		100			100	

We get some further insight by calculating the double ratios of rates. For single-vehicle segment accidents, we obtain:

	Ratio of Straight to Curved Rates	Double Ratio
Level	1:8.4	0.24
Grade	1:2.0	

For multi-vehicle segment accidents and for intersection accidents we get:

Segment Accidents	Ratio of Straight to Curved Rates	Double Ratio
Level	1:1.8	0.33
Grade	1:0.6	
Intersection Accidents		
Level	1:1.5	0.40
Grade	1:0.6	

Tables 4.5.3.2-4 and 4.5.3.2-5 show accident involvements, exposure and rates by highway alignment and surface. For both segment and intersection accidents the rates on wet are much higher than on dry surfaces. Rates in curves are lower than on straight roads with the exception of segment accidents in curves on wet roads, suggesting an interaction of wet surfaces and curves. The double ratio of rates is 6. For intersection accidents it is 0.9, differing only little from 1, which shows lack of interaction.

A close look at the data shows, however, that the high rate for segment accidents in curves is based on a very low exposure figure. The average traffic volume for the underlying observations is also very low. This suggests a possible underestimate of exposure on wet roads in curves. For segment accidents on straight sections, the percentage of exposure on wet surfaces is 11. For intersection accidents it is 14% on straight, 18% on curved sections. For segment accidents in curves, however, it is only 3%. It is highly unlikely that this low percentage is real. Therefore, we will assume a "model" that the percentage of wet surfaces is the same for segment and intersection exposure, and for straight and curved roads. In Tables 4.5.3.2-6 and 7, the exposure is distributed accordingly between dry and wet roads. The overall pattern of the rates is not changed, but the extreme value is reduced. The double ratio of rates for segment accidents is 1.8, for intersection accidents 1.2. Both show an interaction between curvature and surface.

TABLE 4.5.3.2-4
SEGMENT ACCIDENT INVOLVEMENTS AND EXPOSURE BY HIGHWAY ALIGNMENT AND SURFACE

Highway Surface	Straight				Curved			
	Accident Involvements	VMT (10^6)	Rate (per 10^6)	Avg. Volume (vph)	Accident Involvements	VMT (10^6)	Rate (per 10^6)	Avg. Volume (vph)
Dry	1071	254	4.2	239	308	102	3.1	200
Wet	217	30.3	7	511	110	3.4	32	54

TABLE 4.5.3.2-5
INTERSECTION ACCIDENT INVOLVEMENTS AND EXPOSURE BY HIGHWAY ALIGNMENT AND SURFACE

Highway Survace	Straight				Curved			
	Accident Involvements	Maneuver (10^6)	Rate (per 10^6)	Avg. Volume (vph)	Accident Involvements	Maneuver (10^6)	Rate (per 10^6)	Avg. Volume (vph)
Dry	502	1234	0.40	257	58	204	0.28	164
Wet	102	206	0.50	258	14	44	0.32	103

TABLE 4.5.3.2-6
SEGMENT ACCIDENT INVOLVEMENTS AND EXPOSURE BY HIGHWAY ALIGNMENT AND SURFACE.
(Frequency of wet surface "modelled")

Highway Surface	Straight				Curved			
	Accident Involvements	VMT (10^6)	Rate (per 10^6)	Avg. Volume (vph)	Accident Involvements	VMT (10^6)	Rate (per 10^6)	Avg. Volume (vph)
Dry	1071	252	4.2	239	308	93	3.3	200
Wet	217	33	6.6	511	110	12	9.2	54

TABLE 4.5.3.2-7
INTERSECTION ACCIDENT INVOLVEMENTS AND EXPOSURE BY HIGHWAY ALIGNMENT AND SURFACE.
(Frequency of wet surface "modelled")

Highway Survace	Straight				Curved			
	Accident Involvements	Maneuver (10^6)	Rate (per 10^6)	Avg. Volume (vph)	Accident Involvements	Maneuver (10^6)	Rate (per 10^6)	Avg. Volume (vph)
Dry	502	1274	0.39	257	58	220	0.26	164
Wet	102	166	0.61	258	14	20	0.50	103

4.5.3.3 Interaction of Three Highway Factors

We study the interaction of highway alignment, grade and surface condition. Tables 4.5.3.3-1 and 2 show accident involvements, exposure, accident rates, and also speed (for segment exposure) and traffic volume for the observed exposure. No obvious pattern appears for speed; therefore, it is not further considered. Traffic volume and accident rates, however, appear related: rates are high for low volumes. This relation is clearly recognizable in Figs. 4.5.3.3-1 and 2.

It was already noticed in Section 4.5.2 that the relative frequencies of exposure on dry and wet surfaces differed between segment exposure and intersection exposure. A closer inspection of Tables 4.5.3.3-1 and 2 confirms this; it also shows that the highest accident rates are not due to unusually larger numbers of accidents, but to unusually small estimated exposures. Table 4.5.3.3-3 shows the relative frequencies of exposure on dry and wet surfaces. For intersection maneuvers it is relatively constant: about 15% wet surfaces; only for level curves is it higher. For segment exposure it is 13% for straight level segments, very much for other surfaces. It is possible that there are real differences in the frequencies of dry and wet highway surfaces among parts of the study area.* The terrain may be related to the relative frequency of intersections in relation to highway miles, to the occurrence of grades, of curves, and the combination of grades and curves. However, since wet surfaces are relatively infrequent, estimates of the exposure on wet surfaces are subject to relatively large variations. Therefore, no conclusions can be drawn from the observed differences. If detailed weather data were available, one could use them in connection with detailed data on highway characteristics to "model" the frequency of wet surfaces for the various highway characteristics. Here we will use only an extremely simplified "model," namely assume that the relative frequency of wet and dry surfaces is the same for all highway conditions. For intersection exposure, the frequency of wet surfaces is 15%, for segment exposure 8%. Though the number of intersection and segment sites is nearly equal, observation times are not. At a segment site, the observation period was half an hour, at intersection sites half an hour for each approach. The typical

*In Connecticut, the amount of precipitation is correlated to the elevation of a location. See, J.W. Wilson and M.A. Atwater, Storm rainfall variability over Connecticut, J. Geophys. Res., Vol. 77, 1972.

TABLE 4.5.3.3-1
SEGMENT ACCIDENT INVOLVEMENTS AND EXPOSURE BY
HIGHWAY ALIGNMENT, GRADE AND SURFACE CONDITION

	Straight					Curved				
	Accident Involvement	VMT (10^6)	Rate (per 10^6)	Avg. Speed (mph)	Avg. Volume (vph)	Accident Involvement	VMT (10^6)	Rate (per 10^6)	Avg. Speed (mph)	Avg. Volume (vph)
Dry										
Level	881	186	4.8	39	254	164	16.4	10	35	31
Grade	190	86	2.2	32	206	144	86	1.7	40	232
Wet										
Level	164	28.6	6	32	547	54	--	--		
Grade	53	2.2	24	30	40	56	3.4	16	38	54

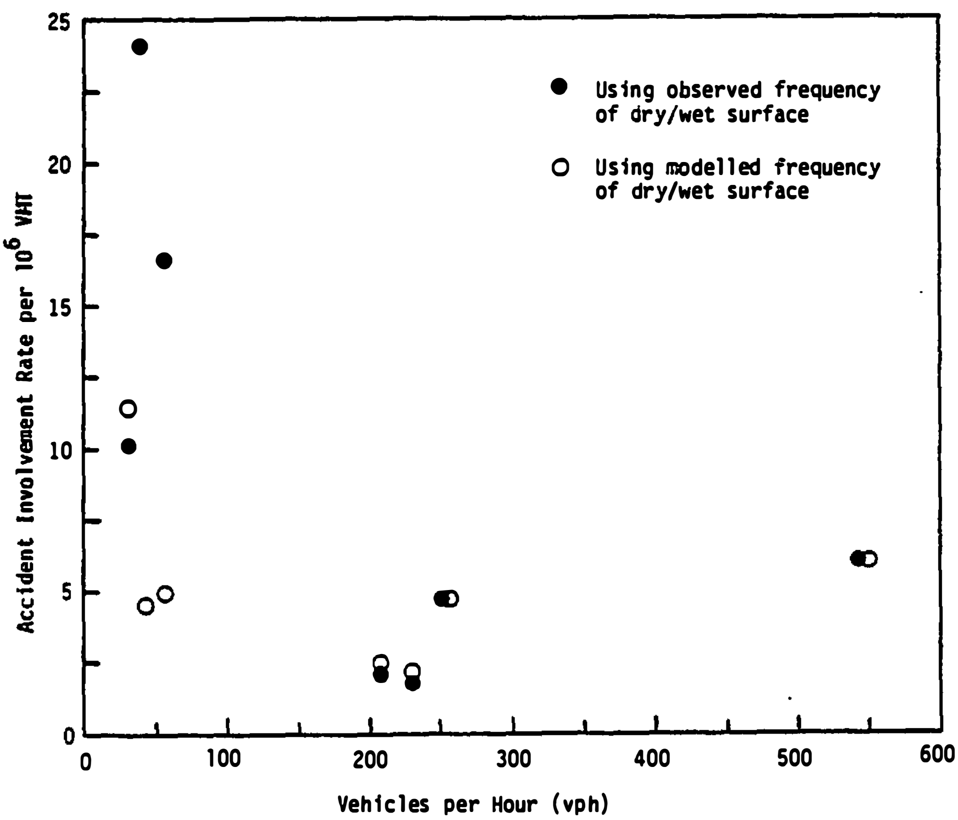


Figure 4.5.3.3-1. Segment accident involvement rates from Tables 4.5.3.3-1 and 4.5.3.3-4 versus traffic volume.

TABLE 4.5.3.3-2
INTERSECTION ACCIDENT INVOLVEMENTS AND EXPOSURE BY HIGHWAY ALIGNMENT, GRADE AND SURFACE CONDITION

	Straight				Curved			
	Accident Involvements	Maneuvers (10^6)	Rate (per 10^6)	Avg. Volume (vph)	Accident Involvements	Maneuvers (10^6)	Rate (per 10^6)	Avg. Volume (vph)
Dry								
Level	406	906	0.45	293	30	38	0.8	96
Grade	96	328	0.30	158	28	166	0.17	179
Wet								
Level	85	150	0.55	155	7	14.4	0.48	50
Grade	17	50	0.34	584	7	30	0.24	129

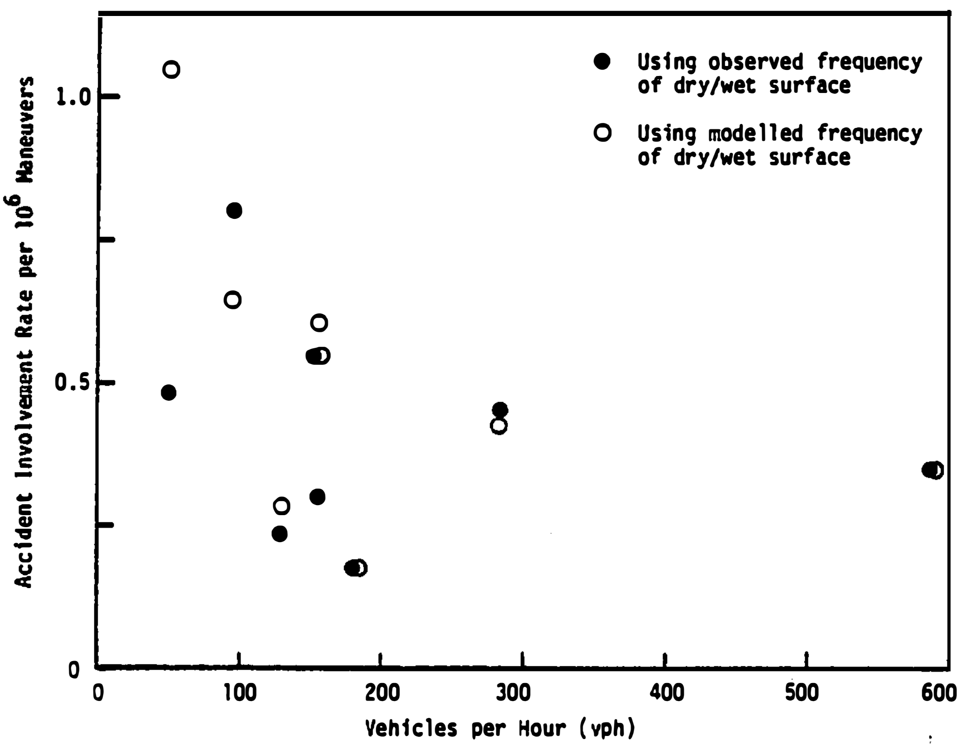


Figure 4.5.3.3-2. Intersection accident involvement rates from Tables 4.5.3.3-2 and 4.5.3.3-5 versus traffic volume.

intersection has three or four approaches (some have only two), however, sometimes two approaches could be observed simultaneously. Therefore, total observation time at intersections is 2.5 to 3 times as much as at segment locations. Consequently, we weight intersection and segment frequencies of wet surfaces as 7:3 and use 13% as representative average. Tables 4.5.3.3-4 and 5 show the resulting "modelled" exposures, and the resultant accident rates. They are also shown in Figs. 4.5.3.3-1 and 2. For segment accidents, the relation between traffic volume and accident rate has become less pronounced, for intersection accidents the changes do not seem to have a pattern.

For segment accidents, separate rates were also calculated for single vehicle, and for multi-vehicle accidents. They are shown in Table 4.5.3.3-6. The numbers of intersection accidents for some categories were so low that a break-down by type would not have been meaningful.

Various models were fitted to the "modelled" accident rates. The basic structure was a log-linear model.

$$\log r = a_o + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_{12} x_1 x_2 + \dots + b \ln v,$$

where

$$x_1 = \begin{cases} -1 & \text{straight} \\ +1 & \text{curved} \end{cases}$$

$$x_2 = \begin{cases} -1 & \text{level} \\ +1 & \text{grade} \end{cases}$$

$$x_3 = \begin{cases} -1 & \text{dry} \\ +1 & \text{wet} \end{cases}$$

$$v = \text{traffic volume.}$$

This model is equivalent to

$$r = A_o A_1^{x_1} \cdot A_2^{x_2} \cdot A_3^{x_3} \cdot A_{12}^{x_1 x_2} \dots v^b$$

First, a set of models was fitted without the term b ln v. At a first level, all main effects (terms with x_1, x_2, or x_3) were included, at a second level the strongest first-order interaction term was added, and at a third level all three first-order interactions included. It turned out that the last two first-order interactions (and also the second-order interaction) were small compared with the terms first used.

Secondly, models with the main effects and the strongest first-order interaction terms and the term b ln v were fitted.

Table 4.5.3.3-7 shows the effects obtained from the models, and Table 4.5.3.3-8 shows the accident rates obtained from the models, together with the actual (modelled) values.

All log-linear models (A in Table 4.5.3.3-7) show the same patterns of effects: the main effects of the factors are strong (with one exception: alignment for intersection accidents), and also the intersection curved x grade. The other first order interactions are much weaker, and the second order interaction also (its coefficient differs relatively little from 1). Since the rates result from dividing actual accident numbers by "modelled" VMT which in turn are based on estimates from a sample, their statistical properties are unknown, and statistical tests can not be applied.

If one compares the magnitudes of the four strongest effects, one finds that they are similar for all accident types, with two exceptions: the effect of alignment is very strong in single vehicle accidents, and it is very weak for intersection accidents. This is very plausible.

For the models including the traffic volume terms, the effects of the other major factors are either unchanged, or become smaller. The effect of traffic volume is difficult to interpret. It is not surprising that the rate for single-vehicle accidents is much higher at low volumes than at high volumes, because "incidents" which result in a single-vehicle accident if no other vehicles are around can become multi-vehicle accidents if other vehicles are there. It is not plausible, however, that the risks of multi-vehicle segment accidents, and of intersection accidents are higher at low volumes than at higher volumes. This may have various reasons. One may be imperfection of the sample: if the volume at certain sites was by chance low, VMT or maneuvers at such sites are underestimated, and accident rates become exaggerated. Another reason could be that other factors which increase the accident risk are correlated with low traffic volumes. A decisive analysis requires volume information also for accidents, not only for exposure.

Table 4.5.3.3-8 compares the modelled rates with the actual rates. Models with the main effects only represent the overall pattern of the rate, but the quantitative differences can be large. Adding the strongest interaction improves the representation somewhat, but not consistently. Adding further the

TABLE 4.5.3.3-3
SEGMENT ACCIDENT INVOLVEMENTS AND "MODELLED" EXPOSURE BY HIGHWAY ALIGNMENT, GRADE, AND SURFACE CONDITION

	Straight				Curved			
	Accident Involvements	VMT (10^6)	Rate (per 10^6)	Avg. Volume (vph)	Accident Involvements	VMT (10^6)	Rate (per 10^6)	Avg. Volume (vph)
Dry								
Level	881	186	4.8	254	164	14.2	12	31
Grade	190	78	2.4	206	144	78	1.8	232
Wet								
Level	164	28.6	6	547	54	2.2	24	--
Grade	53	11.8	4.5	40	56	11.6	4.8	54

TABLE 4.5.3.3-4
INTERSECTION ACCIDENT INVOLVEMENTS AND "MODELLED" EXPOSURE BY HIGHWAY ALIGNMENT, GRADE AND SURFACE CONDITION

	Straight				Curved			
	Accident Involvements	Maneuvers (10^6)	Rate (per 10^6)	Avg. Volume (vph)	Accident Involvements	Maneuvers (10^6)	Rate (per 10^6)	Avg. Volume (vph)
Dry								
Level	406	924	0.44	293	30	46	0.65	96
Grade	96	328	0.30	158	28	170	0.16	179
Wet								
Level	85	138	0.60	155	7	6.8	1.05	50
Grade	17	50	0.34	584	7	26	0.28	129

TABLE 4.5.3.3-5
SEGMENT ACCIDENT INVOLVEMENT RATES (PER 10^6 VMT) FOR
SINGLE- AND MULTI-VEHICLE ACCIDENTS

	Straight		Curved	
	Single-Vehicle	Multi-Vehicle	Single-Vehicle	Multi-Vehicle
Dry				
Level	0.70	4.0	5.0	6.5
Grade	0.34	2.0	0.8	1.1
Wet				
Level	0.53	5.0	11.0	14.0
Grade	0.75	3.8	1.8	3.1

TABLE 4.5.3.3-6
EFFECTS OF HIGHWAY FACTORS ON ACCIDENT INVOLVEMENT RATES

Factor	Segment Accidents					Intersection Accidents	
	All	Single-vehicle		Multi-vehicle			
	A	A	B	A	B	A	B
Straight/curved	1.69	5.0	2.6	1.18	1.21	1.03	0.91
Level/grade	0.33	0.37	0.43	0.35	0.37	0.37	0.43
Dry/wet	1.90	1.66	1.10	1.81	1.54	1.44	1.43
Curved x grade	0.54	0.42	0.87	0.56	0.85	0.67	0.67
Curved x wet	1.24	1.36		1.27		1.13	
Grade x wet	1.16	1.27		1.41		0.97	
3-factor interaction	0.96	0.81		0.98		1.06	
volume (500/50)			2.6		1.8		1.4

Note: The numbers show, e.g., that the segment accident risk on curves is 1.69 times that on straight sections. For intersections, they show, e.g., that for grades in curves it is only 0.54 of that in level curves, after separate consideration of the effects of alignment and grade.

"A" shows estimates of the effects using log-linear models including all factors.

"B" shows the effects resulting from models using only the four strongest factors and traffic volume.

The figures for "volumes" show how much higher the accident rates for 50 vph are than for 500 vph.

TABLE 4.5.3.3-7
"ACTUAL" (A) AND MODELLED (1-4) ACCIDENT RATES*

	Straight					Curved				
	A	(1)	(2)	(3)	(4)	A	(1)	(2)	(3)	(4)
DRY										
Level										
Segment Accidents										
All	4.8	4.8	4.8	---	4.8	11	8	15	---	12
Single-	0.7	0.7	0.7	0.7	0.7	5	3.5	8	5	6
Multi-	4.0	4.0	4.0	4.0	4.0	6	5	8	9	6
Intersection Accidents	0.44	0.44	0.44	0.44	0.44	0.65	0.46	0.7	0.7	0.6
Grade										
Segment Accidents										
All	2.4	1.6	2.9	---	2.5	1.8	2.7	2.7	---	1.8
Single-	0.4	0.26	0.6	0.37	0.48	0.75	1.3	1.3	0.8	0.75
Multi-	2.0	1.4	2.5	1.8	2.1	1.1	1.6	1.6	1.8	1.1
Intersection Accidents	0.3	1.18	0.27	0.24	0.28	0.16	0.18	0.18	0.28	0.16
WET										
Level										
Segment Accidents										
All	6.0	9.0	9.0	---	5.0	24	16	28	---	24
Single-	0.55	1.2	1.2	0.55	0.7	11	6	14	---	11
Multi-	5.0	8.0	8.0	5.0	5.0	14	9	16	---	14
Intersection Accidents	0.6	0.65	0.65	1.0	0.6	1.0	0.65	1.0	0.75	1.0
Grade										
Segment Accidents										
All	4.5	3.0	5.5	---	4.5	4.8	5.0	5.0	---	5.0
Single-	0.75	0.43	1.0	0.8	0.75	1.8	2.2	2.2	1.6	2.2
Multi-	3.8	2.6	4.8	4.2	3.8	3.1	3.2	3.2	4.0	3.2
Intersection Accidents	0.34	0.25	0.39	0.37	0.34	0.28	0.26	0.26	0.28	0.26

*"Actual" accident rates are based on "modelled" frequencies of wet highways. Model (1) uses only the main effects of the three highway factors, model (2) the main effects and the strongest interaction (curved x grade), model (3) the main effects, the strongest interaction and traffic volume, and (4) all effects except the second order interaction. All rates are based on that for straight, level dry roads.

volume terms gives a very good fit for single-vehicle accidents, some improvement for multi-vehicle segment accidents, and little improvement for intersection accidents.

Finally, the model including all terms but the second-order interaction represents the rates quite well. This confirms that the interaction of all three factors has practically no effect--at least not within the error limits of this study.

4.6 Vehicle Factors

4.6.1 Introduction

Many vehicle characteristics can influence the probability of an accident: braking capability, tire characteristics, steering response, response to wind gusts, to name but a few. Some of them are design characteristics, others depend on the status of maintenance, e.g., brakes and tires, and some may be influenced by the weight and distribution of occupants and load. Many of these characteristics are not easily quantified on a one-dimensional scale, and practically all are very difficult to obtain.

Therefore, usually only gross vehicle characteristics are studied--some of which may be related to factors which have a causal effect. Car size is such an obvious factor. Sometimes "sports" cars are recognized as a separate category. Several classifiers for size are in use: weight, wheelbase, and interior volume. We use weight, because it was directly obtainable from the New York Motor Vehicle Department's registration records. Considering the distribution of weights in our exposure sample, we categorized cars somewhat arbitrarily into four groups: less than 2250 lbs, 2250-2999 lbs, 3000-3749 lbs, and 3750 lbs or more. This corresponds roughly to subcompacts, compacts, intermediate and full size cars (using pre-downsizing standards).

Another obvious factor is vehicle age. Some characteristics deteriorate with age. Though maintenance and repair can prevent, at least to some extent, deterioration, it appears plausible that cars are being less well maintained with increasing age.

However, age has probably stronger indirect effects: owners, drivers and vehicle use tend to change with age, and these factors will also influence the accident risk. To some extent one can control for this by classifying exposure, however, some effects are likely to remain.

Vehicle information could not be obtained for 22% of the segment exposure, and for 17% of the intersection exposure. This is the combined effect of illegible license plates (e.g., at night) and license numbers which could not be recognized by the New York Department of Motor Vehicles. In 16% of the accident cases vehicle data could not be retrieved. This can be due to illegible license numbers on the copies of the accident reports which were used, and also due to vehicles no longer being registered when the file was searched.

4.6.2 Vehicle Age and Weight

Figure 4.6.2-1 shows the accident involvement rates per exposure by vehicle age. Age "0" combines cars of the current model year and those of the next model year, the ages 10 and 11, and 12, 13 and 14, and those 15 and over are grouped. Rates for "0" age are much higher than for ages of one or a few years. This could be an artifact, because cars of the current and next model year were being sold during the exposure data collection period, and the differing period of accident data collection. Therefore, we will ignore these points.

For single-vehicle segment accidents a clear pattern appears: the rate is essentially constant for ages 1 through 8 years, and also essentially constant, but much higher--about double--for cars of 9 years and over. For multi-vehicle segment accidents the pattern is similar. However, there is a slight increasing trend in the range of 1 through 9 years; the points would also be compatible with an increasing trend from 1 through 11 years, considering the large year-to-year fluctuations. For intersection accidents, there is an increasing trend from 1 through 9 years, and a drop afterwards.

The figure shows, also, single-vehicle and multi-vehicle accident rates per 10^6 VMT for passenger cars in North Carolina in 1979.* Single-vehicle accident rates increase from 1 through 11 years (the drop at 9 years is a peculiarity of the 1970 model year; it appears also in an earlier study of 1974 North Carolina accidents and exposure) and level off; multi-vehicle accident rates increase from 1 through 10 years and level off thereafter. In North Carolina, the single-vehicle accident rates for "old" cars is also about twice as high as for young cars. For multi-vehicle accidents, it is about 60% higher. In Ulster County, the rate for multi-vehicle accidents is also about 60% higher for "old" cars than for young cars.

---

*J.R. Stewart, C. Lederhaus Carroll, Annual Mileage Comparisons and Accident and Injury Rates by Make, Model. Highway Safety Research Institute, University of North Carolina, October 1980.

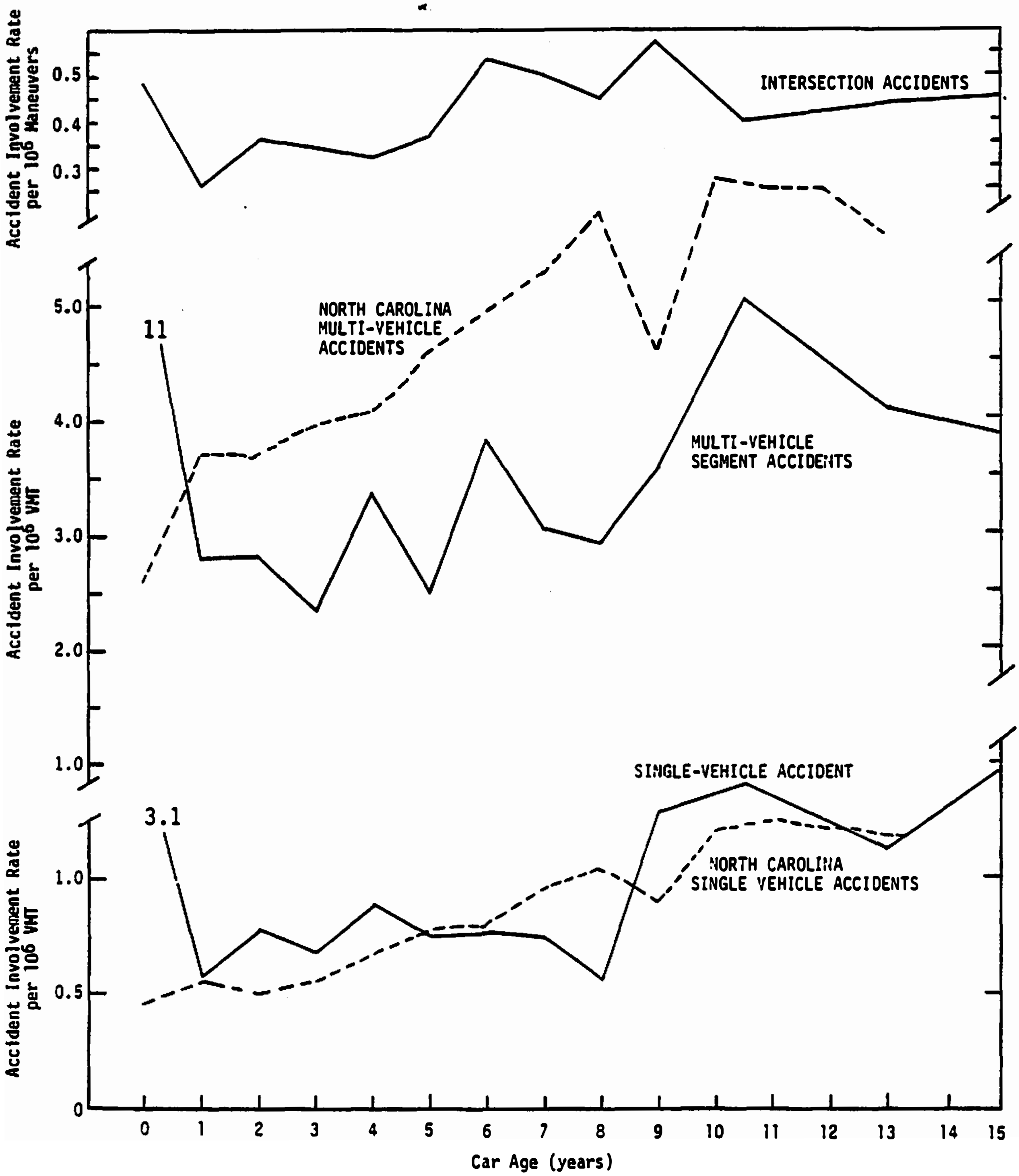


Figure 4.6.2-1. Accident involvement rates by car age.

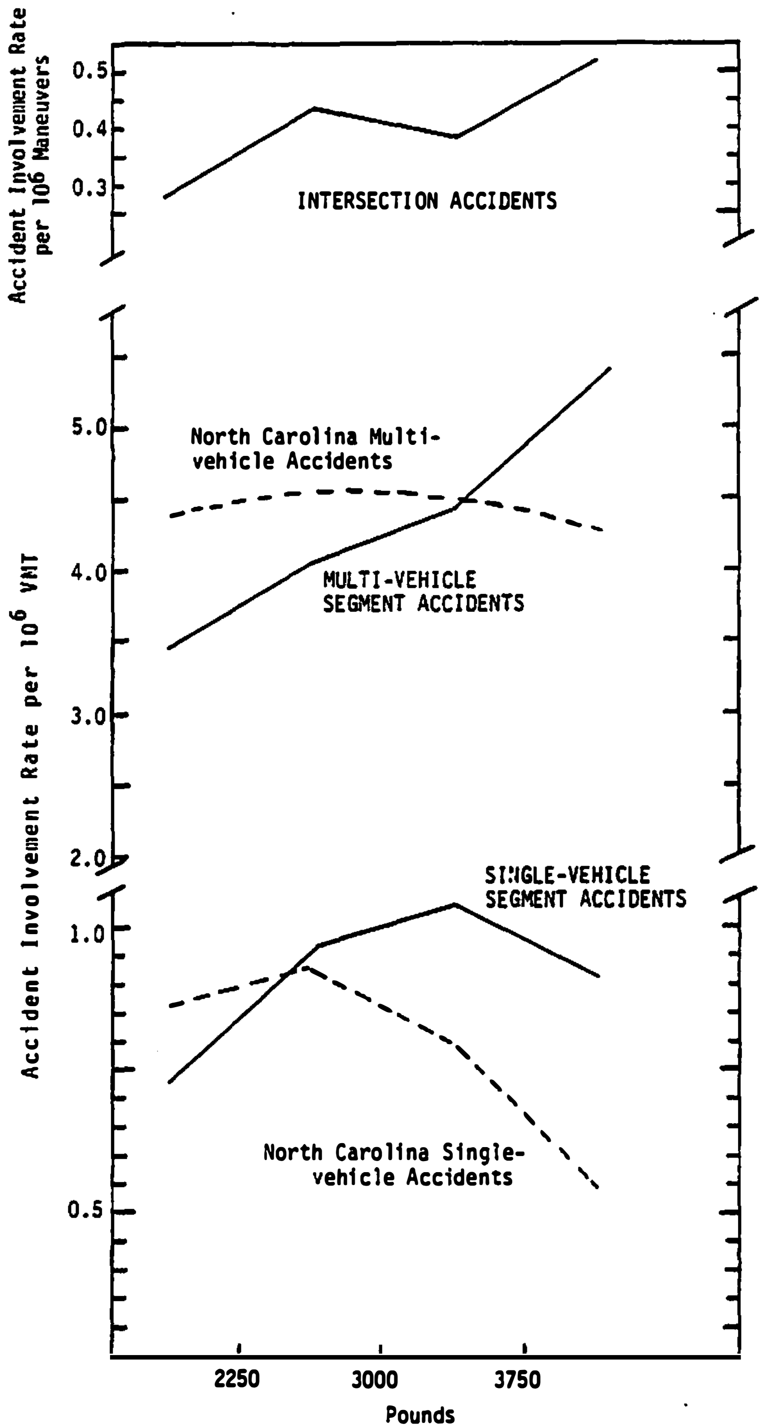


Figure 4.6.2-2. Accident involvement rates by vehicle weight class.

Thus, North Carolina and Ulster County data agree in the relations of accident rates for "old" and "young" cars. However, the time trends differ. In North Carolina, rates change gradually with age, in Ulster County, abrupt changes occur at a car age of about 10 years. Since segment accident rates increase suddenly at this time, but intersection accidents drop, one suspects that there is a fairly rapid change in car use when it becomes 10 years old. Such a use pattern may be specific to a small area, such as a County. If the patterns differ among parts of a state, the strong changes at certain ages will disappear in the average, and a smooth trend with age, as in North Carolina, will appear.

The Ulster County data do not rule out the possibility of a physical effect of vehicle age, because the rates for multi-vehicle accidents show a trend of increase with car age.

Because of the suspicion of age-related difference in car use, and related pre-crash factors, some interactions will be explored in the next section.

Figure 4.6.2-2 shows accident involvement rates in relation to car weight. Surprisingly, involvement rates for all types of accidents tend to increase with car weight. The only exception is that heavy cars have a lower single-vehicle accident rate than middle-weight cars. The figure also shows the North Carolina accident rates; they show a quite different pattern: essentially no relation between multi-vehicle accident rates and car size, and only little difference among the single-vehicle accident rates for subcompact, compact and intermediate cars.

Since one would expect intersection accident rates to be largely independent of vehicle weight, one can again suspect differences in use patterns and consequently other pre-crash factors.

The exposure observations are for 1981, when small cars represented a larger proportion of the car population than in 1980. Therefore, the rate of 1980-81 accidents relative to 1981 exposure is decreased for light cars, increased for heavy cars. This effect, however, is much smaller than the differences found between light and heavy cars.

4.6.3 Interactions with Car Age

The most obvious vehicle use factor is urban/rural driving. Our data base does not contain information on urban/rural environment (and even if available, it tends to be extremely crude). However, the ratio of intersection maneuvers to VMT can be used as an indicator of the relative frequency of urban/rural driving: the higher it is, the higher the proportion of urban driving, where the number of intersections for highway mile is high.

Figure 4.6.3-1 shows the ratio of intersection maneuvers to VMT by car age. With the exception of the 10- to 11-year old group, there is a consistent trend

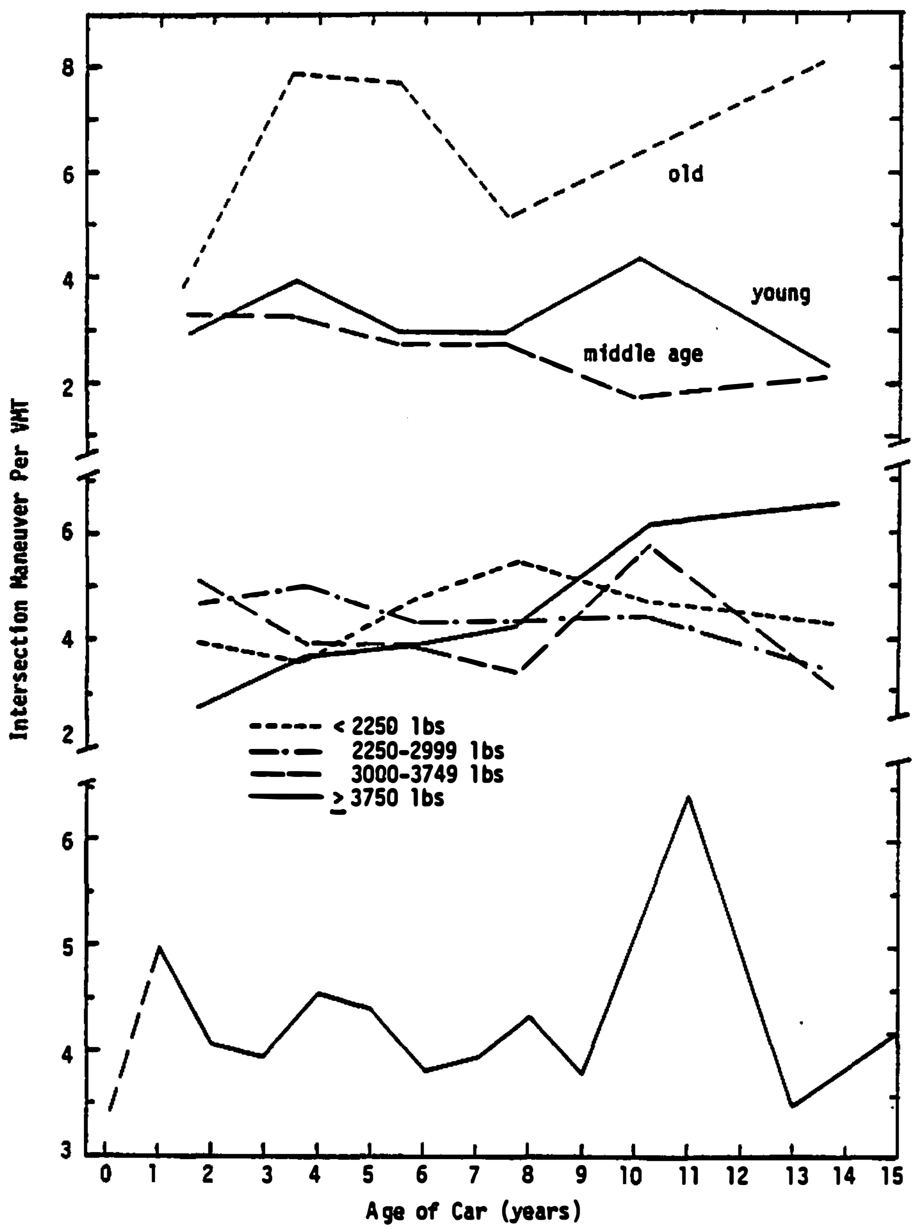


Figure 4.6.3-1. Intersection maneuvers per VMT by car age.

toward more rural driving with increasing car age. The figure shows also the same relation, disaggregated by driver age class, and by car weight class. Disaggregation by driver age class gives no clear picture: possibly urban driving increases with car age for old drivers, but decreases for middle age drivers. In the disaggregation by car weight, one trend is obvious: for heavy cars (≥3750 lbs) urban driving increases with age. For the 3000-3749 lbs class there is a weak opposite trend. For the lighter weight classes no clear trend is apparent.

Figure 4.6.3-2 shows the accident risks by car age separately for the three driver age groups. For single-vehicle accidents, the risks fluctuate greatly with vehicle age, but they tend to be higher for older cars, with middle age and old drivers. For young drivers, the trend may be opposite. The pattern of constant lower rates for younger cars, and constant higher rates for older cars does not appear. The multivehicle segment accident rate for old drivers is practically constant. For young and middle age drivers it appears constant except for the oldest cars. These separate curves do not fit the pattern shown in Fig. 4.6.1-1. Intersection accidents show rates increasing with car age for middle age and old drivers, no, or a slightly decreasing trend for young drivers. Overall, there is some similarity with the pattern of Fig. 4.6.2-1.

Figure 4.6.3-3 shows the risk by car age disaggregated by car weight. For single-car accidents, the risks for older cars tend to be higher than for younger cars in all weight classes, but the scatter of the points is too large to assess whether there is a continuous trend with age, or constant low rates for young, constant high rates for old cars. For multi-vehicle segment accidents the pattern is similar; however, the points do suggest a constant rate for young cars, and a constant higher rate for old cars. This is similar to the patterns observed when disaggregating by driver age.

For intersection accidents, heavy cars show no change in risk with car age, possibly a declining trend; very light cars show a declining trend, the two middle weight classes an increasing trend.

In sum, disaggregation by driver age and car weight gives the same overall trends of accident risk with car age as found in Section 4.6.2. However, some age groups and weight groups appear to differ. Also, the ratio of intersection maneuvers to VMT changes clearly with vehicle age, and with driver age and vehicle weight. These observations suggest strongly that the observed relations of accident risks to car age are not causal relations, but incidental to other use factors which are related to car age.

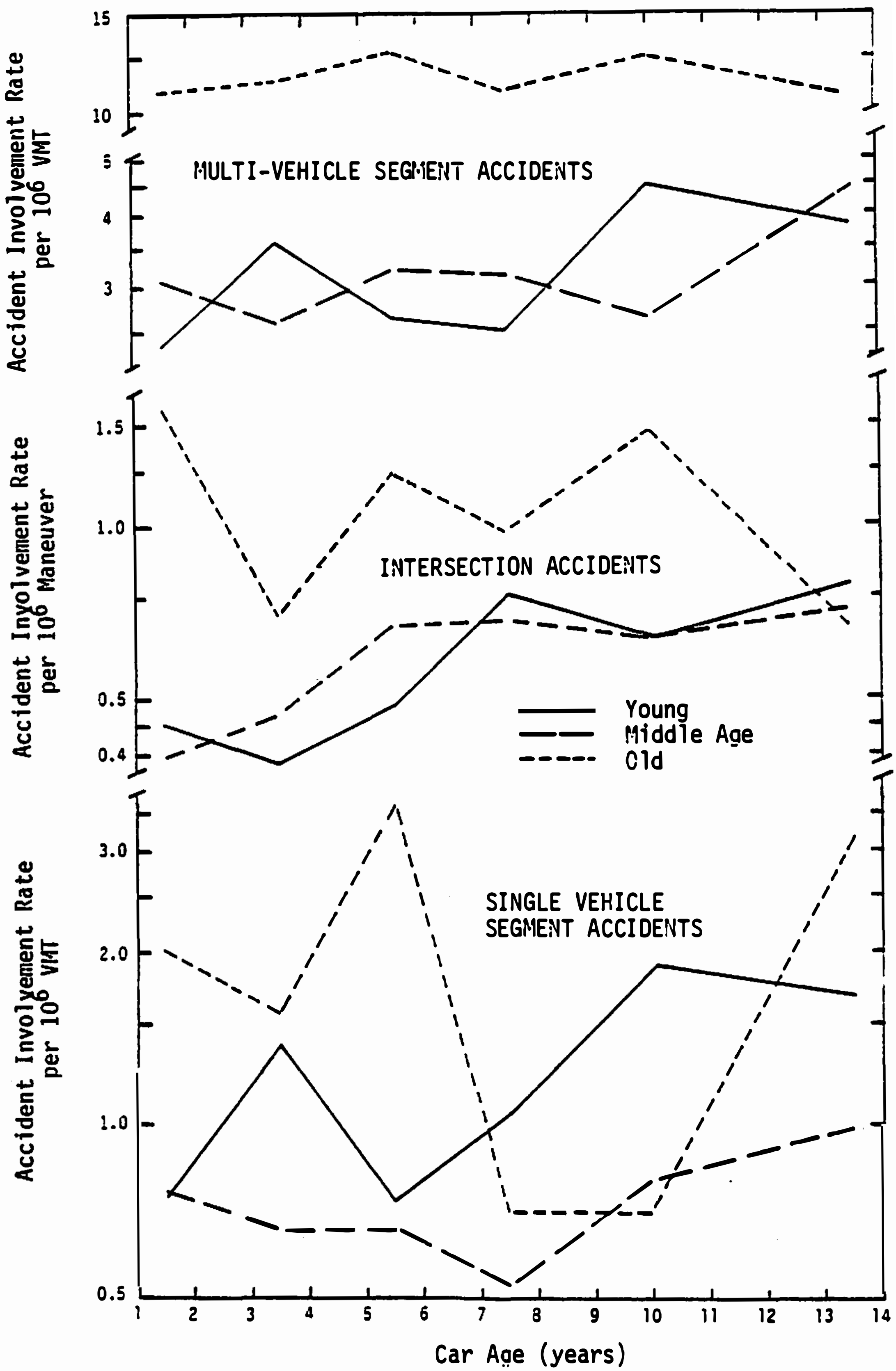


Figure 4.6.3-2. Accident involvement rates versus car age, by driver age.

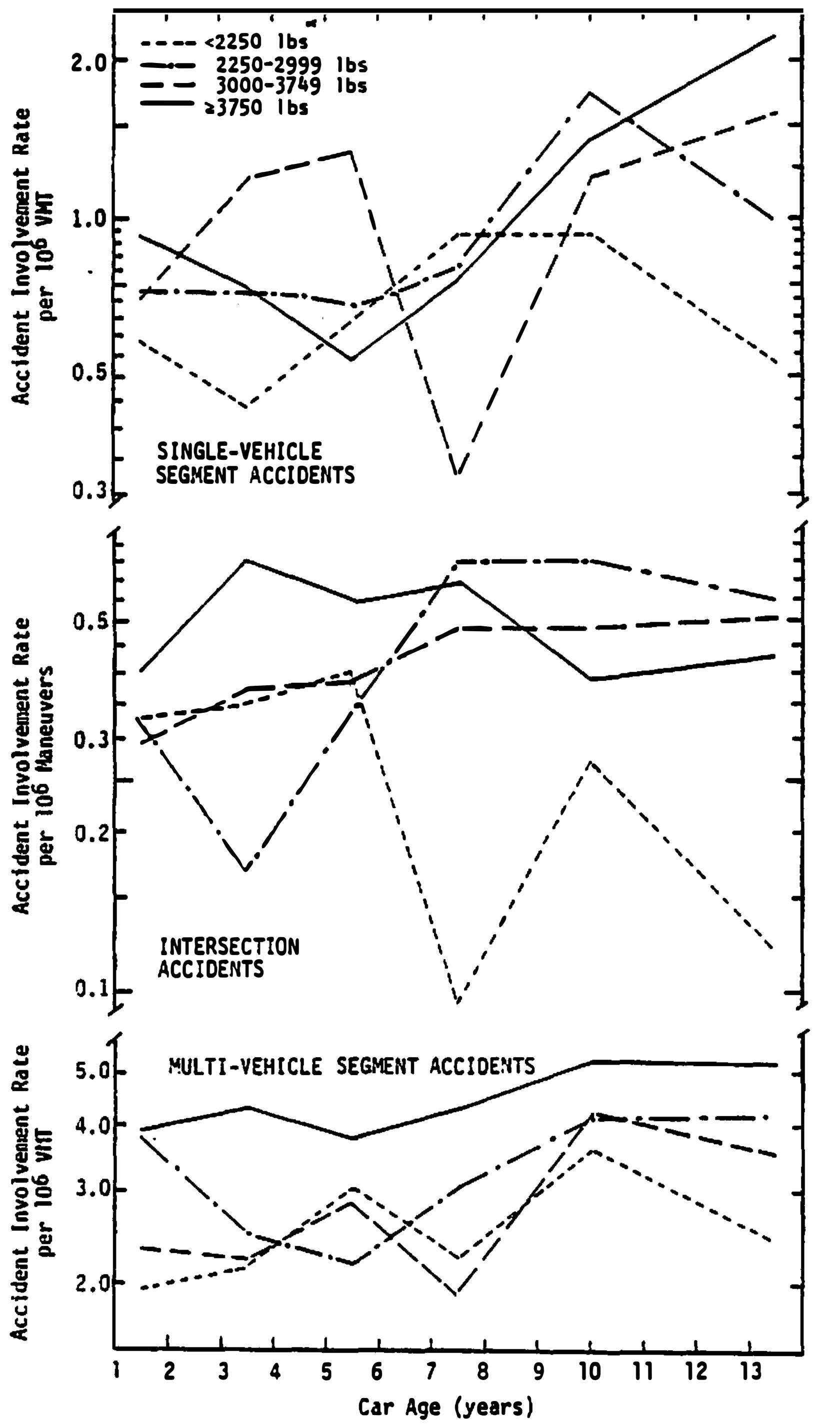


Figure 4.6.3-3. Accident involvement rates versus age of car, by car weight.

4.6.4 Interactions with Car Weight

Figure 4.6.4-1 shows the ratio of intersection maneuvers to VMT by car weight. It is higher for the heavier cars than for the lighter only, suggesting more urban driving for the heavier cars. For young and middle age drivers the same holds; however, old drivers use the very light cars more often in urban environments than the other weight groups.

Figures 4.6.4-2a through c show the accident involvement rates by car weight and driver age classes. For single-car segment accidents, rates for young drivers increase strongly with vehicle weight; for old drivers, they decrease similarly strongly. For middle age drivers there is some decline for the heavier weights--the relation between rate and weight is practically the same as in North Carolina (Fig. 4.6.2-2).

For multi-car segment accidents, the rate for young drivers increases again strongly with vehicle weight. For old drivers, the pattern is not clear, and for middle age drivers the rate is practically constant--again the same as in North Carolina.

The intersection accident rates for middle age and old drivers fluctuate greatly, with some tendency to increase with weight. The rate for young drivers is higher for the higher weights than for the lower weights.

Overall, it appears that for young drivers the rates for all accidents increase with vehicle weight. For middle age and old drivers the rates for multi-vehicle accidents appear not to vary systematically with weight; single-vehicle accident rates decrease with vehicle weight.

Since the relations between weight and accident rates depend on driver age, they can not reflect the effects of vehicle characteristics only, but must include the effects of other factors which are related to vehicle weight and driver age.

One factor which could interact with vehicle characteristics is highway curvature. Figure 4.6.4-3 shows the relations between accident rates and vehicle weight by highway curvature. In general, the relations parallel each other. There are two exceptions: for intersection accidents in curves, no relation between weight and rate is apparent, and for single-vehicle accidents in curves, the rate for heavy cars is much lower than to be expected. There is no suggestion that light cars might have higher accident risks in curves than heavier cars.

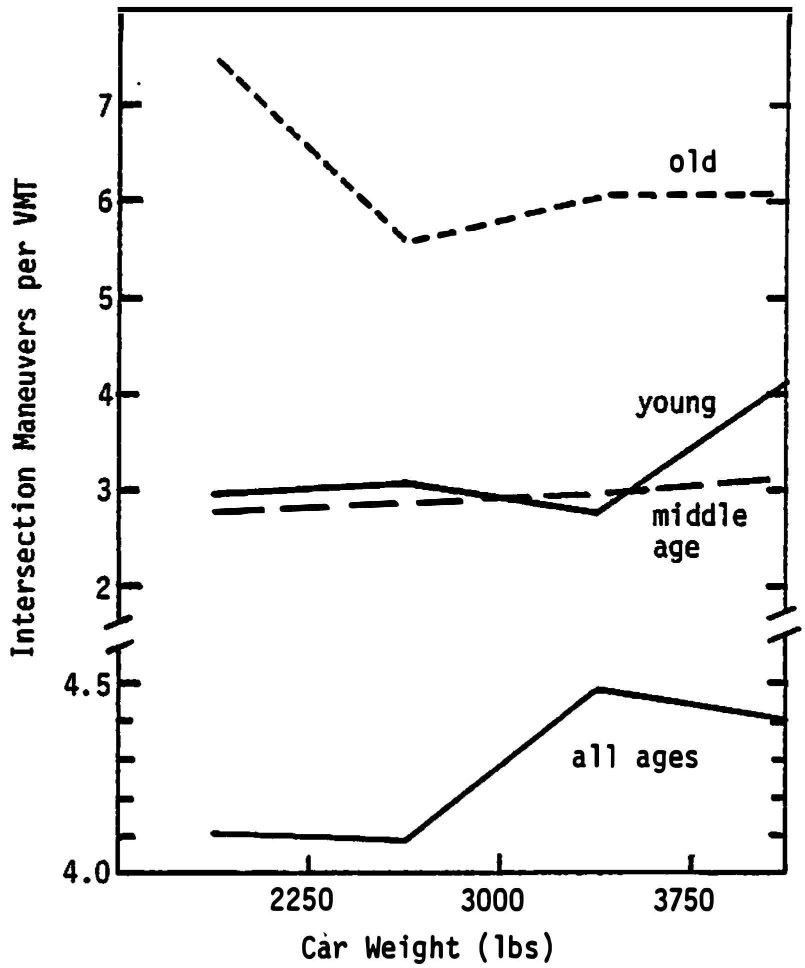


Figure 4.6.4-1. Intersection maneuvers per VMT by car weight and driver age.

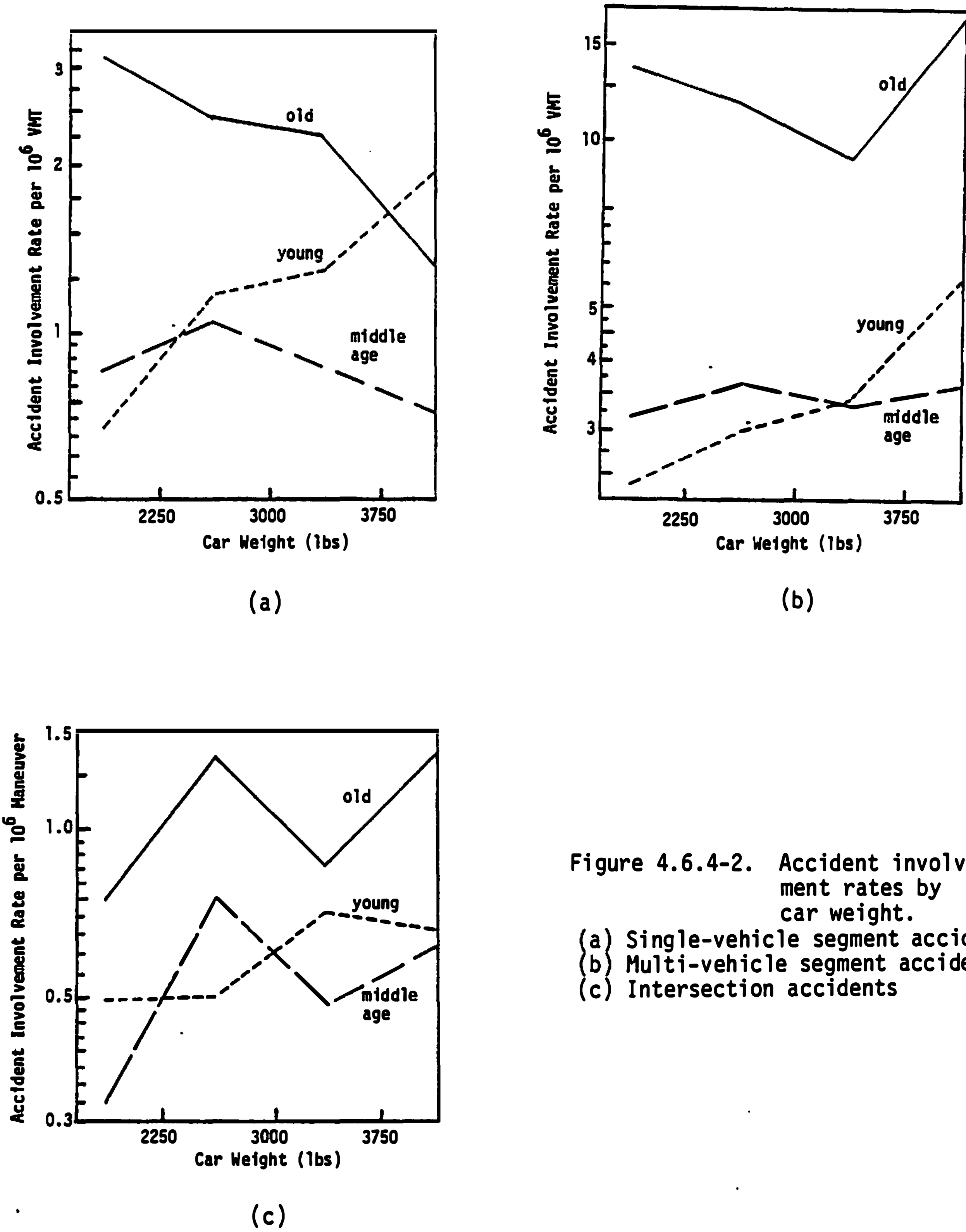


Figure 4.6.4-2. Accident involvement rates by car weight.
(a) Single-vehicle segment accidents
(b) Multi-vehicle segment accidents
(c) Intersection accidents

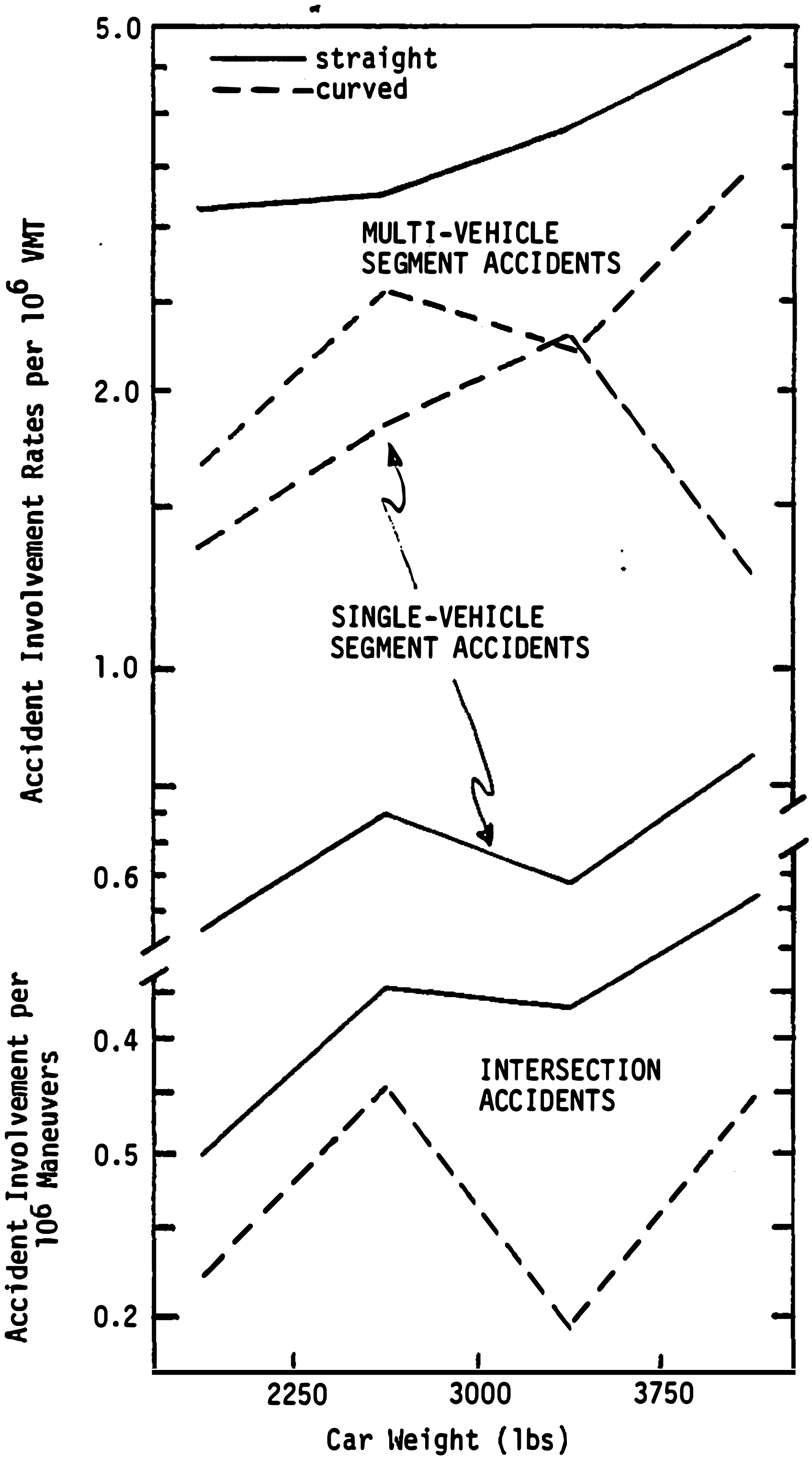


Figure 4.6.4-3. Accident involvement rates by car weight and highway alignment.

4.7 Exploratory Regression Analyses

An alternative to the approach used in Sections 4.5 and 4.6 is to use regression analysis. Each accident, and each exposure observation is treated as one case, accidents representing "failures," exposure observations "successes" (in terms of not being involved in an accident at that time and place). If one assigns a dependent variable $z=1$ to the accident cases, and $z=0$ to the exposure observations, one can fit a model

$$z = a + b_1 x_1 + b_2 x_2 \ldots \qquad (4.7\text{-}1)$$

using pre-crash factors $x_1, x_2, \ldots$ which are known for accidents and exposure observations. $\hat{z}$ values obtained from the model are estimates of the accident probability, given the values of the pre-crash factors $x_1, x_2, \ldots$. Since the exposure observations are only a sample, one has to weight each of them with its expansion factor. In our case, accidents are a census and need not to be weighted. If accidents are sampled, they have also to be weighted with the proper expansion factors.

If one has pre-crash factors with continuous values x, one can use this value directly. If one has discrete factors, one has to introduce "dummy" variables. A dichotomous variable can be represented by a variable with values 0 and 1 for its two levels. A variable with three levels has to be represented by two such variables, etc. This ensures that any pattern of effects among the levels of the variable can be represented by the model. Interactions of factors are represented by products of the corresponding variables.

Regression analysis allows to quickly and efficiently screen a large number of factors and factor combinations. A disadvantage is that the resulting model can give negative probabilities, or probabilities larger than 1 for certain factor combinations. However, even if this happens, the factors selected tend to be valid; only the linear structure of the model is inadequate. Therefore, one can use the regression analysis as an exploratory approach, to identify factors and interactions which merit further study.

Table 4.7-1 lists the factors used for this analysis. Initially, a continuous variable from 0 to 1 approximating light level was also considered. However, most of the observations had values 1 or 0 for the daylight hours and night. Therefore this factor is highly correlated with 19-23 hours. There were relatively few cases during dusk which had values between 1 and 0; a separate analysis was unpromising.

TABLE 4.7-1
FACTORS USED IN EXPLORATORY REGRESSION ANALYSIS

Factors	Levels
Driver age	young, middle, old
Driver sex	male, female
Vehicle occupancy	1, 2, 3 or more
Highway alignment	straight, curved
Highway grade	level, grade
Highway surface	dry, wet
Highway type	state, other
Location	Kingston, other
Time of day	7-11, 11-15, 15-19, 19-23
Day of week	Mo-Th, Fr, Sa, Su

Missing values can be handled in two ways. One is to provide a separate variable for each factor, indicating that it is missing. In our case, this would have meant that these three variables for driver age, sex and vehicle occupancy would have had practically perfect correlation among each other, and a very high one with 19-23 hours. Therefore, we used the other approach, substituting in each unknown case the average values of x_i over all cases. For instance, for cases with unknown driver sex $x_2 = 0.29$ was assumed, averaging x for 71% male and 29% female drivers. This approach is still not perfect: since the percentage of male and female drivers varies over the day, some effects may be distorted. It is possible to refine this approach.

Since only discrete factors were used, a regression analysis using dummy variables is equivalent to fitting an additive model to a high-dimensional contingency table. In our case, it was a 10-dimensional table with 3x2x3x2x2x2x 2x2x2x4x4 = 9216 cells. 6331 exposure observations and 1639 accidents were to be distributed over these cells; intersection and segment cases had to be analyzed separately. This means that this 10-dimensional matrix was only sparsely covered. However, for each of the exposure observation sites the last 7 factors are identical. This means that the 512 cell, 7-dimensional "face" of this table had only 72 (for segments) or 69 (for intersections) cells with exposure observations. This means that any meaningful analysis has to select relatively few factors which in effect collapse the high-dimensional table into a lower dimensional one whose cells are reasonably well covered with observations. In our case, a matrix should not have more than three dimensions, resulting in 8 cells (if each factor is dichotomous) and an average of 9 observation sites per cell.

Table 4.7-2 shows which factors were selected by a regression routine using only factors, no interactions, in the sequence of selection. There is considerable agreement between the factors for segments and intersections. The first four factors agree. That the risk is higher at night, on other than state highways and outside Kingston, is plausible. It is not clear why it should be higher from 11-15 hours. However, one must consider that the sampling design could not be balanced over time within the highway strata (and vice versa). Therefore, part or all of the effect could be due to imbalances of the sampling design.

TABLE 4.7-2
FACTORS SELECTED BY REGRESSION ANALYSIS OF SINGLE FACTOR EFFECTS
(Factors are shown in order of introduction,
"+" indicates increase in accident risk.)

Segments		Intersections	
19-23 hours	+	Not state highway	+
Outside Kingston	+	Outside Kingston	+
11-15 hours	+	19-23 hours	+
Not state highway	+	11-15 hours	+
Wet	+	15-19 hours	+
Grade	+	Saturday	+
Curved	-	Friday	+
Female	+	Wet	+
2 occupants	+	Curved	-

The directions of the effects for highway alignment, and surface agree between segments and intersections. They also agree with those found in the single factor analysis in Section 4.5.2, but that for curvature disagrees with that found in the multifactor analysis in Section 4.5.3. The effect for grade has the opposite sign of that found in Sections 4.5.2 and 4.5.3. However, the simple correlation between grade and accidents has the same sign. This suggests that the discrepancy is due to the effects of the other factors, most likely outside Kingston and not state highway.

The higher risk for female drivers agrees with the finding of Section 4.4.1

Some first order interactions were also explored, namely all 10 interactions between the three highway factors, the highway type, and location in Kingston or outside, and the 16 interactions between the times of day and day of week. Driver and occupant factors which did not or only at a late stage appear in the first analysis were dropped. Table 4.7-3 shows the factors selected by the routine in the order introduced:

TABLE 4.7-3
FACTORS SELECTED BY REGRESSION ANALYSIS OF FIRST ORDER INTERACTIONS
(Driver and vehicle factors are excluded. Factors are shown in order of introduction; "+" indicates increase in risk.)

Segments		Intersections	
19-23 hours	+	Sa, 19-23 hours	±*
Outside Kingston	+	Not state highway	+
11-15 hours	+	Outside Kingston	+
Wet, not state highway	+	Sa, 11-15 hours	+
Fr, 19-23 hours	-	19-23 hours	+
Not state highway	+	Not state highway, outside Kingston	-
Not state highway, outside Kingston	-	Wet*	+
Sa, 11-15 hours	-	Set, outside Kingston	-
Fr, 11-15 hours	-		
Su, 11-15 hours	-		
Grade	+		

*Introduction of the factor "wet" changes the sign of "Sa, 19-23 hours" from + to "-"!

For segments, the first three factors remain the same, the fourth one "not state highway", and the fifth "wet" are here combined into "wet, not state highway." However, because we doubt the representativeness of the observations on wet surfaces (see Section 4.5.2) this factor should not be taken too seriously. "Not state highway" appears in a lower position than in the simple analysis, followed by the interaction "not state highway, outside Kingston" with a negative sign, meaning that the risk on non-state highways outside Kingston is less than the combination of the factors for outside Kingston, and non-state highway. The next three factors indicate that only on weekdays (Mo-Th) the risk is increased between 11 and 15 hours. For intersections, the main difference against the simple analysis is that the interaction Saturday night is first introduced. This, however, is very likely a spurious effect, caused by an observation on a rainy Saturday night, because the sign of this interaction changes when the factor "wet" is introduced.

Overall, we conclude that gross, pre-crash conditions such as highway type, urban/rural environment, and time of day are stronger predictors of accident risk than more specific highway and driver factors. Therefore, we conclude that one must stratify by these gross conditions if one wants to study the effects of more specific pre-crash factors, or that one has to identify those factors which differ among the highway classes, environments and times of the day, and account for the corresponding differences in accident risk.

4.8 Summary of Findings in Ulster County

The data show clearly that female drivers have higher involvement rates in police reported accidents than male drivers (by 0 to 36% higher). The only exception are head-on collisions at night, where the involvement rate for men is higher. There is a suggestion of an interaction between female drivers and wet highway surfaces: the double ratio of the risks varies around 1.6.

The driver age estimates are likely to be biased. However, an implausibly large bias would be required to explain the difference in involvement rates among the age groups. It appears that the involvement rates for young drivers (under 25) do not differ much from those for middle age drivers, and that those for older driver (over 50) are higher.

Of the highway characteristics, grade and surface condition had consistent effects: rates were lower on grades, higher on wet surfaces, whether the factors were considered separately or in interaction with others. Curvature, if considered alone, gave lower involvement rates, but in models with more factors the involvement rates in curves were higher. Curvature and wet surface showed a strong positive interaction, but if grade was also considered, the interaction became weak. Curves and grades showed consistently strong negative interactions.

Traffic speed, and its standard deviation varied little if averaged over the pre-crash conditions studied. Traffic volume, however, varied. When the data were aggregated into two categories, involvement rates appeared to be higher for higher traffic volumes. When more categories were used the relation appeared to be reverse.

Involvement rates increased with increasing car age and also, unexpectedly, with increasing car weight. The increase with age remained for heavy cars and also for old drivers; for other classes the trends were unclear. There is some suggestion that older cars are used in more "rural" environments the younger than younger cars.

On the other hand, lighter cars tend to be used more in "rural" environments than heavier cars. For young drivers, the involvement rate for all types of accidents increases with vehicle weight. For middle age and old drivers the involvement rate decreases with vehicle weight; for other types of accidents it does not appear to vary systematically with vehicle weight. There is no or only a weak interaction between vehicle weight and highway alignment.

A stepwise selection of factors related to involvement rates shows that gross factors such as state highway vs. other highway, Kingston vs. outside Kingston, and time of day had stronger effects than specific highway characteristics, such as alignment, grade and surface. This, together with the observations on the relations between rates and traffic volume, which is related to time-of-day, highway type, and probably also to the location inside or outside Kingston, suggests that one should either stratify the data by these factors, or add other pre-crash factors which sufficiently characterize the differences among these driving environments, when studying the effects of pre-crash factors.

5. VALIDATION WITH SCHENECTADY COUNTY DATA

5.1 The Approach

Validating a methodology has two aspects: (1) to show that it is possible to collect the necessary data, and perform the analyses, and (2) that the results obtained are valid. The result of a validation can be positive, if the required data have been collected, the analyses performed, and results obtained which can be verified. A negative result may not be as simple. If the data could not be collected, it could be simply due to inadequate techniques, or too limited resources. The analysis may fail because of peculiarities of the study population, e.g., if certain factors are highly correlated. Finally, the results may not be confirmed by those from a validation study or other sources because the data base was too small to separate the effects of all factors studied, or because the presence of factors which were not included in the study. Therefore, a positive finding can establish the validity of the methodology, at least in principle, but a negative finding needs to be qualified: one has to distinguish between failure of the methodology, and failure of the specific implementation.

That the necessary data can be collected has been shown in Sections 2 and 3, that the analysis can be performed in Section 4. In this section we examine whether the findings are valid, by using data from Schenectady County.

In the ideal case, one would proceed as follows: One would split the data collection effort so that half of the cases are in one group (area, time period, or a combination of both) half in another. One would perform independent analyses in the two groups. Then one would compare whether the relations between precrash factors and accident risk agree between the two groups, or at least whether a relation found in one group is compatible with the data from another group (if no corresponding relation is found).

We could not proceed in this manner. The extent of the data collection effort was limited. Splitting it evenly over a developmental and a validation area would have given too few observations to allow a balanced sampling design. Therefore, using heuristic arguments and simple calculations, it was decided to split the effort 3:1 among the developmental and validation sample.

The resulting developmental group was just large enough to balance the sampling design. The validation group was concentrated over a shorter time period and on one highway class to allow reasonable balancing of the sampling design.

The main effort was concentrated on analyzing the developmental group. Then, those relations between pre-crash factors and accident rates which were found in the developmental group were examined in the validation group, and agreement or lack of it determined.

5.2 Overall Exposure Estimates and Accident Rates

From the segment observations, one obtains for the study period 24.6×10^6 VMT, and 144×10^6 intersection maneuvers. This gives 5.9 intersection maneuvers per VMT, compared with 4.2 for Ulster County. One reason for the difference is that in Ulster County intersection maneuvers and VMT for all roads were estimated; in Schenectady County, all maneuvers at intersections with state highways were counted including those from approaches not on state highways, but only VMT on state highways.

Two other independent estimates of VMT can be obtained: one from the intersection observations, the other from the Annual Average Daily Traffic given in the state highway inventory. Average hourly passenger car volumes at intersection approaches on state highways combined with the highway miles on each stratum give 32×10^6 VMT for passenger cars in the study period. Combining AADT* for each section with its length gives 41×10^6 VMT for all vehicles. Three quarters of the vehicles observed in Schenectady County were passenger cars. Data from the 1977 NPTS allow to estimate** that 9% of all VMT were travelled between 22:00 and 6:00 hours. As an approximation we will assume the same percentage for the hours 23:00 to 7:00. This gives 28×10^6 VMT for passenger cars during the study period. These independent estimates are 15 to 30% higher than that obtained from the segment observations.

Table 5.2-1 shows accidents, involvements, exposure and accident rates for Schenectady County, and the corresponding rates for Ulster County. Rates for Schenectady County tend to be higher, especially for segment accidents where they are essentially double, and for turning maneuvers at intersections where they are about triple. Such large differences are surprising.

---

*Most of the AADT data were for 1979, some for 1978. The average monthly VMT in the Eastern Region of the US for September/October 1981 was 2% lower than for the annual average for 1979 (estimated from data in "Traffic Volume Trends," published by the Federal Highway Administration).

**H.C. Joksch, Comments on the paper "Rollovers and Serious Driver Injury Differences Among Various Utility Vehicles, Pickup Trucks and Passenger Car Groups," Am. Assoc. for Automotive Medicine, Quarterly Journal, Vol. 5, No. 1, January 1983, 35-43.

TABLE 5.2-1
ACCIDENT INVOLVEMENTS, EXPOSURE AND ACCIDENT INVOLVEMENT RATES, SCHENECTADY COUNTY, AND ACCIDENT INVOLVEMENT RATES, ULSTER COUNTY

	Accident Involvements	Exposure	Accident Involvement Rates (per 10^6)	
			Schenectady	Ulster
Segment accidents		VMT (10^6)		
Single-vehicle	39		1.6	0.9
Multi-vehicle	178		7.2	3.3
Total	217	24.6	8.8	4.2
Intersection accidents		Maneuvers (10^6)		
Going straight	39	128	0.30	0.32
Turning left	24	7.8	3.1	1.25
Turning right	4	6.3	0.6	0.20
Other	0	1.3	0	0.30
Total	67	144	0.47	0.38

The most obvious potential explanation for such a difference is that the Ulster County rates average all highways and rural and urban areas but Schenectady rates are only for state highways outside of the city (though including suburbs). Using accident involvements and estimates of VMT for state highways outside of Kingston in Ulster County gives involvement rates for single- and multi-vehicle and segment accidents of 0.54 and 2.5 per 10^6 VMT, respectively. They are even lower than the countywide rates in Ulster. For intersection accidents on state highways outside of Kingston in Ulster County, the involvement rate is 0.34 per 10^6 manuevers, slightly lower than the countywide figure. Therefore, the hypothesis has to be rejected.

Another possibility is that the VMT obtained from the segment observations are an underestimate, as suggested by the two independent estimates of VMT. This, however, would at most reduce the segment accident involvement rate from 8.8 to 6.8.

There are differences between the traffic conditions in Ulster County, and on the state highways in Schenectady County. Traffic volume in the first is 248 vph, for the latter 480 vph! If the relations between involvement rates and traffic volume suggested by the lines in Fig. 4.5.2-1 were real (which is not very plausible for single-vehicle accidents, and contradicts the observations made in Section 4.5.3.3) one would expect involvement rates of 1.7 for single vehicle, 5.8 for multi-vehicle segment accidents. The first agrees with the actual value 1.6 (but the relation on which the estimate is based is implausible), the second is still lower than the actual value of 7.2.

Figure 5.2-1 shows the distributions of average car speeds in Ulster and Schenectady Counties. Speeds under 30 mph are completely absent in Schenectady County, speeds over 40 mph and even more those above 50 mph are much more frequent in Schenectady County. If there were a nonlinear relation between segment accident risk and speed, one would expect much higher accident risks in Schenectady County.

The unexplained large differences between overall accident involvement rates in Ulster and Schenectady Counties suggest to compare in the following sections not absolute, but relative accident involvement rates among the various pre-crash conditions.

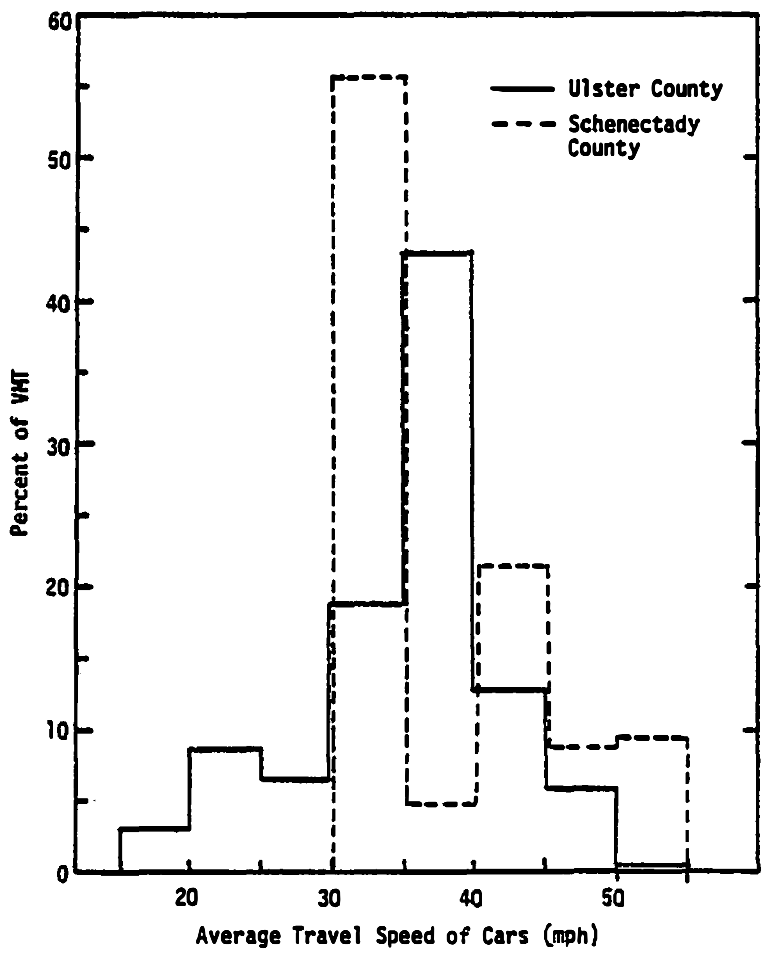


Figure 5.2-1. Distribution of car travel speeds observed in Ulster and Schenectady Counties.

5.3 Driver Characteristics

Table 5.3-1 shows accident involvement and exposure by driver sex for daytime hours (7:00 - 19:00). For 0.5% segment exposure, driver sec could not be observed, and for 7% of intersection exposure. For segments, accident involvement and exposure agree well between Ulster and Schenectady County. For intersections, exposure agrees, but the accident involvement of male drivers is lower in Schenectady than in Ulster County (however, with a standard error of 7 for 47%, the difference is not significant).

This confirms the finding of a higher accident involvement rate for women than for men: 1.3 times as high for segments (the same as in Ulster County), and 1.8 for intersections (1.2 in Ulster County).

TABLE 5.3-1
ACCIDENT INVOLVEMENT AND EXPOSURE BY DRIVER SEX
(Time of day, 7:00 to 19:00 hours)

County	Segments		Intersections	
	Accident Involvements	Exposure	Accident Involvements	Exposure
	Percent of Male Drivers			
Schenectady	56	62	47	62
Ulster	57	63	57	61

Table 5.3-2 shows the distribution of exposure and accident involvements by driver age, for day and night combined. The involvements in segment accidents are distributed as in Ulster County; the distribution of exposure is

TABLE 5.3-2
DISTRIBUTION OF ACCIDENT INVOLVEMENTS AND EXPOSURE BY DRIVER AGE
(Only drivers with estimated age were considered. Age was missing for 24% of segment exposure, 8% of intersection exposure.)

	Driver Age		
	Young	Middle Age	Old
	Percent of all drivers with estimated age		
SEGMENTS			
Accidents	36	39	25
Exposure	36	50	14
Ratio accidents/ exposure	1.0	0.8	1.8
INTERSECTION			
Accidents	31	40	29
Exposure	34	54	12
Ratio accidents/ exposure	0.9	0.7	2.4

similar, but there are fewer young, and more old drivers. Also, differences between accident involvement rates for young and middle age drivers are relatively small, those for old drivers are much higher.

However, because the reliability of driver age estimates in exposure observations is doubtful, no substantive conclusions should be drawn. One can only conclude that the patterns observed in the two counties agree.

5.4 Highway Factors

Tables 5.4-1 through 3 show accident involvements and exposure by highway alignment, grade and surface condition.

The actually observed distribution of surface condition was different: 2.9% of the segment exposure, and 1.6% of the intersection exposure were on wet surfaces. As in Section 4.5.3.3 we used as "modelled" frequency the weighted average of 2%.

The relation between alignment and segment accident involvement rates is the opposite of that in Ulster County: it is much higher in curves than on straight sections. For intersection accidents, the involvement rate is much lower in curves than on straight roads.

The relation between grade and segment accident involvement rates is also the opposite of that in Ulster County: it is higher on grades. For intersection accidents, the relation is the same as in Ulster County.

At first glance, the relation between accident involvement rates and highway surface in Schenectady is similar to that in Ulster County. This agreement, however, should be interpreted with great caution. The various models (Table 4.5.3.3-7) showed that in Ulster County involvements on wet roads were 10 to 90% higher than on dry roads. In Schenectady County, they were about 25 times as frequent as on dry roads! Also, this ratio is essentially the same for segment and intersection accidents. This is very implausible. Therefore, it is likely that the frequency of wet roads in Schenectady County was underestimated due to the sparse observation schedule.

In Ulster County, we found an interaction between highway alignment and grade related to accident involvement rates (Section 4.5.3.2). Table 5.4-4 shows accident involvements, exposure and rate by alignment and grade for Schenectady County. To make the rates more easily comparable they are shown in Table 5.4-5 relative to that for straight level roads, together with those for Ulster County. There is no obvious similarity in the patterns. There is also

TABLE 5.4-1
ACCIDENT INVOLVEMENT, EXPOSURE AND ACCIDENT INVOLVEMENT RATES BY HIGHWAY ALIGNMENT, SCHENECTADY COUNTY

	Straight		Curved	
	Segment	Intersection	Segment	Intersection
Accident involvements	179	55	36	12
Exposure (10^6)	21.8	87	2.8	57
Involvement rate (per 10^6)	8	0.6	13	0.2

TABLE 5.4-2
ACCIDENT INVOLVEMENT, EXPOSURE AND ACCIDENT INVOLVEMENT RATES BY HIGHWAY GRADE, SCHENECTADY COUNTY

	Level		Grade	
	Segment	Intersection	Segment	Intersection
Accident involvements	154	59	61	8
Exposure (10^6)	18.3	98	6.3	46
Involvement rate (per 10^6)	8	0.6	10	0.2

TABLE 5.4-3
ACCIDENT INVOLVEMENT, EXPOSURE AND ACCIDENT INVOLVEMENT RATES BY HIGHWAY SURFACE, SCHENECTADY COUNTY
(The exposure on wet roads is a weighted average of those actually observed for segments and intersection)

	Dry		Wet	
	Segment	Intersection	Segment	Intersection
Accident involvements	137	44	77	23
Exposure (10^6)	24.1	141	0.5	3
Involvement rate (per 10^6)	.6	0.3	150	8

TABLE 5.4-4
ACCIDENT INVOLVEMENTS, EXPOSURE AND ACCIDENT INVOLVEMENT RATES BY HIGHWAY ALIGNMENT AND GRADE, SCHENECTADY COUNTY

	Straight		Curved	
	Segment	Intersection	Segment	Intersection
LEVEL				
Accident involvements	137	49	17	10
Exposure (10^6)	16.1	76	2.2	21
Involvement rate (per 10^6)	8.5	0.64	7.7	0.47
GRADE				
Accident involvements	42	6	19	2
Exposure (10^6)	5.7	10	0.6	36
Involvement rate (per 10^6)	7.4	0.59	32	0.06

TABLE 5.4-5
ACCIDENT INVOLVEMENT RATES BY HIGHWAY ALIGNMENT AND GRADE, RELATIVE TO THAT FOR STRAIGHT LEVEL ROADS, SCHENECTADY COUNTY AND ULSTER COUNTY

	Straight		Curved	
	Segments	Intersection	Segments	Intersection
LEVEL				
Schenectady	1	1	0.9	0.7
Ulster	1	1	2.8	1.5
GRADE				
Schenectady	0.9	0.9	4.1	0.1
Ulster	0.6	0.6	0.5	0.4

no similarity in the interaction which might be less easily recognizable: on segment accidents in Schenectady, the double ratio ("odds-ratio") of rates for segment involvements is 5, for Ulster 0.3. For intersection involvements, the respective values are 0.2 and 0.4: the effect (if any) has at least the same direction in both counties: accident involvement in curves on grades are less frequent than one would expect from the combination of the effects of grades.

The Ulster County data showed also an interaction effect between highway surface and driver sex (Section 4.5.3.1). Table 5.4-6 shows corresponding data for Schenectady County. The double ratio of rates for segment accidents is 1.36, for intersection accidents 0.78. The first is roughly comparable to that found in Ulster County, the second, however, indicates an effect in the opposite direction.

TABLE 5.4-6
ACCIDENT INVOLVEMENTS, EXPOSURE AND AACIDENT INVOLVEMENT RATES BY DRIVER SEX AND HIGHWAY SURFACE, SCHENECTADY COUNTY
(Daytime, 7:00- 19:00 hours)

	Male		Female	
	Segments	Intersection	Segments	Intersection
DRY				
Accident involvements	61	18	49	19
Exposure (10^6)	11.3	81	6.9	49
Involvement rate (per 10^6)	5.4	0.22	7.1	0.39
WET				
Accident involvements	32	9	25	11
Exposure (10^6)	0.045	1.1	0.020	1.04
Involvement rate (per 10^6)	711	8	1250	11

5.5 Vehicle Characteristics

In Ulster County, we made some unexpected observations on the relation between accident involvement rates and car weight: heavier cars tended to have higher accident involvement rates than lighter cars. Table 5.5-1 shows accident involvements and exposure in Schenectady County by car weight for the 85% of accident involvements and 85% of exposure for which car weight was available. Again, heavier cars tend to have higher accident involvement rates than lighter cars.

In Fig. 5.5-1 the accident rates for Schenectady County and Ulster County are presented together. Though the values for Schenectady fluctuate even more than for Ulster, the overall trend is the same: heavier cars tend to have higher rates than lighter cars.

In Section 4.6.4 we found that this pattern changed if interactions between driver age and car weight were considered. With the low numbers of accidents and the high fraction of exposure where car weight or driver age was missing (34% for segments, 20% for intersections) a breakdown of the data as fine as in Section 4.6.4 was not possible. Only "light" (less than 3000 lbs) and "heavy" cars, and "young" and "other" drivers were distinguished. Table 5.5-2 shows the accident involvements, exposures, and rates.

The pattern remains the same within each age group: heavier cars have greater involvement rates than lighter cars. This contradicts, to some extent, the findings of Section 4.6.4 (Fig. 4.6.4-2) that this held for young drivers, but that for middle age or old drivers the relation was weaker or reversed. Table 5.5-3 compares data for Schenectady and Ulster County. To eliminate the effects of the differences in the absolute levels of the rates, the rates for heavy and light cars are shown relative to the average for each driver age class in Table 5.5-3. For young drivers, the relative involvement rates for heavy cars is even greater in Schenectady County than in Ulster. For other drivers, the relative involvement rates for heavy cars are also larger than for light cars in Schenectady, though generally less so than for young drivers. In Ulster County, the differences between involvement rates for light and heavy cars are smaller, and vary in direction.

In sum, one can conclude that the Schenectady County data confirm that the involvement rates for young drivers increase with car weight. For older drivers, the data do not completely agree, though they show some similarities.

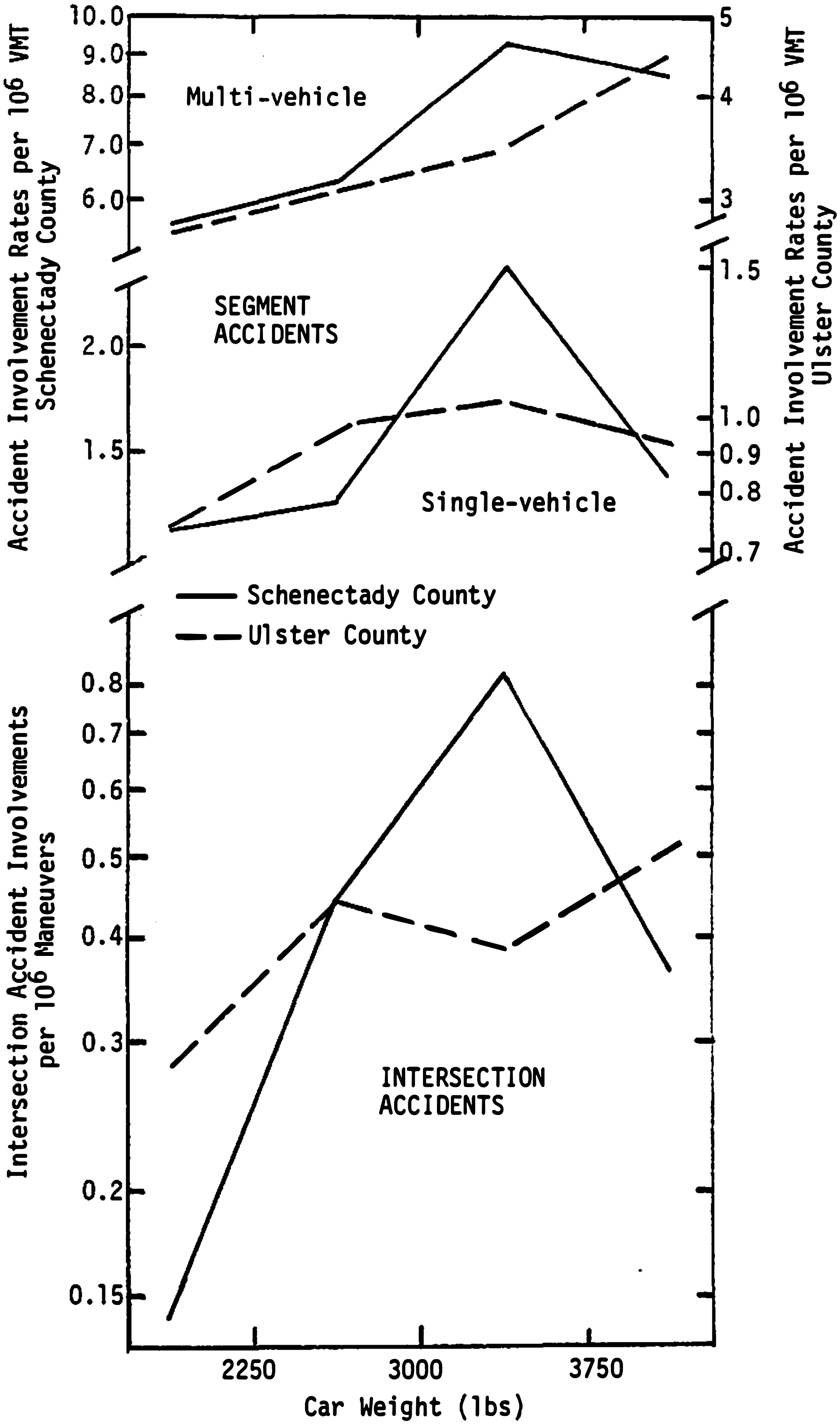


Figure 5.5-1. Accident involvement rates by car weight, Ulster and Schenectady Counties.

TABLE 5.5-1
EXPOSURE AND ACCIDENT INVOLVEMENT BY CAR WEIGHT, SCHENECTADY COUNTY

	Car Weight (lbs)			
	<2250	2250-2999	3000-3749	≥3750
SCHENECTADY COUNTY				
VMT (10^6)	5.0	5.5	5.6	5.0
Segment Accidents				
Single-vehicle	6	7	14	7
Multi-vehicle	28	29	52	42
Intersection Maneuvers (10^6)	35	23	37	28
Intersection Accidents	5	10	30	10
Segment Accident Involvement Rate				
Single-vehicle	1.2	1.3	2.5	1.4
Multi-vehicle	5.6	5.3	9.3	8.4
Intersection Accident Involvement Rate	0.14	0.44	0.82	0.36

TABLE 5.5-2
ACCIDENT INVOLVEMENT, EXPOSURE AND INVOLVEMENT RATES IN SCHENECTADY COUNTY BY DRIVER AGE AND CAR WEIGHT

	Driver Age					
	Young			Other		
	Segments		Inter-sections	Segments		Inter-sections
	Single-Vehicle	Multi-Vehicle		Single-Vehicle	Multi-Vehicle	
LIGHT CARS						
Accident involvements	6	23	6	7	34	9
Exposure (10^6)	3		23.3	5.3		31
Rates (per 10^6)	2.0	7.6	0.26	1.3	6.4	0.29
HEAVY CARS						
Accident involvements	13	26	9	8	68	31
Exposure (10^6)	2.3		15	5.4		45
Rates (per 10^6)	5.6	11	0.60	1.5	13	0.68

TABLE 5.5-3
ACCIDENT INVOLVEMENT RATES BY DRIVER AGE AND CAR WEIGHT
(The rates are relative to the average for each driver age group, over all cars with known weight. Figures cannot be compared between the age groups.)

	Young			Middle Age and Old		
	Segment		Inter-section	Segment		Inter-section
	Single-	Multi-Vehicle		Single-	Multi-Vehicle	
LIGHT CARS <3000 lbs						
Schenectady	0.6	0.8	0.7	0.9	0.7	0.6
Ulster	0.8	0.8	0.8	1.1	0.9	0.9
HEAVY CARS ≥3000 lbs						
Schenectady	1.6	1.2	1.5	1.1	1.3	1.3
Ulster	1.3	1.3	1.2	0.9	1.1	1.1

5.6 Summary of Validation

Accident involvement rates in Schenectady County are higher than in Ulster County. They are about twice as high for segment accidents, and a quarter higher for intersection accidents. However, traffic volume is twice as high in Schenectady as in Ulster County, and average travel speeds of cars tend to be higher. Without accounting for potential effects of these factors, no conclusion can be drawn from the discrepancy in the rates.

Data from the two counties agree that the involvement rates for female drivers are higher than for male drivers. They also agree that older drivers have much higher accident rates than middle age and young drivers (but this observation should not be taken at face value because of the possible age bias which both data sets would have in common).

The relations between involvement rates and highway factors contradict each other to a large extent. The only clear similarity is that in both counties the involvement rates on wet roads are higher than on dry roads.

In both counties the involvement rates increase with car weight. The same holds for young drivers. For middle age and old drivers this relation holds also for multi-vehicle accidents in both counties; for single-vehicle accidents the relations disagree, but both are weak.

It is noteworthy that the findings in both counties agree for those factors which can vary within an observation site, but that they show no agreement or great fluctuations for factors which vary only between observation sites. This suggests strongly that the samples were adequate in terms of the numbers of cars observed, but not in terms of the number of setups.

6. SUMMARY AND CONCLUSIONS

6.1 Overview

This work had the objectives to further develop a methodology for studying accident causation in terms of accident probabilities depending on pre-crash factors, to demonstrate the feasibility of the methodology by implementing it, and to obtain actual results on accident causation. This final report describes the last two parts of the study. The methodology is presented in a separate manual.

6.2 The Methodology

Key areas of the methodology are:

- exposure data collection techniques
- exposure data collection planning
- accident data
- analysis

Exposure data collection consisted of observing cars and recording time, place, license plate, speed, driver characteristics, highway characteristics and ambience, and a few other factors at selected locations. Data were collected visually, photographically, and by radar. During daylight, most of these data could be collected. There were sometimes problems with photographing license plates in dense traffic. Sometimes driver age could not be estimated. Sometimes rain impeded the observations. At dusk or night, license plates can often be visually read when they can no longer be photographed. Driver characteristics could be observed at night only under unusually favorable conditions. To get a better opportunity to read license plates and observe drivers, "chasing" a car by following it with the observers' car was tried. It was unsuccessful because of too high travel speeds or bad roads.

Setting up photographic and radar equipment used considerable time. Using only visual observations, productive time could have been increased by 50% or more. With visual observations alone (without making marks on the road, etc.) one cannot obtain speed. One can read license plates, but has no objective record to check questionable numbers, e.g., distinguish 0, O or Q, etc. Therefore, for a given level of effort, one has to make a trade-off between observation time and setup-time: the one allows to obtain more observations, especially use of more observation sites; the other to obtain more, and more reliable information.

A special problem is that traffic volumes on local roads are low, especially at night, though the total number of accidents under such conditions is not. Exposure observations under such conditions are very unproductive in terms of observations per observer hour. For such conditions more sophisticated techniques need to be developed and tested.

To develop a sampling design, one needs a sampling frame. If no adequate highway inventory is available, one has to use maps, and inspect the prospective observation sites. This should not only verify their existence, but also assign them to the proper stratum. If possible, a rough estimate of traffic volume should also be obtained.

We found that a sampling plan which is adequate for estimating aggregate exposure is not sufficient for estimating exposure in certain pre-crash conditions. If one is studying factors which can vary among vehicles and drivers at each observation site, the number of sites is relatively less important than the total number of observations. On the other hand, if one is studying factors which differ only between sites (or observation periods) the number of observation sites (or periods) becomes relatively more important than the number of vehicles observed. This is especially important for transient conditions, such as wet surfaces or rain. Also, we found that very broad pre-crash conditions, such as urban/rural, and time of day had a very strong relation to accident rates. To study more specific pre-crash factors, stratifying by such general factors might be necessary. This requires balancing the design within each stratum, and therefore a large number of observation sites.

Because of the travel time between observation sites, there is a trade-off between the total number of vehicles which can be observed, and the number of sites at which observations can be made, with a given level of effort. Which combination to select depends on the factors of greatest interest.

When a sampling design is translated into a specific sampling schedule, one has to allow for the uncertainty of travel times for which only rough estimates may be known; with a rigid schedule, some time may be wasted. With a flexible schedule, more observations may be obtainable, especially if observers can quit high-volume sites after observing a certain number of vehicles, and proceed to low-volume sites. Such a procedure, however, may create quality control problems.

The number of accidents investigated by the NASS team was too low to allow any meaningful analysis. Therefore, all police-investigated accidents in the selected area were studied.

NASS obtains more information on the accidents than the police, but most of it is for the crash- and post-crash phase. For the pre-crash data, the difference is not as great. Since it is likely that for the foreseeable future the number of NASS investigated accidents will be relatively small, one will have to rely to a large extent on police-investigated accidents. It appears worthwhile to study how the police-investigated accidents may be used as a basis to extrapolate certain pre-crash factors from the NASS cases.

Our analysis suggested that traffic volume, and possibly also travel speed influence accident risk. Since this information is not available for accidents, these factors could not be rigorously studied. It appears worthwhile to make at least rough estimates of traffic volume and speed for the time and location of each accident.

Two analytical approaches were tried: (1) to aggregate data in multi-dimensional contingency tables, and to represent the accident involvement rates by simple models using pre-crash factors, and (2) treat each exposure observation ("success") and each accident ("failure") as one observation, and fit a regression model in the pre-crash factor to these data. Both approaches worked and gave some plausible results.

6.3 Effects of Pre-crash Factors

Because of the limited number of exposure observations, no reliable quantitative estimates could be made, but some qualitative conclusions could be drawn.

Female drivers have higher accident involvement rates than male drivers. The only exception are head-on collisions at night. To what extent this might be due to other factors which are related to driver sex could not be determined.

Older drivers (over 50 years) have much higher involvement rates than middle age and younger drivers (and the difference between these two groups is relatively small). Though part of the difference may be due to a bias in estimating age, it is very unlikely that the entire difference is due to it. To what extent the remainder is due to other factors which are related to driver age could not be determined.

Accident involvement rates increase with car age. To some extent this might be due to more "rural" driving for older cars. The increase of involvement rates with car weight is most pronounced for young drivers. For middle age and old drivers, the variation with car weight is less, sometimes an increase, sometimes a decrease.

With the exception of wet surfaces, no highway factor had a consistent relation with involvement rates in Ulster and Schenectady County. To some extent this might be due to the very limited information available in the police accident reports. Even the great increase in rates associated with wet surfaces should be interpreted with caution, because of the small number of observations with wet surfaces.

Driving environment in broad terms, such as state highway or other highway, inside or outside Kingston, and time of day, has a much stronger relation with involvement rates than the following factors: driver age, sex, number of occupants, highway grade, highway alignment, and surface. One might speculate that traffic characteristics play an important role.

This speculation is supported by the observation that in Ulster County traffic volume showed—though not consistently--relations with accident involvement rates, and that Ulster and Schenectady Counties differed in involvement rates and in traffic volumes and travel speeds.

APPENDIX A

COLLECTION AND PROCESSING OF EXPOSURE DATA

A.1 COLLECTION OF EXPOSURE DATA

A.1.1 Introduction

This Appendix describes the procedures used for observing traffic and recording the data as a basis for estimating exposure.

At specified times and locations, for passenger cars (sampled, if necessary), data on:

- vehicle maneuvers
- vehicle identity
- vehicle speed
- vehicle driver and occupant attributes,

and information on the highway, traffic and ambient conditions are collected.

The method of observing and recording vehicle identity and vehicle speed data depends on site characteristics as outlined in Table A.1-1. Driver and occupant characteristics are always obtained by direct visual observation and recorded on an audio cassette.

TABLE A.1-1
SUMMARY OF VEHICLE DATA COLLECTION TECHNIQUES

Type of Approach	Day		Night	
	Vehicle	Speed	Vehicle	Speed
Midblock (segment), general	Photography	Radar	Binoculars or visual	Radar
Intersection Approach: Uncontrolled	Photography or visual	Radar	Binoculars or visual	Radar
Intersection Approach: Signal control	Photography or visual	None	Binoculars or visual	None
Intersection: Stop control	Visual	None	Visual	None

Notes:
1. "Visual" means that information observed visually is recorded onto audio cassette.
2. "Photography or visual" means that photographs are taken unless traffic volumes and vehicle speeds are low (generally less than 15 mph) so as to permit reliable observations of a larger sample than would be possible using photography.
3. Night data collection procedures are used when license plates can not be legibly photographed.

The following equipment was used:

- 35mm camera with motor-drive and data back
- 80-200mm zoom lens f4.5
- Radar speed meter
- Stereo tape recorder
- Mono tape recorder

A.1.2 Standard Data Collection Procedure

The "standard" data collection procedure is used for daylight conditions at all midblock (segment) locations and at uncontrolled intersection approaches.

A.1.2.1 Equipment Setup

The team leader identifies the exact location at which data is to be collected. The setup activities include:

- Unpack and prepare equipment.
- Establish necessary screen lines.
- Fill out Data Collection Log - Part A (Fig. A.1-2) and record header information.

The general data collection setup is shown in Fig. A.1-1. Table A.1-2 outlines a six-step procedure for the setup activities. Table A.1-3 lists the required header information.

A.1.2.2 The Data Collection Log

The Data Collection Log is the primary documentation of the data collection activities. It consists of two separate portions:

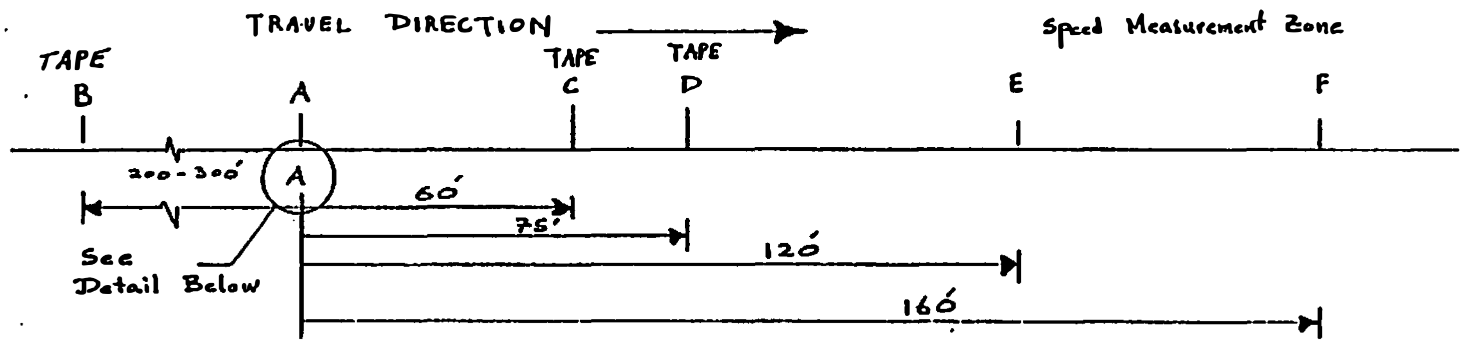
Part A - Identifies and describes the data collection site and the type of data collection. It is completed before data collection starts.

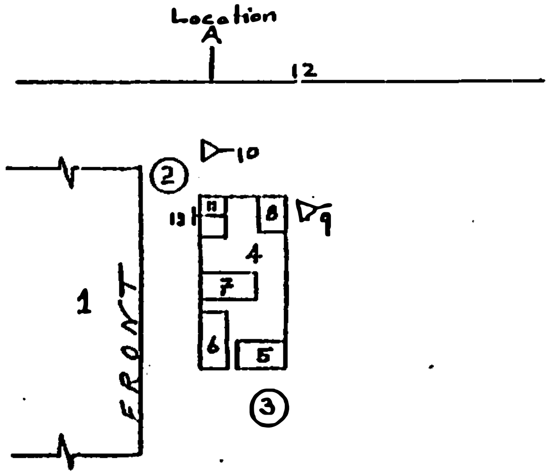
Part B - Identifies and describes the data collected. It will be completed during data reduction taking information from the audio recording.

In addition to completing this log, and the header information for each data cassette used, all unusual circumstances or other items of interest should be recorded on the audio tape and transcribed to the appropriate form.

Part A. Site Description

The site must be described so that the exact location can be identified. In case of midblock locations, the identification should be to the nearest one-tenth of a mile or better for rural areas or city block for urban areas.





LOCATION OF EQUIPMENT

1. Team Vehicle
2. Team Leader
3. Team Member
4. Portable Table
5. Clipboard with Counters
6. Stereo Recorder
7. Mono Recorder
8. Speed Digital Display
9. Radar Antenna
10. Camera
11. Two Boxes
    - new films
    - used films
12. Pavement Edge
13. Check List

SCREEN LINES

A - Location of Data Collection Crew and Equipment

B - Upstream Screen Line

C - Downstream Screen Line for Actuating Camera Shutter

D - Location of Photograph

E - Near Point of Speed Measurement Zone

F - Far Point of Speed Measurement Zone

Figure A.1-1. Data collection setup.

Site No. _____ County __________ City __________ Town/Village __________

Road Name or Number ____________________

at intersection with ____________________ or

_______ miles ____ upstream ____ downstream from ____________

Landmark: ______________________________

(Reference Markers, milepost, traffic sign number, or any other permanent structure)

Data Item	Approach A	Approach B	Approach C	Approach D	Approach E
Roadway Name					
Direction of Travel					
Number of Lanes					
Approach Width					
Land Use (R,U)					
Horizontal Alignment (S,CR,CL,RC)					
Vertical Alignment (F,U,D,C,S)					
Speed limit					
Pavement surface (A,C,G)					
Passing (Y, N)					
Parking (Y, N)					
Control					
Turning Lanes					
Type of Data Collection					
Setup Number					

Data collection crew (initials) 1.___ 2.___ 3.___ Date: __/__/__

Figure A.1-2. Data Collection Log - Part A.

TABLE A.1-2
SCHEDULE AND ASSIGNMENT OF SETUP ACTIVITIES

Time	Activity	Team Leader	Team Member
Start	1	(Both) Take equipment out of the car. Set up the table and tripods. Place the following items on table: • Clipboard with counters • Stereo recorder • Beeper recorder • Film boxes • Digital speed display	
Must be less than 20 minutes	2	Take measuring wheel Install downstream tapes (C,D)	Assemble camera Mount camera on tripod
	3	Focus camera at wheel rod	
	4	Rewind Beeper cassette Label data cassette, and load into stereo recorder	Set zoom and f setting on camera; set number on data back
	5	Fill out site log Reset stereo recorder counter Record header information Set PAUSE	Set up radar equipment Connect to battery Check calibration
	6	Release PAUSE Record time	Take exposure (photo) of the camera's field of view
		Collect data	Collect data

TABLE A.1-3
HEADER INFORMATION

The following will be recorded at the beginning of every audio cassette used during data collection:

1. Site number
2. Location - Route Name or Intersection and approach
3. Setup number
4. Date: Month & Year
5. Time: Hours and Minute, AM or PM (Note: if the tape recorder is ever stopped during data collection, the time must be the first item recorded after restart.
6. Weather: Clear, cloudy, rain
7. Roadway surface: Dry, wet
8. Type of data collection

If there is a significant change in weather or in roadway surface conditions during data collection, this information must be recorded as it happens together with the time of occurrence.

The site description details are given separately for each approach to an intersection.

Direction of Travel - The approximate compass direction in which vehicles are traveling.

Number of Lanes - The effective number of lanes used by vehicles.

Approach Width - The width in feet of roadway used by vehicles.

Land Use - Land use will be defined as

Rural - if there are shoulders
Urban - if there is curb or curb and gutters.

Horizontal Alignment - Use the following codes

ST - Straight
CR - Curve to the right (in the direction of travel)
CL - Curve to the left (in the direction of travel)
RC - Reverse Curve

Vertical Alignment (Grade) - Use the following codes:

F - Flat
U - Upgrade in direction of travel
D - Downgrade in direction of travel
C - Crest - data collection is higher than both upstream and downstream locations
S - Sag - data collection is lower than both upstream and downstream locations

Speed Limit - As posted. If unposted, mark as UP.

Pavement Surface Type - Use the following codes:

A - Asphalt
C - Concrete
G - Gravel or Macadam

Passing - Note by Y (Yes) or N(No) whether passing is permitted. Leave blank for roads with two or more lanes in the direction of travel.

Parking - Note by Y (Yes) or N (No) whether parked vehicles (legal or not) are present during the data collection process within 200 ft in either direction of the data collection point.

Control - Use one of the following designations:

1 - Traffic signal
2 - Red Flasher
3 - Stop sign
4 - Yellow Flasher
5 - Yield Sign
6 - Uncontrolled
7 - Not intersection approach

Turning Lanes - Use the following codes:

1 - Right turn lane
2 - Left turn lane
3 - Both left and right turn lane
4 - Not intersection approach

Type of Data Collection - Use one of the following designations:

Standard
Visual
Night
Stopped Approach
Surrogate Location

Part B. Data Collection

Part B documents the data collection process and the data collected. It is completed during the data reduction activities from information recorded on the audio cassettes. A separate Part B will be completed for every intersection approach on which data was collected.

Date and day of week

Time data collection started - Use twenty-four-hour clock and record time to the nearest minute. Data collection starts when all the equipment is ready and the traffic count starts, not when the first vehicle is sampled or counted.

Time data collection ended - Time last traffic count total was recorded.

Weather & Surface Conditions - Taken from header information.

Interruptions - Complete the film log for all approaches at which photography was used.

A.1.2.3 Data Collection

A.1.2.3.1 Start and Stop Data Collection

The team leader records when data collection starts and stops. He may also interrupt data collection in case of any occurrence that obviously affects the speed of passing vehicles, such as:

- School bus loading or unloading passengers.
- Any vehicle stopping in the travelled lanes within the data collection zone except in obedience to a traffic control device at intersection.
- Passage of an emergency vehicle.

Data collection will also be interrupted for film changes. All starting and stopping times must be recorded.

The tape recorder will be stopped whenever data collection is interrupted and the traffic count suspended.

A.1.2.3.2 Team Leader Activities

When the team is ready to collect data, the Team Leader will state into the recorder (Channel A):

"Data Collection at site ____________ will start at time __________ on (Month) (Day) (Year)".

Subsequently, the Team Leader counts all passing vehicles at Location A using the mechanical counters--one for Lane 1 vehicles, the other for Lane 2 vehicles. Each vehicle is identified as an <u>auto</u>, <u>truck</u> or <u>bus</u>.

Every 10 minutes, or so, when a lull appears in the traffic stream, record the current values of the counts onto the audio cassette as follows:

"Time is ____________"
"Lane 1 autos, _____, trucks _____, buses _____"
"Lane 2 autos, _____, trucks _____, buses _____"

This must also be done at the end of the data collection period. Whenever the tape is stopped for an interruption of data collection, the last item recorded prior to stopping must be the time and cumulative count totals.

A.1.2.3.3 Team Member

The team member collects data only for "sampled" auto vehicles--no other type of vehicle is sampled. He records most of this data on audio cassette using Channel B. He is also responsible for tripping the camera shutter at the correct time.

This procedure is a sequence of events. The sequence begins when a selected oncoming vehicle crosses the tape at Location B. The Team Member "selects" a vehicle if:

- It is an auto.
- The "Beeper" cassette has issued at least one MARK command <u>following</u> the completion of his data collection activity of the prior vehicle.
- There is a trailing gap of a sufficient length to permit unobstructed observation of the license plate.

This selected vehicle is called a <u>sample</u>.

Vehicles which are part of a funeral procession should not be sampled. A special notation "Funeral Procession--Vehicles" should be recorded on tape.

As the sample crosses Location B, the Team Member will record the following onto the cassette using the appropriate entries from Table A.1-4.

1. Vehicle color.
2. Vehicle class.
3. Number of occupants.
4. Driver sex and age classification.

Then focus on the tape marker at Location C. At the instant the rear wheel of the sample crosses the marker position:

5. Trip the camera shutter using the remote cord.
6. Record the sample speed from the radar display.
7. Record the lane occupied by the sample within the speed trap.

TABLE A.1-4
TEAM MEMBER DATA COLLECTION ITEMS

1 - Vehicle Class	3 - Occupancy
1. Subcompact 2. Compact 3. Intermediate 4. Full size 5. Undetermined	1. One 2. Two 3. Three or more 4. Undetermined
2 - Vehicle Color	**4 - Driver Sex**
1. Black 2. Blue 3. Brown 4. Gold 5. Grey 6. Green 7. Maroon 8. Orange 9. Pink 10. Purple 11. Red 12. Tan 13. White 14. Yellow 15. Undetermined	1. Male 2. Female 3. Undetermined **5 - Apparent Drive Age** 1. Child (Under 16) 2. Young (16 - 25) 3. Middle Age (26 - 50) 4. Old (Over 50) 5. Undetermined

8. Record whether the speed of the sample is constrained by a preceding vehicle.
9. Record the turning movement.
10. Record the frame count on the camera.

Step 10 completes the data collection procedure for the sample car. The Team Member then waits for the next MARK command and observes the first auto crossing Location B after that MARK command, proceeding with Step 1.

TABLE A.1-5
SCHEDULE AND ASSIGNMENT OF PACK-UP ACTIVITIES

Time	Activity	Team Leader	Team Member
Start	1	Remove cassette from stereo recorder. Rewind Beeper cassette.	Remove shutter release cord from camera. Dissemble radar equipment, pack in case.
	2	Put all data (cassette, exposed film, data collection log) in envelope marked with site ID, if last setup at site.	Unload camera if film counter reads 21 or higher.
	3	Remove tapes from pavement at B, C, D.	Place equipment on table into car, then table and measuring wheel and radar equipment and battery.
	4	Mark location A with yellow spray on edge of pavement or curb.	Remove camera from tripod, pack tripod in car.
	5	General site clean up.	Disassemble camera and pack into case, if last setup of the time slot; store in car.

A.1.3 Procedure for Data Collection at Night* and During Inclement Weather

In these situations, the crew is seated in the car which is parked on the shoulder or by the curb along the approach which is sampled. The Team Leader is in the driver's seat while the Team Member is seated behind him. The Team Leader has the binoculars, a microphone connected to the stereo recorder and the radar display unit where required; the Team Member has the counters and a microphone connected to the stereo recorder.

*When conditions permit (weather and personal safety), use the procedure of Section A.1.4 at night on intersection approaches controlled with a signal or sign.

The procedure is:

1. The Team Member records license plate of the sampled vehicle using the binoculars (if necessary).

2. The Team Leader selects the sampled vehicles. He records the other required information in the order given in Table A.1-6 and keeps a count of all vehicles by type and lane using the mechanical counters.

This method will be used under all inclement weather conditions even if visibility is reduced to such an extent that license plate data cannot be recognized. Only traffic count data and other visually obtained data; e.g., vehicle class, will be recorded.

TABLE A.1-6
DATA COLLECTION SEQUENCE

Team Leader (Channel A) Night Data Collection and Inclement Weather

1. Color of vehicle
2. Class: Subcompact, Compact, Intermediate, Full
3. Number of occupants
4. Driver sex and age: Child, Young, Middle, Elderly
5. Record speed: XX mph
6. Record lane (two-lane approach, only): Lane 1, Lane 2
7. If time permits: record make and model of vehicles

Team Member (Channel B of stereo recorder)

1. Read license plate and record it on cassette recorder
2. Registration type
3. Registration state

A.1.4 Data Collection - Controlled Approaches

These procedures apply to the following cases:

- Standard Data Collection at intersection approaches controlled by STOP signs, red flashers or other traffic control devices requiring all vehicles to come to a complete stop.
- Selected, controlled, intersection approaches at night.

This type of data collection will be done by a single team member using a cassette recorder. The team member will position himself near and upstream of the stopline. For each vehicle that approaches, he will record a type designation (car, truck, bus). For each sampled auto he will record:

1. Vehicle class
2. Vehicle color
3. Occupancy
4. Driver sex
5. Apparent driver age
6. State of registration
7. License plate number including special plate
8. Vehicle action at stop line (STOP control, only)
   a. complete stop
   b. rolling stop (slow--brake light on--to 10 mph or less)
   c. did not stop
9. Turning movement
10. If time permits, make and model of vehicle

An auto which is not sampled--but is counted--should be recorded as "missed."

For items 1 through 5, use the items defined in Table A.1-4.

At the beginning of each cassette, the full "header information" will be recorded.

A.1.5 Summary of Data Collection Procedures

Tables A.1-7 and A.1-8 summarize, respectively, the data collection responsibility and equipment assignments.

TABLE A.1-7
DATA COLLECTION ASSIGNMENTS

	SITE	PROCEDURE	TEAM LEADER	TEAM MEMBER
DAY	Midblock	Standard	*All vehicles:* Count, type, lane.	*Sampled vehicles:* Color, classification, occupants, driver age and sex, photograph, speed, frame count.
DAY	Inter-section	Signal control	*All vehicles:* Count, type.	*Sampled vehicles:* Color, classification, occupants, driver age and sex, photograph or record turn movement, frame count.
DAY	Inter-section	STOP, Red flasher	On separate approaches, *All vehicles:* Count, type. *Sampled vehicles:* Color, classification, occupants, driver age and sex, license, turn movement, count.	
NIGHT	Midblock	Standard	*All vehicles:* Count, type, lane. *Sampled vehicles:* Classification, occupants, speed, driver age and sex.	License plate, license type, state.
NIGHT	Inter-section	Signal	*All vehicles:* Count, type, lane. *Sampled vehicles:* Classification, driver age and sex, occupants, turn movement.	License plate, license type, state.
NIGHT	Inter-section	STOP, Red flasher	On separate approaches, *All vehicles:* Count, type, lane. *Sampled vehicles:* License plate, occupancy classification, driver age and sex, turn movement.	
NIGHT	Inter-section	YIELD, Uncontrolled, Amber flasher	*All vehicles:* Count, type, lane. *Sampled vehicles:* Classification, occupants, speed, turn movement, driver age and sex.	License plate, license type, state.

TABLE A.1-8
EQUIPMENT ASSIGNMENTS

	SITE	PROCEDURE	TEAM LEADER	TEAM MEMBER
DAY	Midblock	Standard	Counters, stereo recorder	Camera, radar, stereo recorder
	Inter-section	Signal control	Counters, stereo recorder	Camera, radar, stereo recorder
	Inter-section	STOP, Red flasher	Recorder	Recorder
NIGHT	Midblock	Standard	Counters, stereo recorder, radar	Binoculars, stereo recorder
	Inter-section	Signal control	Stereo recorder	Stereo recorder, binoculars
	Inter-section	STOP, Red flasher	Recorder	Recorder
	Inter-section	YIELD Uncontrolled, Amber flasher	Counter, stereo recorder, radar	Binoculars, stereo recorder

A.2 PRELIMINARY REDUCTION OF EXPOSURE DATA

The preliminary data reduction takes information off the audio cassettes and film and tabulates these data in preparation for computer data entry.

The information on the cassette will be transferred to one of three forms, as appropriate.

- Data Collection Log - Part B (Fig. A.2-1)
- Traffic Count Summary (Fig. A.2-2)
- Master Data Reduction Form (Fig. A.2-3)

Data Collection Log - Part B

Part B in the Data Collection Log contains the header information and all data on photography and on the time sequence of data collection. The time at which each interruption begins and ends is recorded to the nearest minute.

The "film used" table is completed from the recorded information whenever photography is used. All comments pertaining to the data collection process are transcribed to this form.

Traffic Count Summary

The Traffic Count Summary contains all data on traffic volumes collected. If mechanical counters are used, cumulative totals at approximately ten minute intervals are recorded on the audio cassette. These totals will be transferred to the form together with the times and tape counter readings.

If mechanical counters were not used and traffic volume data recorded directly--e.g., at stopped approaches--the audio cassette is played in real time and cumulative counts made. They are then recorded on the form at ten minute intervals.

All traffic count data are recorded even if, for any time increment, the counts are zero.

Master Data Reduction Form

The Master Data Reduction Form contains all data recorded by the team member during standard data collection or all data except traffic counts recorded for stopped approaches. The form is filled in while listening to the audio cassette, stopping and starting and rewinding, as required. Since tape counters vary, the same playback unit must be used for the entire data reduction process for a single audio cassette.

Each line in this form contains data describing a single sampled auto.

The form contains 15 columns of data. Columns 1 to 11 refer to cassette data, 12 to 16 are taken from photographs when available.

Site No. _____ Setup No. _____ Date _______ Time: Begin _____
End _____

Road Name or Number: ____________________ Weather: __________

Approach Direction
of Travel: ____________________ Roadway Surface: __________

Interruptions		1	2	3	4	5	6	7
Begin:	Counter							
	Time							
End:	Counter							
	Time							
Duration								
Reason								

Film Used	A	B	C	D	E	F
Film Type						
Beginning Frame No.						
Ending Frame No.						
Initial aperture f/						
Compensation Dial						
Data Back Setting						
Dial 1						
Dial 2						
Dial 3						

NOTES AND COMMENTS:

Figure A.2-1. Data Collection Log - Part B.

TRAFFIC COUNT SUMMARY Sheet ____ of ____

Site Number: _____ Reduced by: ________________ Date: _________

Set Up Number: ____ Checked by: ________________ Date: _________

Time Data Collection Started: _______ Ended: ________ Date: _______

Time	Tape Counter	Lane 1			Lane 2		
		Auto	Truck	Bus	Auto	Truck	Bus

Figure A.2-2. Traffic Count Summary.

MASTER DATA REDUCTION FORM

Site Number: ____________________ Sheet No. ______ of ______

Setup Number(Approach): __________ Reduced by: ______________

Film Roll No: ____________________ Checked by: ____________

Date: ____________ Date: ____________

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Tape Counter	Time	Color	Class	# of Occup.	Driver Sex	Age	Speed	Lane	Turn Mvmt	Frame No.	Lic Plate	State	Type of Regis	Make	Model

Figure A.2-3. Master Data Reduction Form.

(1) Tape Counter - The tape counter is reset to zero at the beginning of each cassette. The number shown on the counter at the time the description of a sampled vehicle starts is recorded.

(2) Time - Only those times recorded by the team member (i.e., Channel B) will be noted for the standard data collection, together with the tape counter reading. For stopped approach data collection, tape recorder readings will be recorded.
If a time is recorded the remainder of that line is left blank.

(3) Color - The color of the sampled vehicle (see Table A.1-4).

(4) Class - The size of the sampled vehicle (see Table A.1-4).

(5) Number of Occupants - The number of vehicle occupants (see Table A.1-4).

(6) Driver Age - The apparent age of the driver (see Table A.1-4).

(7) Driver Sex - The apparent sex of the driver (see Table A.1-4).

(8) Speed - The speed of the sampled vehicle to the nearest tenth of a mph. If speed was not obtained, mark "X". If speed was not measured during the entire data collection period, this column is left blank.

(9) Lane - The lane in which the vehicle is at the moment of speed measurement. The curb lane is No. 1.

(10) Turn Movement - This column is only used at intersections. Use R, S and L for Right, Straight, and Left, respectively. Use X for not observed.

(11) Frame No. - The frame number of the photograph applying to the sampled vehicle. The frame number recorded is that of the next picture taken. If no picture is taken, "X" is shown.

(12) License Plate - The license plate as recorded or as taken from photographs. When all or part of a license plate is recorded at the sites where photographs are also taken, this information is written in lightly when the cassette is reduced.

(13) State - The state of registration.

(14) Type of Registration - This column is filled in for NY registered vehicles only. The type of registration is written in small letters below the license plate number. Codes shown in Table A.2-1 are used.

(15) Make and (16) Model - These columns are filled in only if unequivocal make and model identification is possible from the photograph or has been recorded.

TABLE A.2-1
TYPE OF REGISTRATION

Numeric Code	Type of Registration
09	Foreign Consul-Diplomats
10	Locomotive
11	Special Passenger ($5 fee)
12	Special Passenger (No $5 fee)
13	Governor's Second Car
14	New York Senate
15	New York Assembly
16	Passenger or Suburban (Regular)
17	U.S. Congress
18	U.S. Senate
19	School Car
20	Hearse Coach (Hearse or Hearse Invalid Regular)
21	Historical
22	Special Reg. Hearse ($5)
23	Special Reg. Hearse (No $5 fee)
24	Limited Use Automobile
25	Court of Appeals
26	Special Purpose Commercial
28	Supreme Court (ADJ)
29	Medical Doctors
30	Court of Claims
31	Governor's Additional Car
32	Congressional Medal of Honor
33	Supreme Court Justice

Numeric Code	Type of Registration
36	Motorcycle
37	Limited Use Motorcycle Type A
38	Limited Use Motorcycle Type B
39	Limited Use Motorcycle Type C
46	Farm
51	Ambulance
52	(Special) Omnibus
53	(Public Service) Omnibus
54	(Taxi) Omnibus
55	(Livery) Omnibus
56	(Regular) Omnibus
57	(Vanity) Omnibus
62	Dealer
64	Motorcycle Dealer
66	Transporter
72	Agricultural Truck
76	(Regular) Commercial
77	State Agencies
78	(Household Carrier) Commercial
81	(Regular) Tractor
82	(Household Carrier) Tractor
85	(Commercial Semi-Trailer)
86	(Regular) Trailer
87	House or Coach Trailer
88	Political Subdivision (Municipal or Thruway)

A.3 PROCESSING OF EXPOSURE DATA

A.3.1 Introduction

Basic exposure data are derived from a pre-processed data base, in hard copy obtained and formatted as described in Appendix A.2. The product is a data tape which contains all available information on:

- The individual data collection sites;
- The individual vehicles sampled and observed during the data collection effort.

Some of the procedures and the associated computer software may be specific to the hardware and operating systems used. These include:

- Data Entry and creation of preliminary files--

  Data Entry was performed using a Motorola 6800 microprocessor and a commercially available text editor. This was used to create corrected data files stored temporarily on floppy disks. The data on disk was then spooled to the mainframe computer using proprietary communications software previously developed for the office microprocessor.

- File manipulation and creation of output tapes--

  File manipulation and tape handling software were written in FORTRAN for use on the Brookhaven National Laboratory CDC 6600/7600 computer facility.

The final output tapes merge data obtained by field observation with data obtained from the computerized files of the New York State Department of Motor Vehicles (NYDMV). The content and format of the NYDMV data are described in a document entitled "Jurisdiction Guide for Motor Vehicle Registration Information Requests--Revised: March 1980" obtainable from NYDMV.

A. 3.2 Overview

The flow of the data processing activities is shown in Fig. A.3-1. Data processing consists of:

- Quality control of field data reduction
- Enter data into computer storage
- Prepare tape for DMV
- Merge DMV (VIN) and field data and make validity checks
- Prepare final output tape

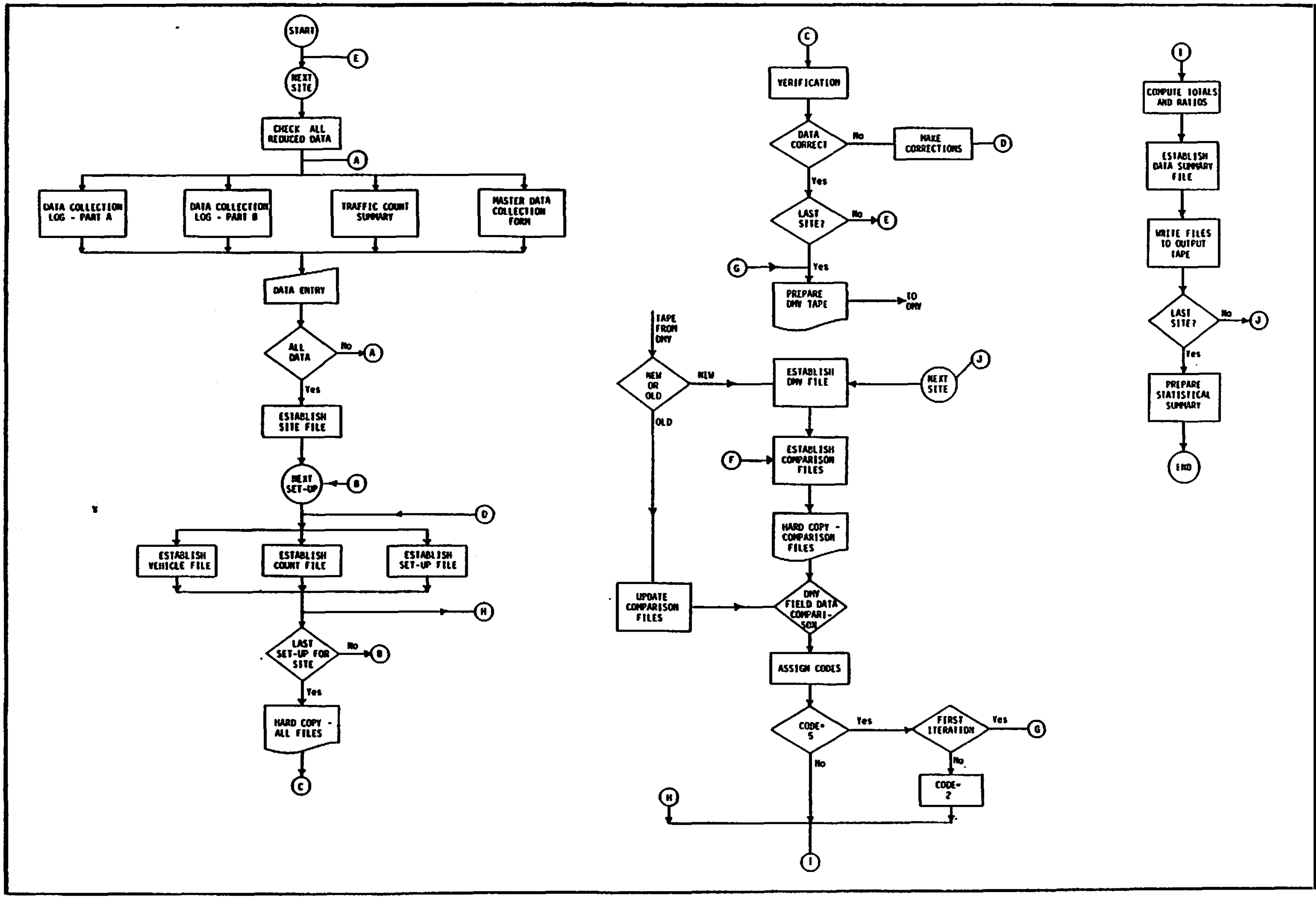


Figure A.3-1. Data Processing Flow Chart.

A.3.3 Quality Control of Field Data Reduction

The data obtained in the field, are on:

- Audio cassettes
- Developed film
- Data Collection Log
- Traffic Count Summaries
- Master Data Reduction Forms

Quality control checks of the data are performed before the data are entered into computer storage.

A visual examination of all completed data reduction forms is made to spot:

- Missing entries
- Data values outside of expected or admissible range
- Inconsistent entries
- Incorrect codes

Discrepancies and errors are corrected by re-examination of the original field data.

A portion of the data selected randomly is reduced again. The output of this effort is compared, on an entry by entry basis, with the output of the original field data reduction effort.

Errors uncovered during this quality control check are corrected.

A.3.4 Enter Data into Computer Storage

All data collected in the field are entered into diskettes via terminal. These data are then used to establish four separate files.

A.3.4.1 Site File

The site file contains all information on the physical characteristics of the site and some aspects of the data collection process. It is taken from the Data Collection Log - Part A. The format of this file is defined in Table A.3-1.

A.3.4.2 Set-Up File

The set-up file contains information on the data collection process. It is taken from the Data Collection Log - Part B. The format of this file is shown in Table A.3-2.

A.3.4.3 Count File

The count file contains information on vehicle counts made during the data collection process. It is taken from the Traffic Count Summary. The format of this file is defined in Table A.3.3.

TABLE A.3-1
FORMAT - SITE FILE
(For each site, eighteen lines are used. For lines 4 through 18, one field is available for each approach.)

Line 1	field 1	Site Number
	field 2	County
	field 3	City
	field 4	Town/Village
Line 2	....	Road Name and Number
Line 3	....	Cross Street for Intersection or exact location and landmark for mid-block location.
Line 4	(fields 1-4 .... as needed)	Roadway Name
5	....	Direction of Travel
6	....	Number of Lanes
7	....	Approach Width
8	....	Land Use
9	....	Horizontal Alignment
10	....	Vertical Alignment
11	....	Speed Limit
12	....	Pavement Surface
13	....	Passing
14	(field 1-4 .... as needed)	Parking
15	....	Control
16	....	Turning Lanes
17	....	Type of Data Collection
18	....	Set Up Number

The explanation of each item, and the codes used, will be found in Appendix A.1.

The indication for missing or inapplicable data items is -999.

A.3.4.4 Vehicle File

The vehicle file contains all the data recorded on the Master Data Reduction Form. The format for this file is shown in Table A.3-4.

A.3.4.5 Verification

After all data for one site have been entered, and these four files established, a hard copy printout of all the files is made. It is used for data verification to identify all anomalies and "suspicious" data items, to trace the source of the error, if any, and to correct it. Original data sources, cassettes and film are used to identify and resolve errors. After corrections are made, all files are stored on magnetic tape.

TABLE A.3-2
FORMAT - SET-UP FILE
(For each set-up, eight lines are used. For lines 3 through 8, one column applies to each interruption in the data collection process.)

Line 1	field 1	Site Number
	field 2	Set-Up Number
	field 3	Date of data collection
	field 4	Time data collection began
	field 5	Time data collection ended
Line 2	field 1	Road Name or Number
	field 2	Weather
	field 3	Approach Direction of Travel
	field 4	Roadway Surface
Line 3	(fields 1-4 .... as needed)	Tape counter reading at beginning of interruption
Line 4		Time at beginning of interruption
Line 5		Tape counter reading at end of interruption
Line 6		Time at end of interruption
Line 7		Duration of interruption
Line 8		Reason for interruption

The indication for missing or inapplicable data items is -999.

TABLE A.3-3
FORMAT - COUNT FILE
(For each set-up (approach) the number of lines used is two more than the number of count intervals recorded.)

Line 1	field 1	Site Number
	field 2	Set-Up Number
Line 2	field 1	Start Time
	field 2	End Time
Line 3	field 1	Time first cumulative count is recorded
	field 2	Tape recorder counter reading
	field 3	Passenger cars in lane 1
	field 4	Trucks in lane 1
	field 5	Buses in lane 1
	field 6	Passenger cars in lane 2
	field 7	Trucks in lane 2
	field 8	Buses in lane 2

Additional lines follow the format of line 3 for subsequent count intervals. The indication for missing or inapplicable data items is -999, except that fields 6 to 8 are left blank for one lane approaches.

TABLE A.3-4
FORMAT - VEHICLE FILE
(For each set-up the number of lines used is one more than the number of vehicles sampled.)

Line 1	field 1	Site Number
	field 2	Set-up Number
Line 2	field 1	Tape Counter
	field 2	Intermediate check time if recorded
	field 3	Color
	field 4	Vehicle Class
	field 5	Occupancy
	field 6	Driver Sex
	field 7	Driver Age
	field 8	Vehicle Speed/Stop Sign Observance
	field 9	Lane occupied
	field 10	Turning movement
	field 11	Photographic frame number
	field 12	License Plate
	field 13	State of registration
	field 14	Type of registration
	field 15	Vehicle make
	field 16	Vehicle model

Additional lines to follow the format of line 2 for subsequent vehicles. The indication for missing or inapplicable data items is -999.

A.3.5 Prepare Tape for DMV

Using the data contained in the Vehicle File, a tape is prepared in accordance to NYDMV specifications.

A.3.6 Make Validity Checks

Validity checks are made by comparing the data for each vehicle, which is returned by DMV, with the data collected in the field. A computer program, Program CODE, processes the tape containing the vehicle file (see Section A.3.4.4 above) and the tape returned by DMV. It creates a merged file which contains both field-recorded and DMV-furnished data for each vehicle sampled. The format of this output is defined in Table A.3-5.

The validity check compares vehicle descriptors as recorded in the field with the vehicle descriptors provided by the DMV files and determines if those two sets are in sufficient agreement.

TABLE A.3-5
FORMAT - PROGRAM CODE OUTPUT

Item No.	Start in Field	Line	Description
1	1	1	Set-Up Number
2	8	1	Sequential Vehicle Number
3	9	1	State of Registration
4	12	1	Data Collection Code
5	15	1	License Plate
		2	License Plate
6	26	1	Registration Type
		2	Registration Type
7	31	2	County of Registration
8	38	2	Zip Code of Registered Owner
9	44	2	VIN
10	64	2	Model Year
11	71	1	Make
		2	Make
12	81	2	Body Type
13	88	1	Color
		2	Color
14	95	2	Propulsion
15	101	1	Size Class
		2	GVW
16	107	1	Body Type
		2	Date of Registration

Note: Line 1 represents field data; line 2 represents DMV data.

The following should be noted:

- DMV will not return any data if the license plate is not valid; i.e., does not correspond to any record in its registration file. In that case, line 2 on the printout will contain the notation "NO LICENSE PLATE MATCH" and no other data. The original data, tape or photograph, are then checked to determine if the correct registration information, number, type, and format, had been transmitted to DMV. If any error is discovered, the appropriate correction is made in the Vehicle File and the revised record included in the next submission to DMV.

- The principal items of comparison are color and make (when available). Implausible combinations of make, body type and GVW, from the DMV tape, with class, make and model from the field data are noted as are differences in color. Model information can be obtained from the DMV tape by decoding the VIN. County of registration information is used to resolve uncertainty.

This comparison process assigns each vehicle a numerical code ranging from 1 to 8. Code definitions are as follows:

Code 1 - The observed license plate was matched by DMV and the two sets of vehicle descriptions are in adequate agreement.

Code 2 - The observed license plate could not be matched by DMV. Code 2 is not assigned unless there has been at least one resubmission to DMV.

Code 3 - Vehicle not registered in NY State.

Code 4 - The observed license plate was matched by DMV, however, there is insufficient agreement between the two sets of vehicle descriptions. [Note: This lack of correspondence could be due to a change in registration (transfer of plates) between the time the vehicle was observed and the time the data was obtained from DMV. It is possible to obtain information concerning prior registrations from the DMV data; however, the software to accomplish this was not developed].

Code 5 - To be rechecked with DMV.

Code 6 - Dealer or Transporter Plate - No vehicle data is available since these plates can be transferred without DMV knowledge.

Code 7 - The license plate was not observed in the field.

Code 8 - Police vehicle for which a DMV issued plate is not required.

The codes are added to the files. License plates assigned Code 5 are resubmitted to DMV.

A.3.7 Prepare Final Output Tape

After the validity checks, a final output tape is prepared. It consists of two parts. The first contains all the header information--data pertaining to the data collection site and the data collection process. The format for this portion of the output is defined in Table A.3-6. The second part of the tape contains all the vehicle data. The format for this portion is defined in Table A.3-7.

TABLE A.3-6
DATA TAPE FORMAT
Header Data

Record	Start in Field	Format	Description	Code for: Not Appl.	Code for: Unknown	Comments
1	1	A1	Record ID	-	-	"H"
1	2	A4	Site No.	-	-	No prefix-Ulster, Rural Prefix K-Kingston Prefix S-Schenectady
1	6	A1	County	-	-	U-Ulster S-Schenectady
1	7	A10	City	-999	-	
1	17	A10	Town/Village	-999	-	
2	1	2A10,A2	Road Name or Number	-	-	
3	1	8A10	Location Identification	-	-	Cross Street Name only if intersection
4	1	I1	Number of Approaches	-	-	
5	1	I3	Set-up Number	-	-	
5	4	A10	Road Name or Number	-	-	
5	14	A10	Type of Data Collection	-	-	
5	24	A2	Direction of Travel	-	-	
5	26	I1	Number of Lanes	-	-	Total Approach
5	27	A4	Approach Width	-	-	Both Directions
5	31	A1	Land Use	-	-	R-Rural, U-Urban
5	32	A2	Horizontal Alignment	-	-	See Note 1
5	34	A2	Vertical Alignment	-	-	See Note 2
5	36	A2	Speed Limit	-	-	U-Unposted
5	38	A2	Pavement Surface	-	-	See Note 3
5	40	A1	Passing Permitted	-	-	Y-Yes, N-No
5	41	A1	Parking Present	-	-	Y-Yes, N-No
5	42	A6	Control	X	-	
5	48	A1	Turning Lanes	X	-	
6	1	A8	Date	-	-	
6	9	I4	Time Started	-	-	Military Time
6	13	I4	Time Ended	-	-	Military Time
6	17	A6	Weather	-	-	
6	23	A6	Road Surface	-	-	
6	29	I5	Interruptions	-	-	Total Length of all interruptions
7	1	I3	Traffic Count	-	-	Cars-Lane 1
7	4	I3	Traffic Count	-	-	Trucks-Lane 1
7	7	I3	Traffic Count	-	-	Buses-Lane 1
7	10	I3	Traffic Count	-	-	Cars-Lane 2
7	13	I3	Traffic Count	-	-	Trucks-Lane 2
7	16	I3	Traffic Count	-	-	Buses-Lane 2
8	1		Same as Record 5 for 2nd set-up			See Note 4
8	49		Same as Record 6 for 2nd set-up			
8	82		Same as Record 7 for 2nd set-up			

(continued)

Table A.3-6 (Continued)

Record	Start in Field	Format	Description	Code for: Not Appl.	Code for: Unknown	Comments
9	1		Same as Record 5 for 3rd set-up			
9	49		Same as Record 6 for 3rd set-up			
9	82		Same as Record 7 for 3rd set-up			
10	1		Same as Record 5 for 4th set-up			
10	49		Same as Record 6 for 4th set-up			
10	82		Same as Record 7 for 4th set-up			
11	1		Same as Record 5 for 5th set-up			
11	49		Same as Record 6 for 5th set-up			
11	82		Same as Record 7 for 5th set-up			

Notes:

1. S or ST — Straight
   CR — Curve to Right
   CL — Curve to Left
   RC — Reverse Curve

2. F — Flat
   U — Upgrade
   D — Downgrade
   C — Crest
   S — Sag

3. A — Asphalt
   C — Concrete
   G — Gravel or Dirt

4. Provision is made for a maximum of five set-ups (approaches) for each site. If less than five approaches exist, the record is filled with blanks and a zero is entered for the set-up number (Record 8, 11, 14 & 17, Field 1-3) and for the number of lanes (Record 8, 11, 14 and 17, Field 26).

TABLE A.3-7
DATA TAPE FORMAT
Vehicle Data

(Each vehicle is one record)

Start in Field	Format	Description	Code for: Not Appl.	Code for: Unknown	Comments
1	A1	Record ID	-	-	"V"
2	I3	Set-up No.	-	-	
5	I3	Vehicle No.	-	-	Consecutive within set-up
8	A2	State	-	XX	
10	I2	Registration Type	-	0	DMV Code
12	A8	License Plate Number	-	-999	
20	I1	Recognition Code	-	-	See Note 1
21	A10,A7	VIN	-	-999	
38	I2	Year	-	0	Last two digits
40	A5	Make	-	Blank	
45	A4	Body Type	-	Blank	
49	2A2	Color	-	Blank	DMV Code
53	I1	No. of Cylinders	-	0	
54	A3	Propulsion	-	Blank	
57	I5	GVW	-	0	
62	I3	County of Registration	-	0	DMV Code
65	I5	Zip Code	-	0	
70	A4	Observed Car Size	-	-999	
74	I1	Vehicle Occupancy	-	0	
75	A1	Driver Sex	-	X	
76	A5	Driver Age	-	X	
81	A2	Stop Sign Observance	XX	XX	See Note 2
83	A4	Speed	XX	-999	
87	I1	Lane	-	0	
88	A1	Turning Movement	X	X	

Notes:

1. RECOGNITION CODE

    1. Observed license plate matched with DMV records.
    2. Observed license plate could not be matched with DMV records.
    3. Not NY registered vehicle.
    4. License plate match is questionable.
    5. To be rechecked with DMV.
    6. Dealer or Transporter Plate - No vehicle data available.
    7. License plate not observed.
    8. Police vehicle.

2. STOP SIGN OBSERVANCE

    0 Full Stop
    10 Rolling Stop
    50 No Stop

APPENDIX B
ACCIDENT DATA CODING

This appendix presents instruction for coding accident information (B.1), vehicle and driver information (B.2) from the police accident reports, and the code book for the resulting data tape (B.3).

Accidents were excluded from data base:

- If none of the vehicles involved is a passenger car registered in New York.
- If accident occurred off-roadway (i.e., on private property, in a parking lot, etc.).
- If accident occurred on interstate, limited access or dividing highway.
- If a driver was absent from all vehicles involved in the accident.
- If any pedestrian, bicyclist or moped was directly involved (i.e., struck or was struck by vehicle).
- If the accident involved the following:
  - three or more vehicles
  - hot pursuit
  - deliberate collision (malicious)
  - object falling from train crossing above
  - operator was shot
  - fire breaking out in vehicle, provided that the vehicle does not subsequently hit something or run off the road involuntarily.

The following is the accident report form used by the New York police agencies.

MV-104A (6/75)

State of New York – Department of Motor Vehicles

POLICE ACCIDENT REPORT

DMV COPY

Page ___ of ___ Pages

Local Codes

ACCIDENT DATE MO / DA / YR | DAY OF WEEK | TIME AM PM | NUMBER OF VEHICLES | NO. INJURED | NO. KILLED | NON-HIGHWAY | NOT INVESTIGATED AT SCENE | LEFT SCENE | POLICE PHOTOS YES NO

VEHICLE 1 | VEHICLE 2

LAST NAME DRIVER 1 | FIRST NAME | MIDDLE INITIAL | LAST NAME DRIVER 2 | FIRST NAME | MIDDLE INITIAL

NUMBER AND STREET | NUMBER AND STREET

CITY | STATE | ZIP CODE | CITY | STATE | ZIP CODE

DATE OF BIRTH MO / DA / YR | SEX | UNLICENSED | NUMBER OF OCCUPANTS | PUBLIC PROPERTY DAMAGED | DMV USE | DATE OF BIRTH MO / DA / YR | SEX | UNLICENSED | NUMBER OF OCCUPANTS | PUBLIC PROPERTY DAMAGED | DMV USE

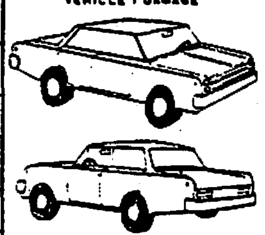
LAST NAME OWNER 1 | FIRST NAME | MIDDLE INITIAL | LAST NAME OWNER 2 | FIRST NAME | MIDDLE INITIAL

NUMBER AND STREET | NUMBER AND STREET

CITY | STATE | ZIP CODE | CITY | STATE | ZIP CODE

PLATE NUMBER | STATE OF REG. | YEAR & VEHICLE MAKE | VEHICLE TYPE | INS. CODE | PLATE NUMBER | STATE OF REG. | YEAR & VEHICLE MAKE | VEHICLE TYPE | INS. CODE

VEHICLE 1 DAMAGE

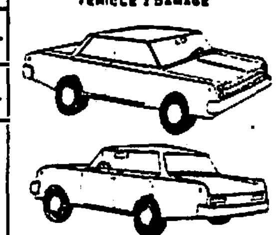


NO DAMAGE | UNDERCARRIAGE

ACCIDENT DIAGRAM

1. REAR END | 2. LEFT TURN | 3. RIGHT ANGLE | 4. RIGHT TURN | 5. HEAD ON | 6. OVERTAKING | 7. LEFT TURN | 8. RIGHT TURN | 9. SIDESWIPE

VEHICLE 2 DAMAGE



NO DAMAGE | UNDERCARRIAGE

VEHICLE BY TOWED TO

VEHICLE BY TOWED TO

REFERENCE MARKER | COUNTY | CITY / TOWN / VILLAGE OF | ADDRESS/LANDMARKS AT SCENE

ROUTE NO. OR STREET NAME | OR | MILES / FEET / AT INTERSECTION WITH | N S E W OF | ROUTE NO. OR STREET NAME

TICKET/ARREST: OPR 1 | OPR 2 | PEDESTRIAN | OTHER | TICKET/ARREST NUMBER(S) | VIOLATION SECTION(S)

ACCIDENT DESCRIPTION/OFFICER'S NOTES

USE COVER SHEET

B

		8	9	10	11	12	13	14	15	16	17	18	NAMES — IF DECEASED GIVE DATE OF DEATH
	A												
	B												
	C												
	D												
	E												
	F												
	G												

SIGN HERE | OFFICER'S RANK AND NAME | BADGE NO. | DEPARTMENT | PRECINCT/POST TROOP/ZONE | STATION/BEAT/SECTOR | REVIEWING OFFICER | DATE/TIME REVIEWED

ACCIDENT HEADER RECORD

2/23/82

A				B				C*	D*								E*							F
	Accident ID			Acc. Checklist					Date of Accid.			Time of Accident					Numbered Boxes							
Type of Record	Batch Number →	Sequence Number →		NY - Case Vehicles	Accident Type	Driver Presence	No Pedestrian/ Bicycle	1=City 2=Town 3=Village	Month (MM)	Date (DD)	Day of Week (first 2 letters)	Hour (HH)	Minute (MM)	'A'=AM 'P'=PM 'M'=military time	# vehicles in accident	# injured	# killed	#3	#4	#5	#6	#7	#24	#28	1=state nyw 2=other
A			0																						
01	02 03 04	05 06 07	08	09	10	11	12	13	14 15	16 17	18 19	20 21	22 23	24	25	26	27	28 29	30	31	32	33	34	35 36	37

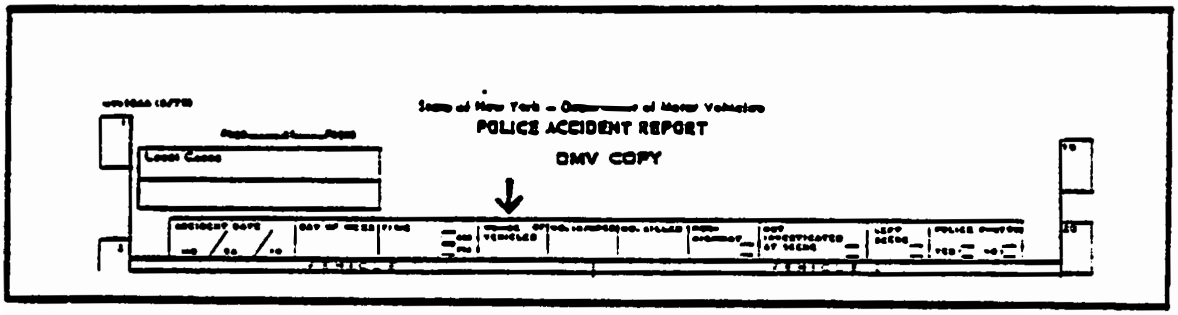
G					H																								I		
Int. Checklist					Vehicle #1 Status						Vehicle #2 Status						Vehicle #3 Status						Vehicle #4 Status									
In intersection?	# of intersection legs	> 1 vehicle?	Vehicle turning?	2 enter at angle?	(A)	(B)	(C)	(D)	(E)	(F) Case veh?	(A)	(B)	(C)	(D)	(E)	(F) Case veh?	(A)	(B)	(C)	(D)	(E)	(F) Case veh?	(A)	(B)	(C)	(D)	(E)	(F) Case veh?	Miscellaneous	# of case vehicles	County	Card Number
																															1	1
38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67 68	69	70	71

* See Accident Header Record Coding Guide for location of information on Police Accident Report.

(A) CODE BATCH NUMBER AND SEQUENCE AS FOLLOWS:

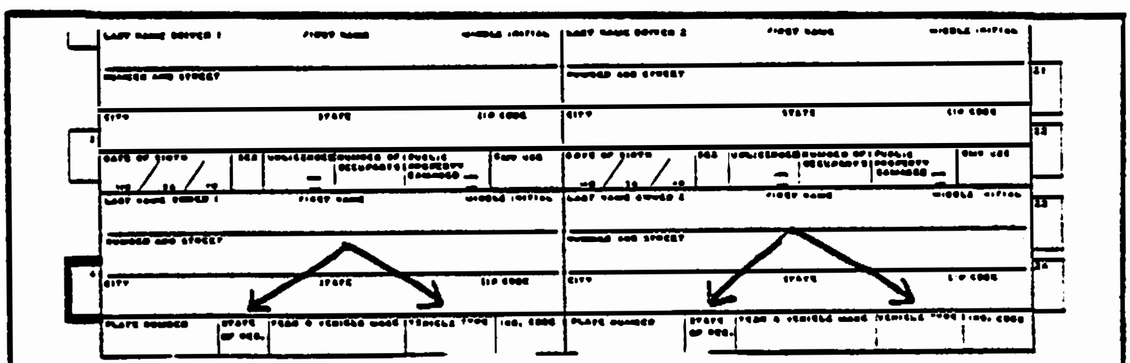
<u>Batch Number</u> - OBTAIN from front of envelope containing accident forms.

<u>Sequence Number</u> - Starting with 001, assign a unique sequence number to each accident report (in ascending order) within the batch being coded. Also record this sequence number on each accident form (in the upper right hand corner). For each <u>new</u> batch being coded, sequence numbers should start at 001.



(B) CODE ACCIDENT CHECKLIST (4 ITEMS) AS FOLLOWS)

- <u>NY - Case Vehicles</u>. If <u>none</u> of the vehicles involved in the accident is a passenger car registered in New York, code '1'. Otherwise, code '0'.

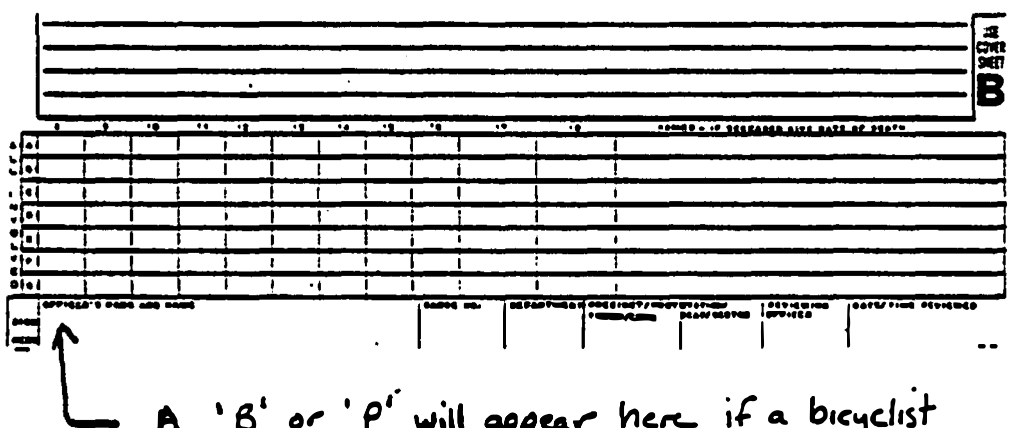


(B) CODE ACCIDENT CHECKLIST (CONTINUED)

- Accident Type. If accident is due to one of the following, code '1'. Otherwise, code '0'.
  - on private property or occurring off-highway
  - deliberate collision (malicious)
  - object falling from train crossing above
  - operator is shot
  - fire breaking out in vehicle, if the vehicle does not subsequently hit something or run off the road involuntarily.

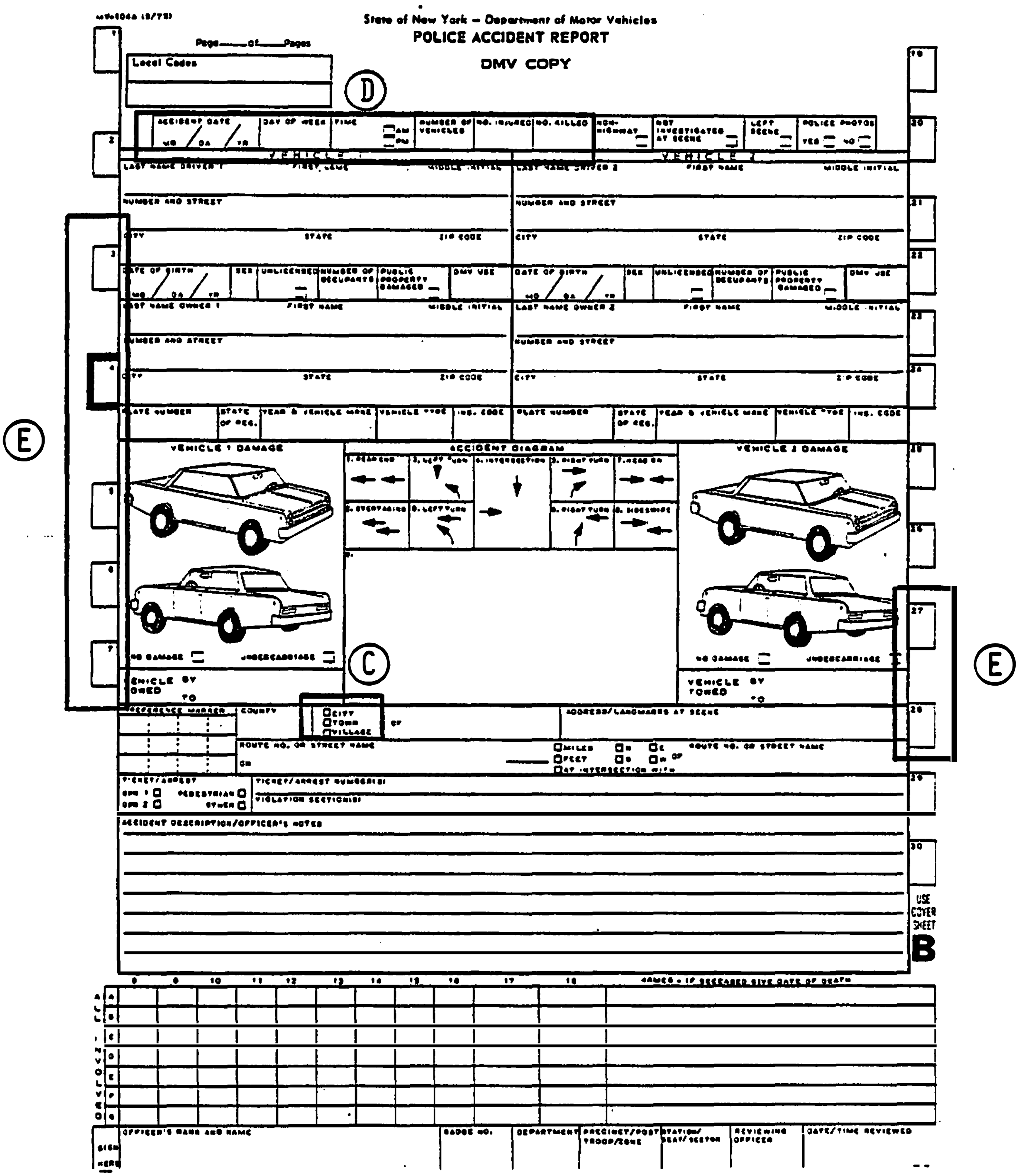
- Driver Presence. If driver is absent in all vehicles involved in the accident, code '1'. Otherwise, code '0'.

- No Pedestrian/Bicycle. If any pedestrian, bicyclist, or moped is involved (i.e., strikes or is struck by a vehicle), code '1'. Otherwise, code '0'.



A 'B' or 'P' will appear here if a bicyclist or pedestrian is involved. Check the 'Vehicle Type' boxes to see if a moped is involved.

NOTE: If a '1' has been coded for any of the 4 accident checklist items, do not code any further. Instead, clip the coding form to the accident report and set aside.



Coding Guide for Accident Header Record.

C See ACCIDENT HEADER RECORD CODING GUIDE for location of city/town/village information.

If unable to determine, leave blank.

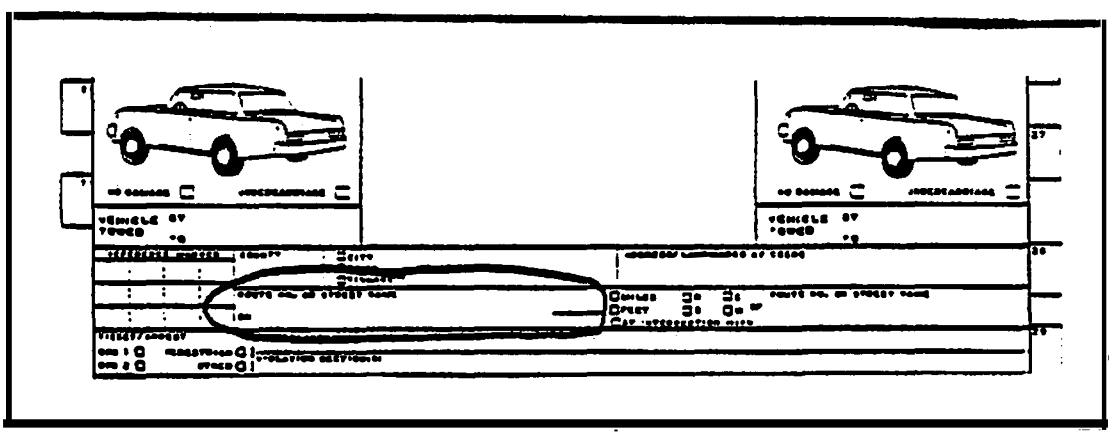
D See ACCIDENT HEADER RECORD CODING GUIDE for location of information. If the accident report does not contain this information (or if it is not legible), leave blank.

E See ACCIDENT HEADER RECORD CODING GUIDE for location of numbered boxes.

If box is blank, leave blank.

F If Route No. or Street Name appears in the LIST OF STATE HIGHWAYS, code '1'.

Otherwise, code '2'.



(F) LIST OF STATE HIGHWAYS

- **For All Batch Numbers:**

Rt 9W	Rt 44	Rt 212
Rt 28	Rt 55	Rt 213
Rt 28A	Rt 44/55	Rt 299*
Rt 32	Rt 208	Rt 375
Rt 32A	Rt 209	

* Does not include "Old Rt 299".

- **Non-Numbered State Highways:**

  - For Batch Numbers 401 to 404: Ulster Avenue
Partition Street
Malden Avenue (not Turnpike)
    - 404: Main Street
  - For Batch Numbers 601 to 602: Lawrence Road
Main Street
Freer Street
  - For Batch Numbers 101 and 102: South Chester Street
North Chester Street
Main Street
South Manhein
South Chestnut
  - For Batch Numbers 501 and 502: Albany Avenue
Boulevard
Broadway
East Chester Street
Flatbush Avenue
Wilbur Avenue
Wurts Street
  - For Batch Numbers 201 and 202: Tinker Street
  - For Batch Numbers 301 and 302 Ulster Avenue Mall

(G) CODE INTERSECTION CHECKLIST AS FOLLOWS:

- In Intersection? If accident occurred in the intersection between two roadways, code '1'.

  Otherwise, code '0' and go directly to Vehicle #1 status (leave columns 39-42 blank).

  If unable to determine, Code '9'.

- # of intersection legs. Code the total number of approaches to the intersection, if it can be determined from the accident diagram.

  If unable to determine, code '9'.

- >1 vehicle? If more than one vehicle (excluding vehicle(s) stopped at roadside or parked) was involved in the accident, code '1'. Otherwise, code '2', unless unable to determine, in which case, code '9'.

- Vehicle turning? If any vehicle was turning, code '1'.

  Otherwise, code '2'--unless unable to determine, in which case, code '9'.

- 2 enter at angle? If at least 2 vehicles entered the intersection at an angle (i.e., →↑) code '1'.

  Otherwise, code '2'--unless unable to determine, in which case, code '9'.

(H) CODE VEHICLE STATUS FOR EACH VEHICLE, AS FOLLOWS:

(A) If vehicle is not a passenger vehicle (i.e.,a truck, pick-up, van, motorcycle, commercial vehicle, etc.), code '1'.
Otherwise, code '0'.

(B) If the vehicle's state of registration is not New York, code '1'.
Otherwise, code '0'.

(C) If a driver was not present in the vehicle at the time of the accident, code '1'.
Otherwise, code '0'.

(D) If the vehicle was parked (as opposed to stopped in traffic or moving), code '1'.
Otherwise, code '0'.

(E) If the vehicle is off the roadway or on the shoulder, code '1'.
If vehicle is in roadway, code '0'.

(F) Case Veh: If a '1' is coded for either (A), (B), (C), (D) or (E), code '0'.
Otherwise, code '1'.

NOTE. If any of the above items are unknown, Code '9'.

(I) # OF CASE VEHICLES.

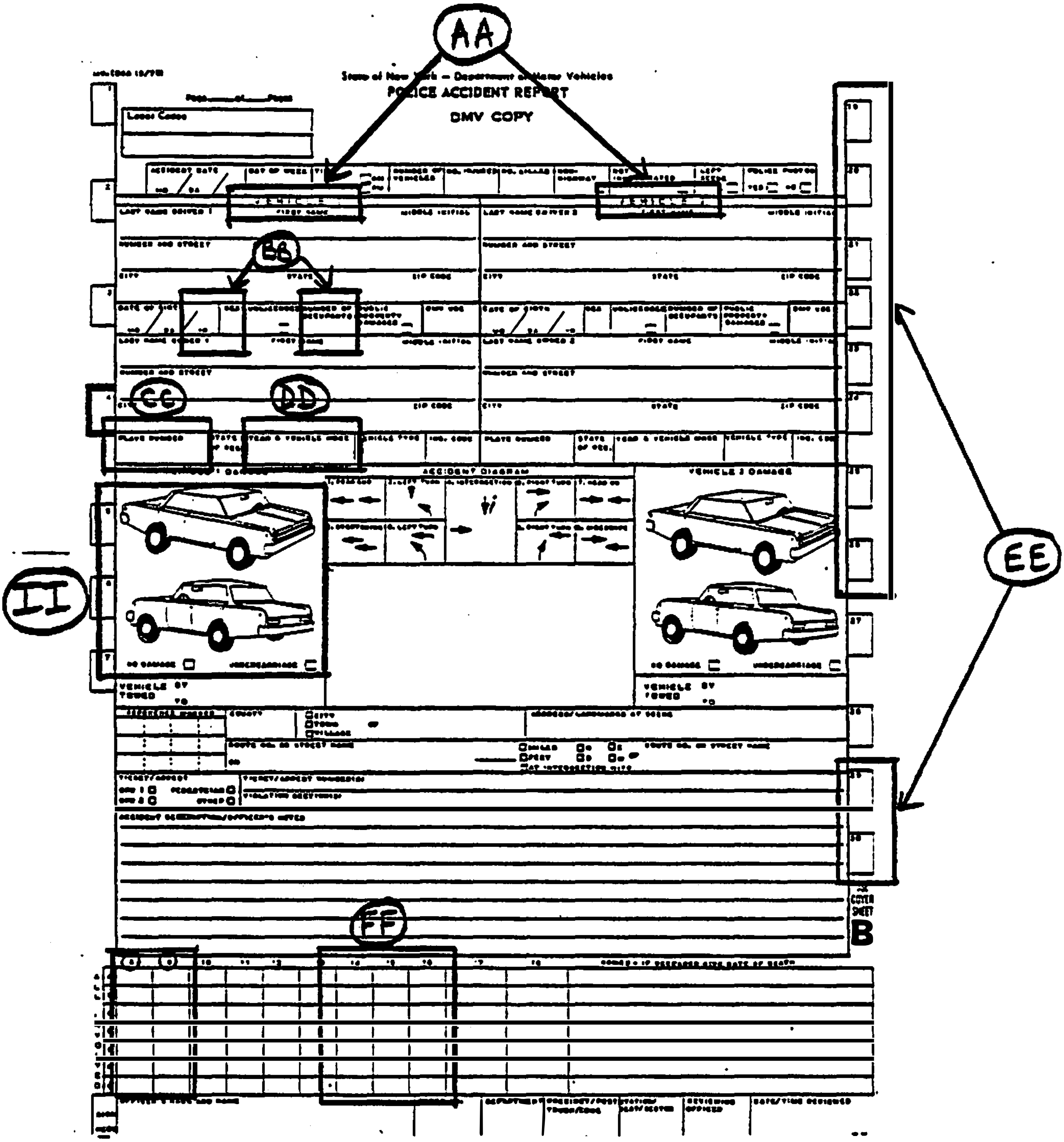
Count the number of '1' codes in columns 48 and 54, and enter the sum in column 69.

VEHICLE TRAILER RECORD

		AA*	XX	BB*			CC* 2/23/82	DD*	
	Vehicle ID						Vehicle Characteristics		
Type of Record	Batch Number / Sequence Number	Vehicle Number	Case Vehicle Code	Driver Year of Birth (YY)	Driver Sex 1=Male 2=Female	Number of Occupants	N.Y. License Plate Number (left justify; no embedded blanks)	Vehicle Model Year (last 2 digits)	Vehicle Make (1st 4 chars.)
V									
01	02 03 04 / 05 06 07	08	09	10 11	12	13 14	15 16 17 18 19 20 21 22	23 24	25 26 27 28

	EE*					FF*			GG		HH						II*		JJ
	Numbered Boxes										Pre-Accid.Factors								
If Veh. # is ODD	#19	#20	#23	#25	#29	Code only if Box #8 = Vehicle # and Box #9 = '1'			'Y' or "T" Inter.	Vehicle Towing?	Basic Maneuver	Turn Direction	Special Maneuver	Passing Direction	Lane Position	Outcome	Initial Point of Impact	Miscellaneous	Card Number
If Veh. # is EVEN	#21	#22	#24	#26	#30	#14	#15	#16											
	29 30	31 32	33	34 35	36 37	38 39	40 41	42	43	44	45	46	47	48	49	50	51 52	53 54	55

* See Vehicle Trailer Record Coding Guide for location of information on Police Accident Report.



CODING GUIDE FOR VEHICLE TRAILER RECORD

(XX) CASE VEHICLE CODE.

Check "Vehicle Status (F)" on Accident Header Record Form. If '1' is coded, code '1' here. Otherwise, do not code vehicle ... set aside.

(CC) NEW YORK LICENSE PLATE NUMBER

No embedded blanks --i.e., 'BARB F' is coded as 'BARBF'. Left justify all plate numbers.

(5G) 'Y' OR 'T' INTERSECTION

Code as follows:

0 = accident did not occur in intersection.
1 = accident did not occur in 'Y' or 'T' intersection.
2 = vehicle approaching on trunk of a 'Y' intersection.
3 = vehicle approaching on leg of a 'Y' intersection.
4 = vehicle approaching on trunk of a 'T' intersection.
5 = vehicle approaching on leg of a 'T' intersection.
9 = unable to determine.

Vehicle Towing?

If vehicle is towing something (i.e., boat, trailer, campter, etc.), code '1'; otherwise, code '0'.

(HH) PRE-ACCIDENT FACTORS (see accident diagram and accident description/officer's notes)

- Basic Maneuver.

  1 = vehicle following road (includes vehicle topped in traffic, but only if it is not parked or experiencing breakdown, etc.).

  2 = vehicle turning.

  3 = vehicle entering or crossing traffic way (from off-road position-driveway, parking lot).

  4 = other (set accident aside).

  9 = unable to determine from accident report.

- Turn Direction. If vehicle is not turning, code '8'. Otherwise, code as follows:

  1 = right turn.

  2 = left turn.

  3 = 'U' turn.

  9 = Direction not specified in accident report.

- Special Maneuver.

  1 = starting from parked position.

  2 = backing (including backing into parking position).

  3 = parking (except backing into parking position).

  4 = passing/overtaking.

  8 = none of the above.

  9 = unable to determine from accident report.

- Passing Direction. If vehicle is not passing/overtaking, Code '8'. Otherwise, code as follows:

  1 = to the right.

  2 = to the left.

  9 = direction not specified in accident report.

(HH) PRE-ACCIDENT FACTORS (continued)

- <u>Lane Position</u>.

  1 = vehicle in lane(s) for travel direction.

  2 = vehicle in lane(s) for opposite direction of travel.

  3 = vehicle is straddling center line/center of road (i.e., partially in lanes for <u>both</u> directions of travel).

  4 = vehicle is entering or crossing lane(s) at angle (does <u>not</u> include changing lanes; <u>does</u> apply to turning maneuvers).

  9 = unable to determine from accident report.

- <u>Outcome</u>

  1 = colliding with vehicle travelling in same direction.

  2 = colliding with vehicle travelling in opposite direction.

  3 = colliding with vehicle crossing (at angle) or entering.

  4 = striking object <u>in</u> raodway. A parked vehicle in road (including legally parked at curbside) is an "object in the road").

  5 = running off road and/or striking roadside object.

  6 = other (including combination of 2 or more of the above). If '6' (other) is coded, set accident aside.

  9 = unable to determine from accident report.

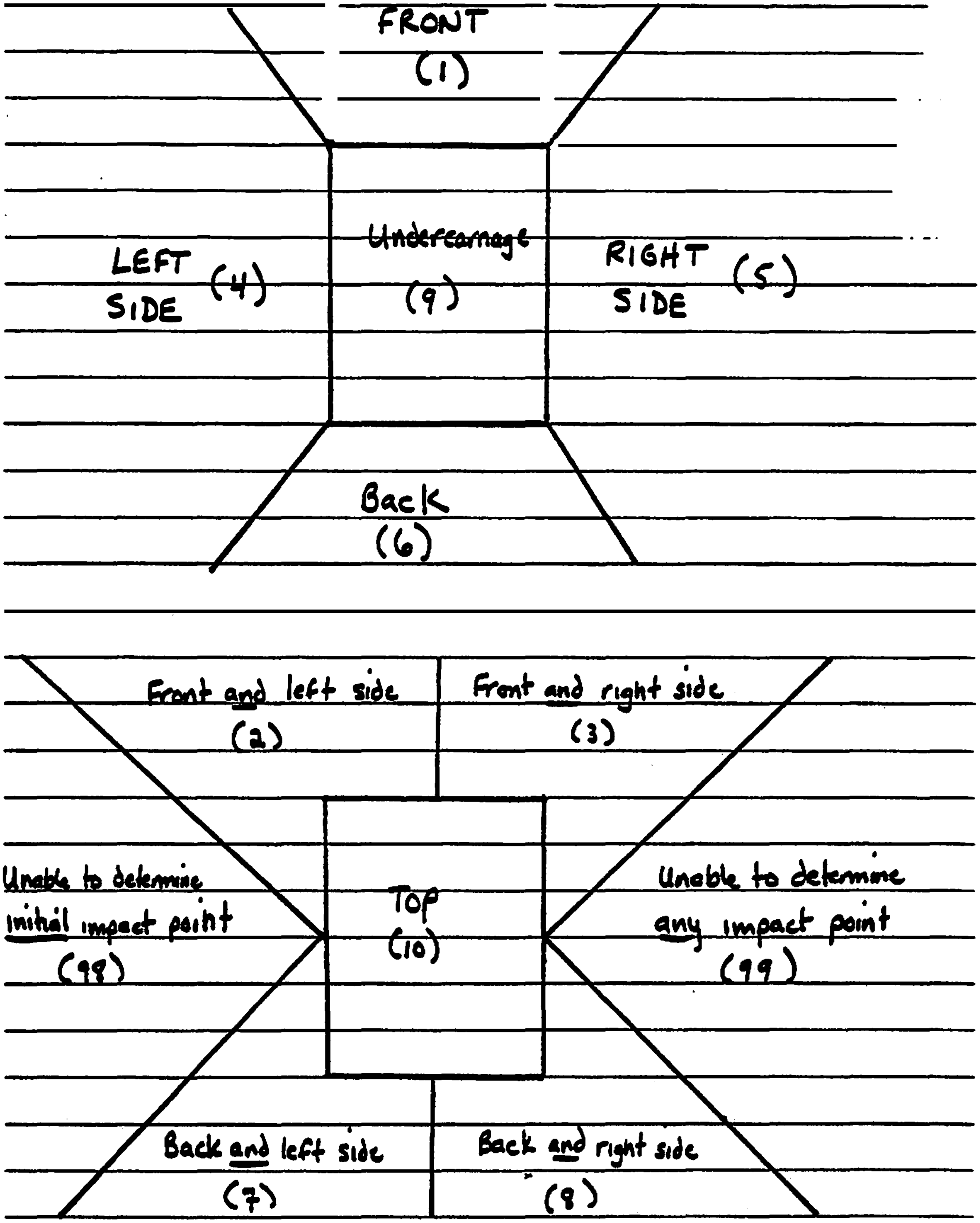
(JJ) CARD NUMBER

Code as follows:

2 = first vehicle coded for accident.

3 = second vehicle coded for accident.

II INITIAL POINT OF IMPACT (<u>not</u> necessarily point of most extensive damage).



96 = No damage

97 = Overturn

B.3 Accident Data Code Books

See following pages.

ACCIDENT HEADER RECORD

Column(s)	Variable	Code(s)
1	Type of Record	'A'
2-4	Batch Number	See Table A-1
5-7	Sequence Number	-
8	Vehicle Number	'0'
9-12	Filler	-
13	Jurisdiction	0 = Missing 1 = City 2 = Town 3 = Village
14-15	Month of Accident	0 = Missing
16-17	Day of Month	0 = Missing
18-19	Day of Week	First 2 letters coded
20-21	Hour	0,99 = Missing
22-23	Minute	0,99 = Missing
24	Time Code	'A' = AM 'P' = PM 'M' = Military Time 'Blank' = Missing
25	# of Vehicles in Accident	0 = Missing
26	# Injured in Accident	-
27	# Killed in Accident	-
28-29	Traffic Control	0 = Missing 1 = None 2 = Traffic Signal 3 = Stop Sign 4 = Flashing Light 5 = Yield Sign 6 = Officer/Flagman/Guard 7 = No Passing Zone 8 = RR Crossing Sign 9 = RR Crossing Flashing Light 10 = RR Crossing Gates 11 = Stopped School Bus - Red Lights Flashing 20 = Other

ACCIDENT HEADER RECORD

Column(s)	Variable	Code(s)
30	Light Condition	0 = Missing 1 = Daylight 2 = Dawn 3 = Dusk 4 = Dark-Road Lighted 5 = Dark-Road Unlighted
31	Roadway Character	0 = Missing 1 = Straight and Level 2 = Straight and Grade 3 = Straight at Hillcrest 4 = Curve and Level 5 = Curve and Grade 6 = Curve at Hillcrest
32	Roadway Surface Condition	0 = Missing 3 = Muddy 1 = Dry 4 = Snow/Ice 2 = Wet 5 = Slush
33	Weather	0 = Missing 1 = Clear 2 = Cloudy 3 = Rain 4 = Snow 5 = Sleet/Hail/Freezing Rain 6 = Fog/Smog/Smoke
34	Location of First Event	0 = Missing 1 = On Roadway 2 = Off Roadway
35-36	Type of Accident (First Event)	0 = Missing COLLISION WITH: 1 = Other Motor Vehicle 2 = Pedestrian 3 = Bicyclist 4 = Animal 5 = Railroad Train 10 = Other Object (Not Fixed) COLLISION WITH FIXED OBJECT 11 = Light Support/Utility Pole 12 = Guide Rail 13 = Crash Cushion 14 = Sign Post 15 = Tree 16 = Building/Wall 17 = Curbing 18 = Fence 19 = Bridge Structure 20 = Culvert/Head Wall 21 = Median/Barrier

ACCIDENT HEADER RECORD

Column(s)		
35-36 (Cont'd)		22 = Snow Embankment 23 = Earth Embankment/Rock Cut/Ditch 24 = Fire Hydrant 30 = Other Fixed Object <u>NON-COLLISION</u> 31 = Overturned 32 = Fire/Explosion 33 = Submersion 34 = Ran Off Roadway Only 40 = Other
37	Type of Highway	0 = Missing 1 = State Highway 2 = Other
38	Did Accident Occur in Intersection Between Two Roadways?	0 = No 1 = Yes 9 = Unable to Determine
39	Total Number of Approaches to the Intersection	0 = Does Not Apply 1 = One 2 = Two 3 = Three 4 = Four 9 = Unable to Determine
40	# of Vehicles Involved in Intersection Accident (excluding vehicles stopped at at roadside or parked)	0 = Does Not Apply 1 = More than One 2 = One Only 9 = Unable to Determine
41	Was Any Vehicle Turning at Intersection?	0 = Does Not Apply 1 = Yes 2 = No 9 = Unable to Determine
42	Did at Least 2 Vehicles Enter the Intersection <u>at an Angle</u> (on <u>separate</u> approaches)?	0 = Does Not Apply 1 = Yes 2 = No 9 = Unable to Determine
Veh. 1: 43, Veh. 2: 49, Veh. 3: 55, Veh. 4: 61	Vehicle Type	0 = Passenger 1 = Non-Passenger 9 = Unknown
44 50 56 62	Registration	0 = New York 1 = Other 9 = Unknown

ACCIDENT HEADER RECORD

Column(s)				Variable	Code(s)
Veh. 1	Veh. 2	Veh. 3	Veh. 4		
45	51	57	63	Driver Present?	0 = Yes 1 = No 9 = Unknown
46	52	58	64	Vehicle Parked?	0 = No 1 = Yes 9 = Unknown
47	53	59	65	Vehicle Off Roadway?	0 = No 1 = Yes 9 = Unknown
48	54	60	66	Case Vehicle?	0 = No 1 = Yes
67-68				Filler	-
69				# of Case Vehicles Coded	-
70				County	1 = Ulster 2 = Schenectady
71				Card Number	'1'

TABLE A-1
BREAKDOWN OF BATCH NUMBERS

Batch No.	Jurisdiction*	County	Police	Year	Months
101	New Paltz	Ulster	Local	1980	Apr-Nov
102	New Paltz	Ulster	Local	1981	Apr-Oct
201	Woodstock	Ulster	Local	1980	Apr-Nov
202	Woodstock	Ulster	Local	1981	Apr-Oct
301	Ulster	Ulster	Local	1980	Apr-Nov
302	Ulster	Ulster	Local	1981	Apr-Oct
401	Saugerties	Ulster	Local	1980	Apr-Nov
402	Saugerties	Ulster	Local	1981	Apr-Oct
403	Saugerties (Village)	Ulster	Local	1980	Apr-Nov
404	Saugerties (Village)	Ulster	Local	1981	Apr-Oct
501	Kingston (City)	Ulster	Local	1980	Apr-Nov
502	Kingston (City)	Ulster	Local	1981	Apr-Oct
601	Rosendale	Ulster	Local	1980	Apr-Nov
602	Rosendale	Ulster	Local	1981	Apr-Oct
701	Lloyd	Ulster	Local	1980	Apr-Nov
702	Lloyd	Ulster	Local	1981	Apr-Oct
751	Duanesburg	Schenectady	State	1981	Sep-Oct
752	Glenville	Schenectady	State	1981	Sep-Oct
753	Niskayuna	Schenectady	State	1981	Sep-Oct
754	Rotterdam	Schenectady	State	1981	Sep-Oct
755	Scotia (Village)	Schenectady	State	1981	Sep-Oct
801	Ulster County	Ulster	Sheriff	1980	Apr-Nov
802	Ulster County	Ulster	Sheriff	1981	Apr-Oct
851	Ulster County	Ulster	State	1980	Apr-Nov
852	Ulster County	Ulster	State	1981	Apr-Oct

*Town, unless otherwise indicated.

Vehicle Trailer Record

Column(s)	Variable	Code(s)
1	Type of Record	'V'
2-4	Batch Number	See Table A-1
5-7	Sequence Number	-
8	Vehicle Number	-
9	Case Vehicle Code	'1'
10-11	Driver Year of Birth	0,99 = Missing
12	Driver Sex	0 = Missing 1 = Male 2 = Female
13-14	# of Occupants	0,99 = Missing
15-22	NY License Plate Number	-
23-24	Vehicle Model Year	0,99 = Missing
25-28	Vehicle Make	First 4 characters
29-30 31-32	Apparent Contributing Factors - I Apparent Contributing Factors - II	0 = Missing HUMAN 2 = Alcohol Involvement 3 = Backing Unsafely 4 = Driver Inattention 5 = Driver Inexperience 6 = Drugs (Illegal) 7 = Failure to Yield Right-of-Way 8 = Fell Asleep 9 = Following Too Closely 10 = Illness 11 = Lost Consciousness 12 = Passenger Distraction 13 = Passing or Lane Usage Improper 14 = Pedestrian's Error/ Confusion 15 = Physical Disability 16 = Prescription Medication 17 = Traffic Control Disregarded 18 = Turning Improperly 19 = Unsafe Speed 40 = Other Human

Column(s)	Variable	Code(s)
29-30 31-32 (Cont'd)		<u>VEHICULAR</u> 41 = Accelerator Defective 42 = Brakes Defective 43 = Headlights Defective 44 = Other Lighting Defects 45 = Oversized Vehicle 46 = Steering Failure 47 = Tire Failure/Inadequate 48 = Tow Hitch Defective 49 = Windshield Inadequate 60 = Other Vehicular <u>ENVIRONMENTAL</u> 61 = Animal's Action 62 = Glare 63 = Lane Marking Improper/ Inadequate 64 = Obstruction/Debris 65 = Pavement Defective 66 = Pavement Slippery 67 = Shoulders Defective/ Improper 68 = Traffic Control Device Improper/Non-Working 69 = View Obstructed/Limited 80 = Other Environmental
33	Direction of Travel	0 = Missing 5 = S 1 = N 6 = SW 2 = NE 7 = W 3 = E 8 = NW 4 = SE
34-35	Pre-Accident Vehicle Action	0 = Missing 1 = Going Straight Ahead 2 = Making Right Turn 3 = Making Left Turn 4 = Making U Turn 5 = Starting from Parking 6 = Starting in Traffic 7 = Slowing or Stopping 8 = Stopped in Traffic 9 = Entering Parked Position 10 = Parked 11 = Avoiding Object in Roadway 12 = Changing Lanes 13 = Overtaking 14 = Merging 15 = Backing 20 = Other

Column(s)	Variable	Code(s)
36-37	Second Event	See Accident Header Record - Type of Accident (First Event), Cols. 35-36
38-39	Location of Driver's Most Severe Physical Complaint	0 = Missing or N/A 1 = Head 2 = Face 3 = Eye 4 = Neck 5 = Chest 6 = Back 7 = Shoulder-Upper Arm 8 = Elbow-Lower Arm-Hand 9 = Abdomen - Pelvis 10 = Hip-Upper Leg 11 = Knee-Lower Leg-Foot 12 = Entire Body
40-41	Type of Driver's Physical Complaint	0 = Missing or N/A 1 = Amputation 2 = Concussion 3 = Internal 4 = Minor Bleeding 5 = Severe Bleeding 6 = Minor Burn 7 = Moderate Burn 8 = Severe Burn 9 = Fracture - Dislocation 10 = Contusion - Bruise 11 = Abrasion 12 = Complaint of Pain 13 = None Visible
42	Driver's Physical and Emotional Status	0 = Missing 1 = Apparent Death 2 = Unconscious 3 = Semiconscious 4 = Incohorent 5 = Shock 6 = Conscious
43	'Y' or 'T' Intersection Approach	0 = Accident did not occur in intersection 1 = Accident did not occur in 'Y' or 'T' intersection 2 = Vehicle approaching on trunk of 'Y' intersection 3 = Vehicle approaching on leg of 'Y' intersection 4 = Vehicle approaching on trunk of 'T' intersection 5 = Vehicle approaching on leg of 'T' intersection 9 = Unable to determine

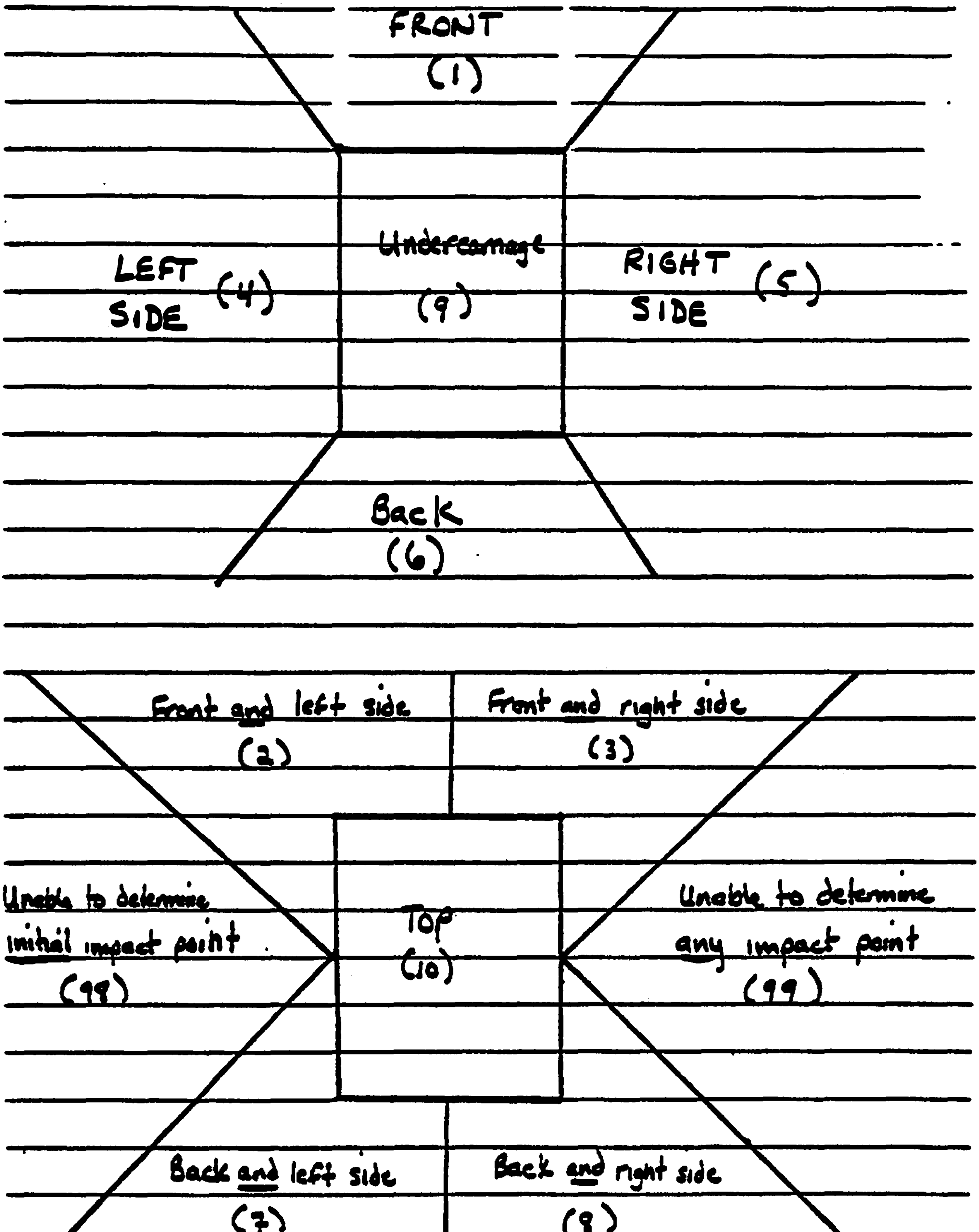
Column(s)	Variable	Code(s)
44	Was Vehicle Towing Something?	0 = No 1 = Yes
45-49	Pre-Accident Factors:	
45	Basic Maneuver	1 = Vehicle following road (includes vehicle stopped in traffic, but only if it is not parked or disabled) 2 = Vehicle turning 3 = Vehicle entering or crossing traffic-way (from off-road position--i.e., driveway, parking lot, etc.) 9 = Unable to determine
46	Turn Direction	1 = Right Turn 2 = Left Turn 3 = 'U' Turn 8 = Vehicle Not Turning 9 = Unable to Determine
47	Special Maneuver	1 = Starting from Parked Position 2 = Backing (including backing into parking position 3 = Parking (except backing into parking position 4 = Passing/Overtaking 8 = None of the Above 9 = Unable to Determine from Accident Report
48	Passing Direction	1 = To the Right 2 = To the Left 8 = Vehicle Not Passing/Overtaking 9 = Unable to Determine
49	Lane Position	1 = Vehicle in lane(s) for travel direction 2 = Vehicle in lane(s) for oppo-direction of travel 3 = Vehicle straddling center-line/center of road (i.e., partially in lanes for both directions of travel 4 = Vehicle entering or crossing lane(s) at angle (does not apply to changing lanes; does apply to turning maneuvers 9 = Unable to Determine

Vehicle Trailer Record

Column(s)	Variable	Code(s)
50	Outcome	1 = Colliding with vehicle traveling in same direction 2 = Colliding with vehicle traveling in opposite direction 3 = Colliding with vehicle crossing or entering trafficway 4 = Striking object in roadway (includes vehicles parked at curbside) 5 = Running off road and/or striking roadside object 9 = Unable to Determine
51-52	Initial Point of Impact	See Table V-1
53-54	Filler	-
55	Card Number	2 = First vehicle coded for accident 3 = Second vehicle coded for accident

TABLE V-1

INITIAL POINT OF IMPACT CODES



96 = No damage

97 = Overturn