Generation and Assessment of Incident Management Strategies

Volume II: Analysis of Freeway Incidents in the Seattle Area

WA-RD 204.3

Technical Report
January 1990

Washington State Department of Transportation
Planning, Research and Public Transportation Division

in cooperation with the
United States Department of Transportation
Federal Highway Administration

TECHNICAL REPORT STANDARD TITLE PAGE

1. REPORT NO. WA-RD 204.3	2. GOVERNMENT ACCESSION NO.	3. RECIPIENTS CATALOG NO.
4. TITLE AND SUBTITLE GENERATION AND ASSESSMENT OF INCIDENT MANAGEMENT STRATEGIES, VOL. II		5. REPORT DATE January 1990
		6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Fred Mannering, Assistant Professor, Civil Engineering Bryan Jones, Dan H. Garrison, Brad Sebranke, and Lester Janssen, Research Assistants, Civil Engineering		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Washington State Transportation Center (TRAC) University of Washington, JE-10 The Corbet Building, Suite 204; 4507 University Way N.E. Seattle, Washington 98105		10. WORK UNIT NO.
		11. CONTRACT OR GRANT NO. GC8286, Task 23
12. SPONSORING AGENCY NAME AND ADDRESS Washington State Department of Transportation Transportation Building, KF-10 Olympia, Washington 98504		13. TYPE OF REPORT AND PERIOD COVERED Technical Report
		14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

This study was conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration.

16. ABSTRACT

This four-volume technical report describes a study of freeway incidents and incident management strategies in the Seattle area. The study statistically analyzed the frequency and duration of freeway incidents on sections of I-5 and SR 520 in Seattle. In addition, a traffic simulation model was operationalized to assess the traffic related impacts of incidents. The findings show that Seattle-area incident management currently responds well to inclement weather and special events (e.g., major sporting games) but has problems with severe accidents. The ongoing operationalization of accident investigation sites and incident equipment storage sites can be expected to improve severe accident management, but response personnel training and the addition of more dedicated tow truck service are also needed. Finally, the study shows that, from a traffic impact perspective, the section of I-5 in downtown Seattle is in need of the most incident management attention.

17. KEY WORDS Incident evaluation, incident management, incident frequency, incident duration, traffic impacts, accident analysis	18. DISTRIBUTION STATEMENT No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22616

19. SECURITY CLASSIF. (of this report)	20. SECURITY CLASSIF. (of this page)	21. NO. OF PAGES	22. PRICE
None	None	85	

DOT FORM 310-022 (11/86) -1216-

Technical Report

Research Project GC 8286, Task 23
Incident Management Strategies

GENERATION AND ASSESSMENT OF INCIDENT MANAGEMENT STRATEGIES

VOLUME II: ANALYSIS OF FREEWAY INCIDENTS IN THE SEATTLE AREA

by

Bryan Jones
Research Assistant

Lester Janssen
Research Assistant

Brad Sebranke
Research Assistant

Dr. Fred Mannering
Associate Professor of Civil Engineering

Department of Civil Engineering
University of Washington, FX-10
Seattle, Washington 98195

Washington State Transportation Center (TRAC)
University of Washington, JE-10
The Corbet Building, Suite 204
4507 University Way N.E.
Seattle, Washington 98105

Washington State Department of Transportation
Technical Monitors
Susan Everett and John Conrad

Prepared for

Washington State Transportation Commission
Department of Transportation
and in cooperation with
U.S. Department of Transportation
Federal Highway Administration

January 1990

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Washington State Transportation Commission, Department of Transportation, or the Federal Highway Administration. This report does not constitute a standard, specification, or regulation.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF TABLES

SUMMARY

The collection of appropriate data for incident frequency analysis is relatively straightforward. Accident reports are available from the State Department of Transportation on diskette. Duration data are harder to come by, and the collection of data from dispatch logs is tedious and extremely time consuming. However, the Washington State Patrol installed a Computer Aided Dispatch system in March 1988, and data collection for subsequent studies should be greatly facilitated, perhaps as simple as collection of accident report data. The study team collected accident and dispatch data, found matching accidents, and used them for the incident and duration analysis.

Frequency and duration models were estimated for one of six zones in the study area. The study team found that there were considerable differences in model parameters from zone to zone because of different zone characteristics.

Frequency data were aggregated for each day of the study and used in a Poisson regression analysis for accidents and disablements. Logged times from the WSP dispatch logs were used to determine the lengths of accidents and subsequently used for a duration analysis by Survival Analysis.

The incident frequency models indicated that seasonal, day of week, special events, and environmental factors are important in forecasting accidents per day. Models were compared by coefficient sign, magnitude, and significance, as indicated by their t-statistic. Rho-squared values for the Poisson models indicated that a considerable amount of variability in accidents per day could be explained.

The duration models showed that accident duration is not so much a function of season or day of week, but of time of day variables such as rush hour or traditional recreational times and of specific accident characteristics, including number of vehicles involved, number of injuries, or truck involvement.

This work has uncovered many of the variables that contribute to incident frequency and duration on an area-wide basis; further work utilizing Poisson regression and Survival analysis could be done to determine the frequency and duration of different incident types (e.g., truck accidents or injury accidents). This kind of investigation could contribute to a more detailed evaluation of specific accident type and to more detailed site management policies.

CHAPTER 1

INTRODUCTION

Traffic congestion has become one of the most pressing public issues of this decade, and the Seattle area is no exception. The major newspapers in the Seattle area have added regular columns about traffic and transportation (e.g., "Getting Around" in *The Seattle Times*). A local television station (KIRO, channel 7) has included traffic congestion as one of its top three editorial priorities, and at least one Seattle councilperson was elected on a pro-light rail platform, since light rail is seen as a possible solution to the increasing congestion. While traffic congestion becomes a hotter political issue, federal matching funds for construction of new highways are drying up, and local tax dollars are more difficult to acquire in a more competitive atmosphere. One alternative to increasing capacity is to maximize existing capacity. Unfortunately, expansion of existing facilities beyond traffic management strategies and routine maintenance is less feasible than in the past.

Much of the congestion around Seattle is recurrent; that is, it can be expected every day at the same times. Traffic demand strategies have been introduced to address this recurrent congestion (e.g., HOV lanes and ride-sharing), but these have proven relatively ineffective in relieving the everyday rush hour crunch. Non-recurrent congestion, which is caused by seemingly random events such as car accidents or disablements, frequently shuts down at least one lane of traffic, which results in a capacity reduction highlighted by the creation of temporary merge and weave sections. The primary objective of this study was to identify and quantify the factors that increase or decrease the frequency and duration of non-recurrent congestion in the Seattle area.

This report begins with documentation of the study's data collection effort, followed by a summary of the kinds of analysis that have historically been used to quantify the factors that affect incident frequency and duration (i.e., the factors affecting non-recurrent congestion) and a review of papers introducing some new analysis techniques. Next, univariate and cross tabulated assessments of the collected data are presented. Multivariate statistical models of incident frequency and duration are then presented and discussed in detail. Finally, a review of the data collection and analysis process is undertaken and suggestions for further research are made.

CHAPTER 2

RESEARCH APPROACH: DESCRIPTION OF THE STUDY AREA AND DATA COLLECTION PROCEDURES

DESCRIPTION OF THE STUDY AREA

The study area included 20 miles of urban freeway, as shown in Figure 2.1. There were two main sections: Interstate 5 and State Route 520. Both sections vary considerably in average congestion and in their geometrics. For example, SR 520 has no shoulders and includes a floating bridge. SR 520 has relatively short sight distances at the floating bridge "highrises," which are near some of the major merging sections used for the p.m. peak hour commute. Moreover, SR 520 and Interstate 90 are the only east/west links between Seattle and her eastern suburbs. As a consequence, east/west route diversion is quite limited. Further limiting diversion possibilities are the burgeoning suburbs to the north of the lake and major employment centers to the south of the lake. Should an accident or disablement occur on SR 520, in the absence of shoulders, capacity is immediately reduced. Although courtesy tow trucks are now stationed at the bridge at rush hours, these trucks must maneuver through stopped traffic to clear the incident. Most Seattle commuters find disabled vehicles on SR 520 to be a source of intense frustration, but they manage to do what they can to assist the driver.

In contrast, I-5 has 11- or 12-foot lanes and a shoulder for most of the portions lying in the study area. Room is available to pull off of the road in most locations, should there be an incident. The factors affecting accident frequency and duration are quite different, but the impact of a lane blocking accident or disablement is still very high, since the highway is now near capacity and an incident contributes an additional weaving section in an area already noted for weaving problems.

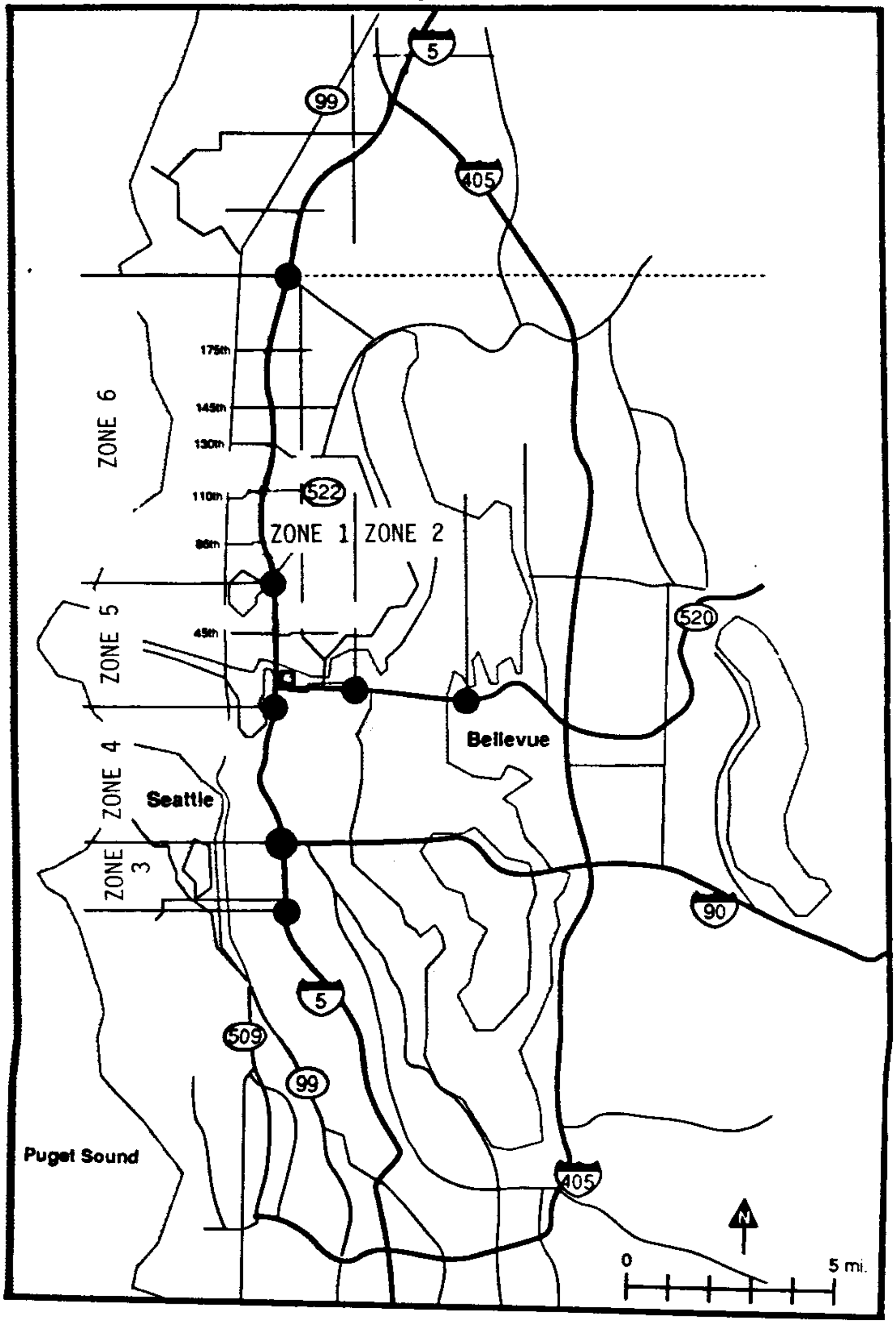

Figure 2.1 Incident Location Zones

To account for the differing geometrics within the defined study area, since they may affect both the frequency and duration of incidents, the area was subdivided into six roughly homogeneous zones, as illustrated in Figure 2.2. Zones 1 and 2 were on SR 520, and Zones 3 through 6 were on I-5. These zones formed the basis for the incident data collection and analysis.

DATA REQUIREMENTS

The data needed for this study had to be suitable for developing statistical models to estimate the daily number (frequency) of incidents (both vehicular accidents and disablements), as well as the duration of specific incidents (accidents only, for reasons to be discussed later). Data suitable for modeling daily frequency could be any data unique for a particular day, or perhaps for a sequence of days. For example, average weather, day of week, month of year, or occurrence of special events would have been appropriate. However, time of day, vehicle type, duration, hourly volumes, and location geometrics would not have been useful in estimating a daily frequency model, since they either do not vary from day to day or are conditional on the occurrence of an incident. For the duration analysis, the researchers attempted to model incident time duration in minutes. Suitable explanatory data had to have some relationship with the duration of the incident duration, (i.e., accident severity, time of day), and could not be a response to the accident characteristics, (i.e., number of tow trucks used). If the latter had been the case, the statistical integrity of the models would have been violated (see Chapter 5).

Frequency Data

Incidents per day were used as the basis for two modeling efforts: (1) the accidents per day model and (2) the disablements per day model. Explanatory variables were selected to describe a particular event or circumstance that was specific to each day. The explanatory variables may have been discrete (i.e.,

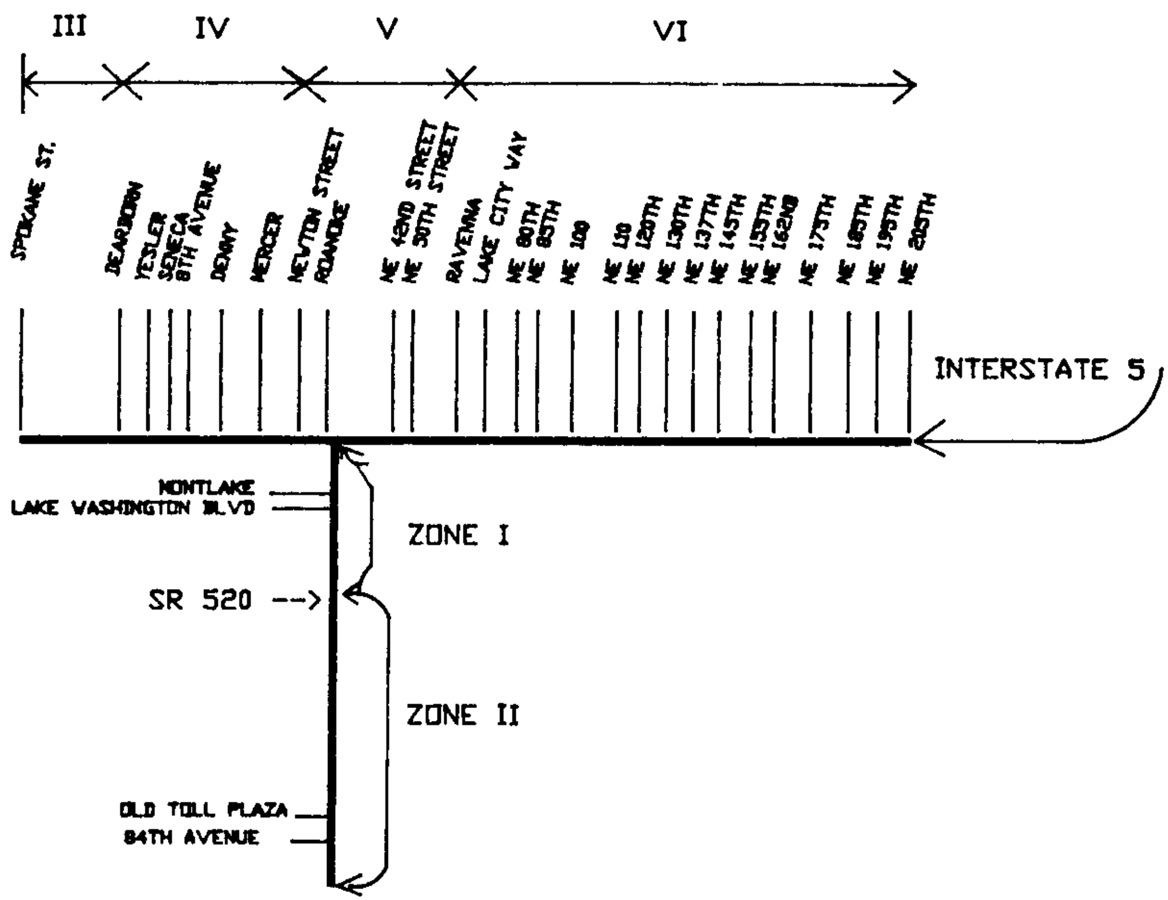

Figure 2.2 Study Area Zones

indicating a particular class such as wet days or days on which a football game took place), or continuous. The continuous variables were given by a number that described the magnitude of the effect of an event, such as the traffic volume on a particular day.

Duration Data

The duration analysis sought to model the duration of an incident (in minutes). The time an officer left the scene of an accident was used as the "end" of an incident. The duration was defined as the difference between when the officer was called and when he/she radioed back that he/she was leaving the scene. Time was a continuous variable that assumed positive values. The explanatory variables might have been either discrete or continuous, as described above.

AVAILABLE DATA SOURCES

A number of potential data sources were identified. Some could provide extensive and detailed information on traffic, weather, and special events; the difficulty was to select a source or sources that provided an adequate level of detail and was consistent over time. Potential sources for incident data (accidents and disablements) were identified. Table 2.1 lists the data sources that were investigated for use in this study.

Source for Duration Information

Recall that the variables to be modeled were duration of accidents and number of accidents and disablements per day. The most consistent and comprehensive source for information on accidents per day was the WSDOT Accident Report Data Base. Other sources (e.g., WSP Accident Dispatch Logs), under-reported the number of accidents. Unfortunately, the Accident Report Data Base included only the start time as recollected by involved parties some time after the accident (i.e., when the report was filled out). The State Patrol Dispatch Logs

TABLE 2.1
DATA SOURCES INVESTIGATED FOR THE
INCIDENT MANAGEMENT PROJECT

Washington State Department of Transportation

Accident Report Data Base
Ramp and Roadway Volumes: Summary Report
State Route Log: Planning Report
TSMC Video Log
Loop Detector 15 minute counts

Washington State Patrol

Accident Dispatch Logs
Incident Dispatch Logs

National Oceanic and Atmospheric Administration

Local Climatological Data: Monthly Summary

Local Professional and College Sports Teams

Husky Football Schedules
Seahawks Football Schedules
Sonics Basketball Schedules
Mariners Baseball Schedules

---

were therefore much more accurate and consistent sources for the duration of accidents, even though they understated the lengths of some accidents because of the unreported time that transpired between accident occurrence and contact with the Patrol for assistance. A key disadvantage was that the dispatch data were not as comprehensive a source for possible explanatory variables (as was the WSDOT Accident Report Data Base), and there was no standard notation for dispatchers' use when logging accidents. However, together the dispatch and accident report data bases made an ample data set for an accident analysis. The duration time data (and some other variables) were collected at the Bellevue Office of the Washington State Patrol Dispatch Center and entered into a Dispatch Data Base and later were combined with the WSDOT Accident Report Data Base. Accidents per day, used in the accident frequency models, were derived solely from the WSDOT accident reports.

A two-year time frame was used for this analysis. This two-year study period was chosen to factor out some annual and seasonal effects. One of the two years was a leap year, so the total number of days studied was 731. Recall that two types of freeway incidents were defined: (1) vehicular accidents, and (2) vehicular disablements. As discussed above, dispatch data from the Washington State Patrol were used to determine the duration of accidents. Since the Patrol periodically eliminates unnecessary documentation, accident data were collected from the earliest month available (April 1986), and collection ended two years later at the end of March 1988. The data for the frequency of vehicular disablements are extremely extensive, and the Patrol only keeps one year's back records. The data used for disablement estimation were therefore only for one year (April 1987 to March 1988). Also, since information relating to the duration of vehicle disablements was extremely poor, only data on the frequency of disablements could be considered.

Explanatory Variables

The complete data base for the frequency analysis was derived from the WSDOT Accident Data Base, WSDOT Ramp and Roadway Volumes, and the WSP Incident Logs, and from game schedules of the major sports teams in the Seattle area. Explanatory variables were taken from the Accident Report Data Base, the Incident Report Data Base, Sporting Event Calendars, and Ramp and Roadway Volumes.

Data Not Used

Loop count data were considered for use in assessing the impacts of specific incidents along the study area, but the data base proved to be too large to manipulate without a large investment in computer resources and research time.

Further, many loops were not reliable within the study area and were scheduled for replacement.

The Video Monitoring system located at the Traffic Systems Management Center (TSMC) was made available for use in this study. Several deficiencies in the usefulness of this system for data collection for our purposes were identified. First, the cameras had a limited sight distance. The researchers had hoped that queue length data could be collected with this system, but queues were soon out of sight of the cameras. Second, the number of incidents actually detected was small in comparison to the accident report and the dispatch data bases. Since only a few locations could be monitored at a time, no one working at the Center could recall actually seeing the start of an incident, and they usually "found" it after a few minutes - at that time they would contact the state patrol if appropriate. This meant that the dispatch data base would include the TSMC data, and therefore would be redundant.

Weather data collected by the National Oceanic and Atmospheric Administration (NOAA) has been used by other researchers to document weather and road surface conditions in their study area. However, this study area was much smaller and included only one NOAA station. Preliminary analysis of rainfall and accident duration indicated that the NOAA data were too general to be of value. Average weather and road surface conditions derived from the Accident Report data base were found to be much better predictors for both the frequency models and for the duration models. Their usefulness is probably due to the relatively small size of the study area.

DATA MATCHING

Dispatch reports were matched with accident reports on the basis of the date, time of day, hour and minute, and milepost. When a match was questionable, the number of vehicles and travel directions were verified. Each matched record was assigned a unique "match flag" number, and data from the dispatch data base were transferred to the accident report data base using this flag to insert additional data for the matched records. Some 2,156 of the 5,637 accident reports were matched with dispatch records in this way. This was a much higher proportion of matched accident records than found in other similar studies (Guilano, 1988 for example), in which much smaller stratified samples of the dispatch data were randomly selected until minimum sample sizes were obtained. One concern was how much difference between dispatch and report start times could be allowed when records were matched. The longest single difference was 95 minutes; for very long differences all available accident characteristics had to match between dispatch and accident data bases. The average difference between accident report and dispatch times was 5.07 minutes, with a standard deviation of 9.64 minutes. The possibility existed that longer accidents might be over-represented, since they had a greater amount of detail in both the accident and dispatch files, making a match easier to identify. Another difficulty was the use of different location data in each data base. The accident reports used mileposts; the dispatchers used the nearest cross-street, and sometimes approximate distance from a cross-street. To resolve this, all streets crossing Interstate 5 and SR 520 within the study area were assigned a milepost in the dispatch files on the basis of the WSDOT induction loop counter location guide. This guide included both the cross-streets and the mileposts.

DETAILS OF DATA COLLECTION

Two kinds of data collection were required. The accident report data were available on diskette from the DOT as ASCII files in a continuous format, 132-character record length data set (see Table 2.2 for a description of the actual data available from this source). To use this data, the researchers had to write a program to reformat the fields into columns separated by spaces, so that other applications programs could use it easily. This was done with a short Fortran

TABLE 2.2

VARIABLES AVAILABLE FROM ACCIDENT REPORTS

Year	Year of accident
Month	Month of accident
Day	Day of Accident
Day of Week	Indicator for day of week Monday to Sunday
Hour	Hour accident report started
Minute	Minute accident report started
Sign Route	State route highway designation number
SR Milepost	State route mile post
Accident Sev	Accident severity index; property only, injury accident, or fatality
N. Injured	Number of persons injured in the accident
N. Fatal	Number of persons killed in the accident
Light	Indicator for illumination level at accident site: daylight, dawn, dusk, dark (with and without street lights, and other)
Collision Type	Code for various possible collision types including pedestrian/vehicle, vehicle/vehicle, parked vehicle and others kinds
Object Struck	Kind of object struck, if any (e.g. light standard)
M. Sev.Inj.	The most severe injury caused by the accident (no injury, fatal, disabling, non-disabling, possible, unknown)
N. Veh.	Number of vehicles involved in the accident
P.Dam.$	Property Damage measured in dollars
R. Char.	Roadway character - grades and curves
L. Char.	Location character - codes for various intersections, under and over passes and other facilities

TABLE 2.2

VARIABLES AVAILABLE FROM ACCIDENT REPORTS (Continued)

R. Sur.	Road surface character: not stated, dry, wet, snow, ice, other
Weather	Weather at the accident site:clear/cloudy, rain, snow, fog, or other
Res. Prox.	Residence proximity of involved drivers: within 15 miles, elsewhere in state, or out of state
Sobriety	Sobriety of the drivers in the accident: 7 codes for had been drinking — ability impaired to had not been drinking
A. Sev.	Alcohol severity: drunkest driver involved in accident
Con.Circ	24 codes indicating different possible RCW violations or indicating no violation
D.V.Act	Driver Vehicle Action: codes indicating evasive or non evasive actions taken by the involved drivers
Veh. typ.	Vehicle type: vehicle type code
Age	Age for each of the involved drivers
Haz. Mat.	Kind of hazardous material involved, if any
Fuel	Fuel Spill (yes/no)
Fire	Fire Resulted (yes/no)

program. The accident dispatch and disablement data were collected manually at the District One State Patrol office in Bellevue. Data for special events were obtained from the business offices of local sports teams in the form of game schedules. The accident data, disablement data, and game schedules were entered into separate data bases to keep separate different sets of input data. Transactions between data bases were used to share information between databases. A third kind of data was generated by one of two methods: either counts taken from the original data or estimates created in a modeling process. Count data included variables such as accidents per day and accidents per day by zone; estimated data were the estimated number of disablements expected each day. The latter were generated by operationalizing a model of disablements per day, as will be discussed in Chapter 5. The researchers had anticipated that this variable would assist in indicating the level of activity during an accident and thus affect the duration of accidents. Altogether, three databases were used to generate the data for the project. A total of 5,637 accident records, 4,648 accident dispatch records, and 8,429 disablement records were collected. To collect one month's worth of accident data required about five person-hours; approximately eight person-hours were needed to collect one month of disablement data.

DISPATCH DATA COLLECTION PROCEDURES

Table 2.3 is a sample of the data available from the State Patrol's accident dispatch and disablement logs. Initially, lap-top computers were considered for data collection, but practically the machines proved to be more of a hindrance than a help.

The following is a description of the dispatch data collection task. The Washington State Patrol now has a computerized dispatch system, and subsequent researchers will have access to electronic data sources much like the accident reports available from the Washington State DOT. However, for those who do not

TABLE 2.3

DATA COLLECTED FROM DISPATCH LOGS

Year	Year of accident
Date	Month and day of the accident
DOW	Day of week:Monday to Sunday
TOD	Time of day: one of eight 3 hour time slots
Rcvd	Dispatcher received call for assistance
Enroute	Time Trooper was en route to accident scene
Atscene	Time Trooper arrived at the accident scene
Roadcl	Time road was cleared, if it had been blocked
Troopcl	Time Trooper cleared the scene of the accident
Tow	Number of tow trucks called to accident site
Amb	Number of ambulances called to the accident site
Other	Number of other emergency vehicles called to the accident site (e.g. Fire Department)
Exp	Indicates that accident occurred on the express lanes
Loc. C	Location Code: code for the cross streets on I-5 and SR 520 within our study area
Dir	Direction of travel
Lane	Lanes(s) involved
N.Veh	Number of vehicles involved
N.L.Block	Number of lanes blocked
Inj	Number of injuries
Ftl	Fatality accident (yes/no)
T/B	Truck or bus involved in the accident

have such a system, this description will provide an idea of the level of effort required to collect dispatch data and match them to accident reports.

Guidelines for the Accident Dispatch Collection Task

1. Make an appointment with the appropriate officer.
2. A half day must be taken to explore and identify the useful data in the dispatch logs.
3. Develop a data collection form on the basis of the exploratory data collection effort.
4. During the first session, the most important thing to learn is how that particular office abbreviates or codes its dispatch cards. Except for the preprinted areas and the time stamp, all information is recorded by hand. The abbreviations are often cryptic, since individual dispatchers have different levels of experience, and since the dispatcher's workload and dispatch priorities change with time.
5. Revise the data collection form if appropriate.
6. Make sure the necessary information is available before using the incident. For this study, the project team needed at least the highway, the cross street, the dispatch time, and the trooper clear time.
7. Data must be transferred to a data base to avoid a large backlog.
8. It was helpful to speak with the dispatch personnel to clear up questions on an informal basis and maintain good communications.

Guidelines for the Disablement Dispatch Task

Collection of the disablement data meant that a great many more data cards had to be examined than during the collection of accident data. The disablement files were approximately four times longer than the accident files. Again, cards were selected by zone and cross streets were checked. These data were to be used only in a frequency model, so the number of disablements in the study area for each day (classified by disablement type) were collected. Abbreviations and data card format were much the same as those of the accident dispatch cards.

CHAPTER 3

FINDINGS: INCIDENT FREQUENCY AND DURATION ANALYSIS TECHNIQUES

Incident data for disablements and accidents are required by several levels of government for highway safety and management. The State Department of Transportation needs to know where improvements in geometric design or when, where, and which traffic management options are most appropriate for incident management. The State Patrol needs to know when and where to allocate resources to appropriate "beats" and the potentially effective incident site management techniques. Incident frequency and duration data analysis is used to answer important questions regarding the characteristics of incidents. Below is a discussion of common data analysis techniques, a review of some recent papers using these techniques, and a summary.

ANALYSIS TECHNIQUES

Analysis techniques include univariate analysis, cross tabulation analysis, indexing schemes, several different regression analysis techniques, and duration analysis. Each is appropriate for analysis of incident frequency, duration, or either, depending on the level of detail required, funding available, or model complexity needed to adequately describe incident frequency or duration for a particular purpose. Frequency analysis may be done with any of the above techniques. Duration analysis may use most of these techniques, but some regression specifications can be inappropriate.

Univariate Analysis

Univariate analysis focuses on a single variable. For frequency analysis, the means, standard deviations, and number of observations for a particular variable are noted and discussed with respect to its probable significance without reference to other accident characteristics. For example, the average number of accidents or disablements for each day of the week may be used to allocate State Patrol troopers. The standard deviation and the number of observations give decision makers an idea of how reliable that measure may be for that purpose.

Cross Tabulated Analysis

This is much like univariate analysis except that a class variable is used to stratify the data. Averages, standard deviations, and frequencies can found for each class. An example is to separate examination of accidents with different severity levels and comparison of another variable such as number of "vehicles per accident" from class to class to determine whether there is a significant difference among the means of different classes.

Indexing Schemes

Many state DOTs use accident indexing schemes. These use the DOT's accident report data base to compile an accident index for each milepost within the agency's jurisdiction. The index is based on the number of accidents near the milepost. Accident severity, frequency,and average daily traffic are a few of the variables used to determine the index. Once an index has reached a threshold level it is reported as a high accident area and appropriate action is taken by the DOT. Usually geometric or pavement characteristics are the problem; sometimes operational deficiencies need correction.

Regression Models

Linear regression is the most widely used model for modeling the frequency of incidents. This model assumes normal distribution of incident frequency. The dependent variable is the number of accidents per unit time, and the independent variables that have been modeled include vehicle miles traveled, weather conditions, gross national product, and many others. The dependent variable's time frame is the control for the independent variables. For example, it would make little sense to model accidents per day using gross national product as an independent variable; it makes sense, however, if the dependent variable is accidents per year.

Other distributions have been assumed for regression analysis of accident data. Log-normal distributions have been used in the analysis of the time duration between accidents. The Poisson model assumes an exponential distribution of the time between accidents and has been used for accident frequency estimation.

Survival Models

Survival models are a recent development in the area of accident analysis. These models describe the conditional probability of an event that has continued a given length of time to end soon. These models have been used in frequency and duration analysis of truck accidents and general traffic accidents.

RECENT INCIDENT FREQUENCY AND DURATION PAPERS

The following overviews some of the more important papers in the area of incident frequency and duration modeling.

Frequency Analysis

Two recent papers have described models for incident frequency analysis that have not been commonly discussed in the accident analysis literature. "Modeling the Relationship of Accidents to Miles Traveled" describes a Poisson regression model that avoids some of the theoretical problems encountered using linear regression (Jovanis and Chang, 1987). The data are from 157 miles of the Indiana Toll Road. An exact count of vehicles classified by passenger car or truck was available for each toll paid. The advantages to using the Poisson regression are in the discreteness of the response variable (i.e., no fractional accidents are allowed), the lack of negative responses, the possibility of a zero response, and reduced homoscedacisity. The model assumes a Poisson distribution.

"Disaggregate Model of Highway Accident Occurrence using Survival Theory" uses survival analysis to determine the factors important in accident frequency (Jovanis and Chang, forthcoming in *Accident Analysis and Prevention*). Submodels include an Injury and Fatality model, a Property Damage Only model, and two collision type models: Single Vehicle and Multiple Vehicle models. Significant independent variables affecting truck accidents were Winter, Night, Age, Experience of the Driver, Weight of Vehicle, Hours of Driving, and Hours Since Last Shift. The dependent variable was time between accidents.

Duration Analysis

"An Analysis of the Severity and Incident Duration of Truck Involved Freeway Accidents" provides both a cross classification analysis for frequency analysis and a log-normal regression analysis for investigation of the time duration of truck accidents (Golob, Recker, and Leonard, 1986). Data were obtained from the California State DOT Accident report data base and from the California State Patrol dispatch records. Accident frequency by collision type, severity and other classifications were provided. The regression analysis suggests that a log-normal distribution is superior for regression and provides a detailed empirical analysis of durations classified by accident type as evidence.

"Incident Characteristics Frequency, and Duration on a High Volume Urban Freeway" is the only paper in the literature that discusses both accident and disablement duration (Giuliano, 1988). Giuliano used Analysis of Variance by Regression to estimate models of duration as a function of various independent variables. Two models were estimated: one for accident duration and one for disablement duration. Significant variables for disablement duration included incident type, lanes closed, time of day, and lane type. The accident model included accident type, time of day, and truck involvement. A log-normal form for the model was tested and loosely justified with a Kolmogorov-Smirnov statistical test.

SUMMARY

There are several uses for incident frequency and duration analysis. Depending on the level of detail needed for a particular study, one of several techniques may be employed for data analysis: univariate, cross-tabulation, indexing, regression modeling, and survival analysis. Recent studies have used Poisson regression and survival analysis techniques for accident frequency analysis; disablement and accident duration have been modeled using Analysis of Variance by Regression and log-normal regression. The methods that will be described in this paper build upon those used in previous work, with appropriate advances and improvements following recent developments in statistical techniques.

CHAPTER 4

FINDINGS: OVERVIEW OF THE INCIDENT DATA

As seen in Chapter 2, many possible factors were gathered from the dispatch and disablement logs that could affect the frequency and durations of incidents. This chapter is a presentation of the factors that appeared to be significant determinants of incident frequency and duration, and those factors generally considered important in the traffic accident analysis and prevention literature. A univariate analysis examining the common statistics of each of the significant factors is presented with the aid of figures and tables. Table 4.1 shows each of the factors to be presented, categorized by frequency (broken down to month, day of week and time of day); and characteristics (broken down to driver, vehicle type, environment, and accident). Reasons for not presenting the remaining collected factors stem from either small sample size geometric features unique to a specific location, variables too similar to each other, or overly complex classification groupings.

TABLE 4.1

ACCIDENT AND INCIDENT FACTORS AND VARIABLES

Frequency	Characteristics
Month	Driver
Day of Week	Age
Time of Day	Sobriety
	Residency
	Vehicle Type
	Environmental
	Road Surface
	Weather
	Location and Frequency
	Accident
	Injuries
	Property Damage
	Frequency and Duration

FREQUENCY

Cyclical trends are important factors in incident analysis. A few of the most important are examined below.

Month

Figure 4.1 shows the number of disablements per day by month. Disablements included both blocking and non-blocking disabled vehicles and blocking hazards on the roadway. The counts were fairly consistent, although the second half of the year, summer and fall, was slightly higher than the first half. Since the second half increase was not very large, and given the characteristics of the region, the increase may have been due to simple factors of activities carried on during these months. Examples of these factors include higher temperatures causing older cars and recreational vehicles not used often in the winter months to overheat, or portions of poorly secured loads falling on the freeway.

Figure 4.2 shows the number of accidents per day by month. The slight decrease of accidents in the summer most likely came from the factors of better weather and less traffic due to vacations and school breaks. The increase of accidents in the spring probably resulted from spring rain and storms, while the increase of accidents in the fall resulted from an increase in overall traffic as people returned to their normal schedules.

Day of Week

Figures 4.3 and 4.4 show what was expected for disablements and accidents on a weekly basis. Rates increased from Monday through Friday, then decreased significantly on Saturday and decreased even more on Sunday. This typical pattern was caused by both the volume of traffic using the freeway as a commuter route to work over the normal Monday through Friday work week and the anxiety that also seemed to build over the work week in expectation of the coming weekend. The accidents plot was more clearly defined than the disablements plot, which was

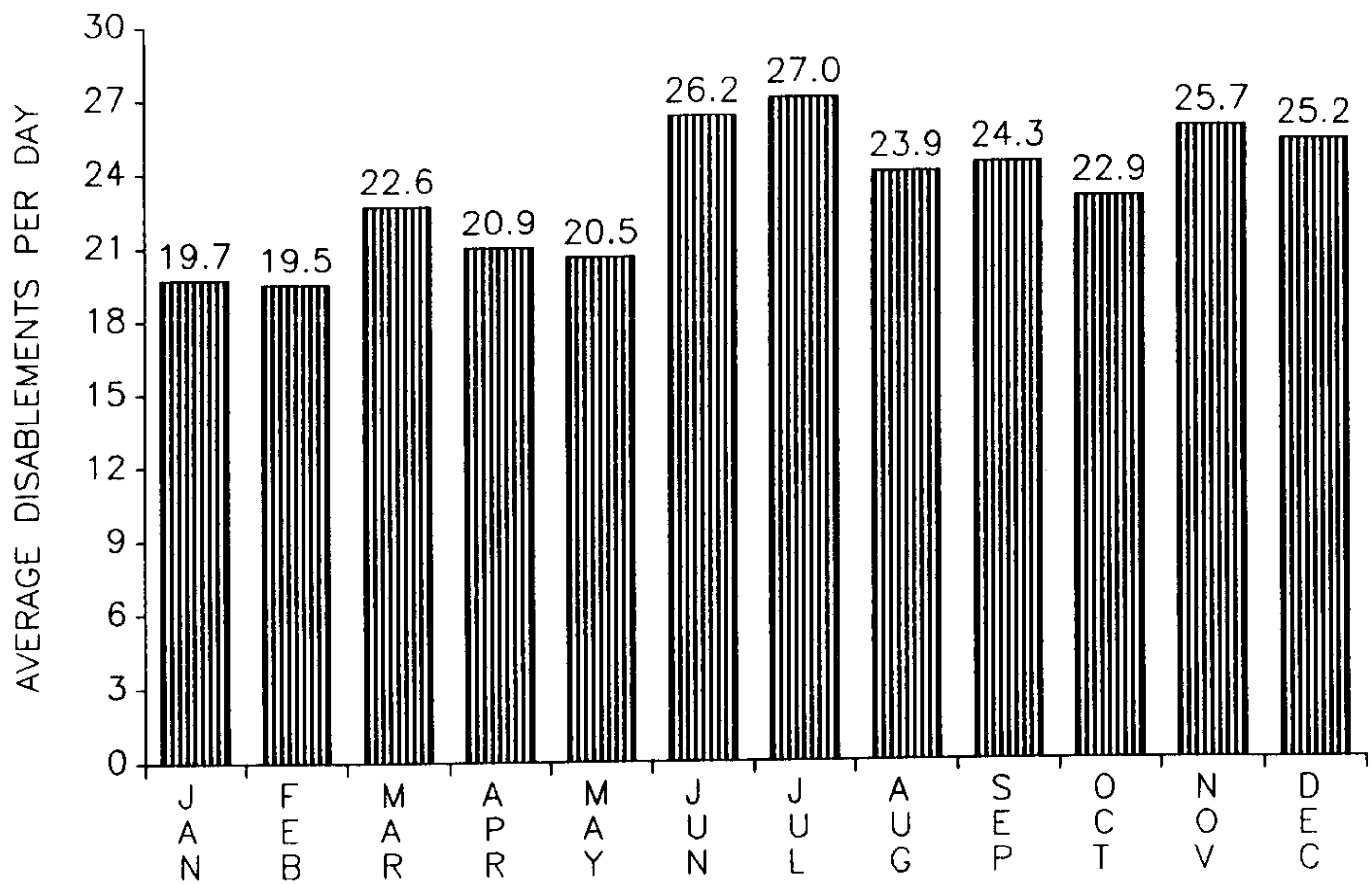

Figure 4.1 Disablements Per Day by Month.

SOURCE: WSP ACCIDENT REPORTS
A One Year sample

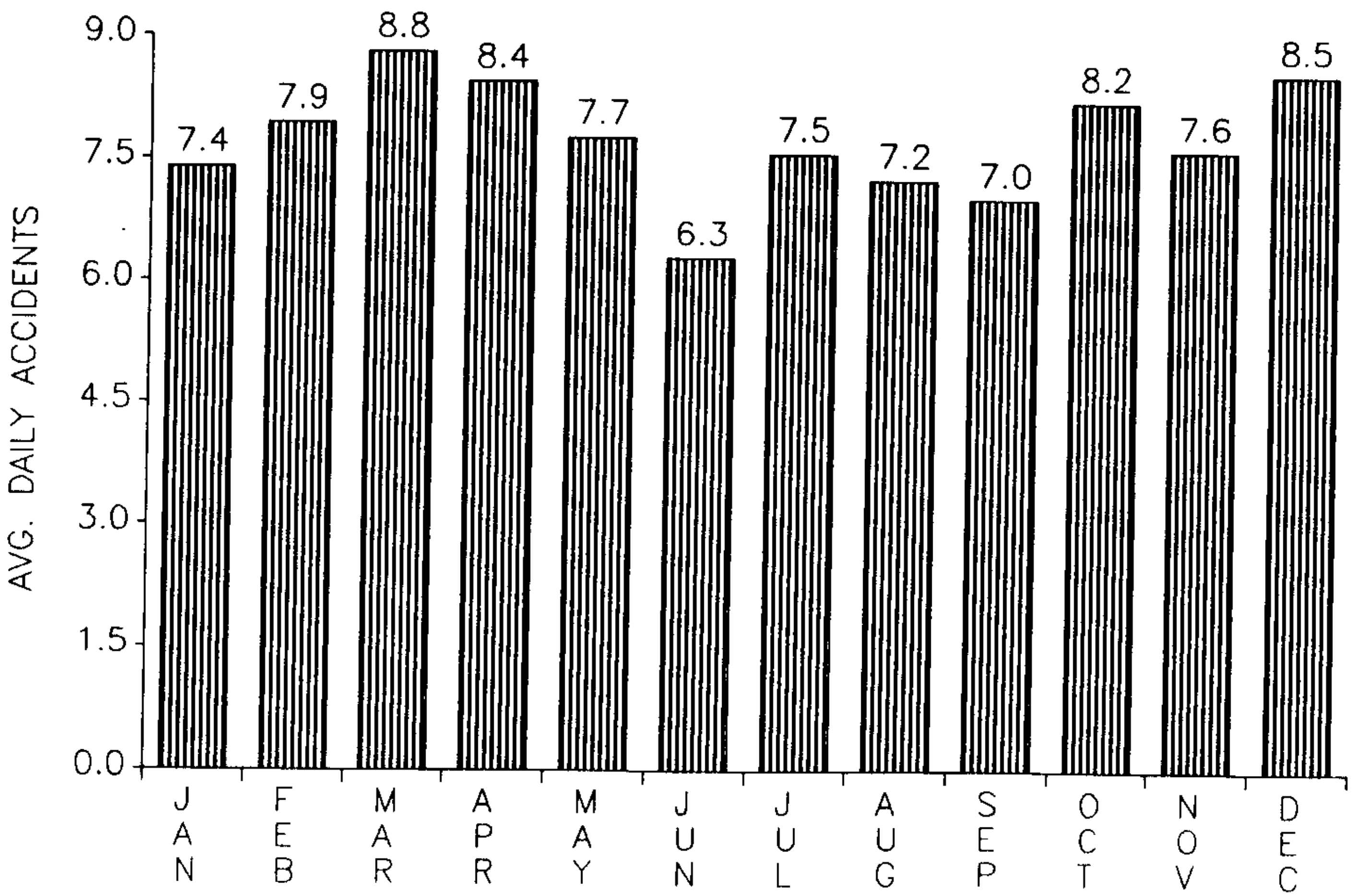

Figure 4.2 Accidents Per Day by Month.

SOURCE: WSP ACCIDENT REPORTS
A Two Year Sample

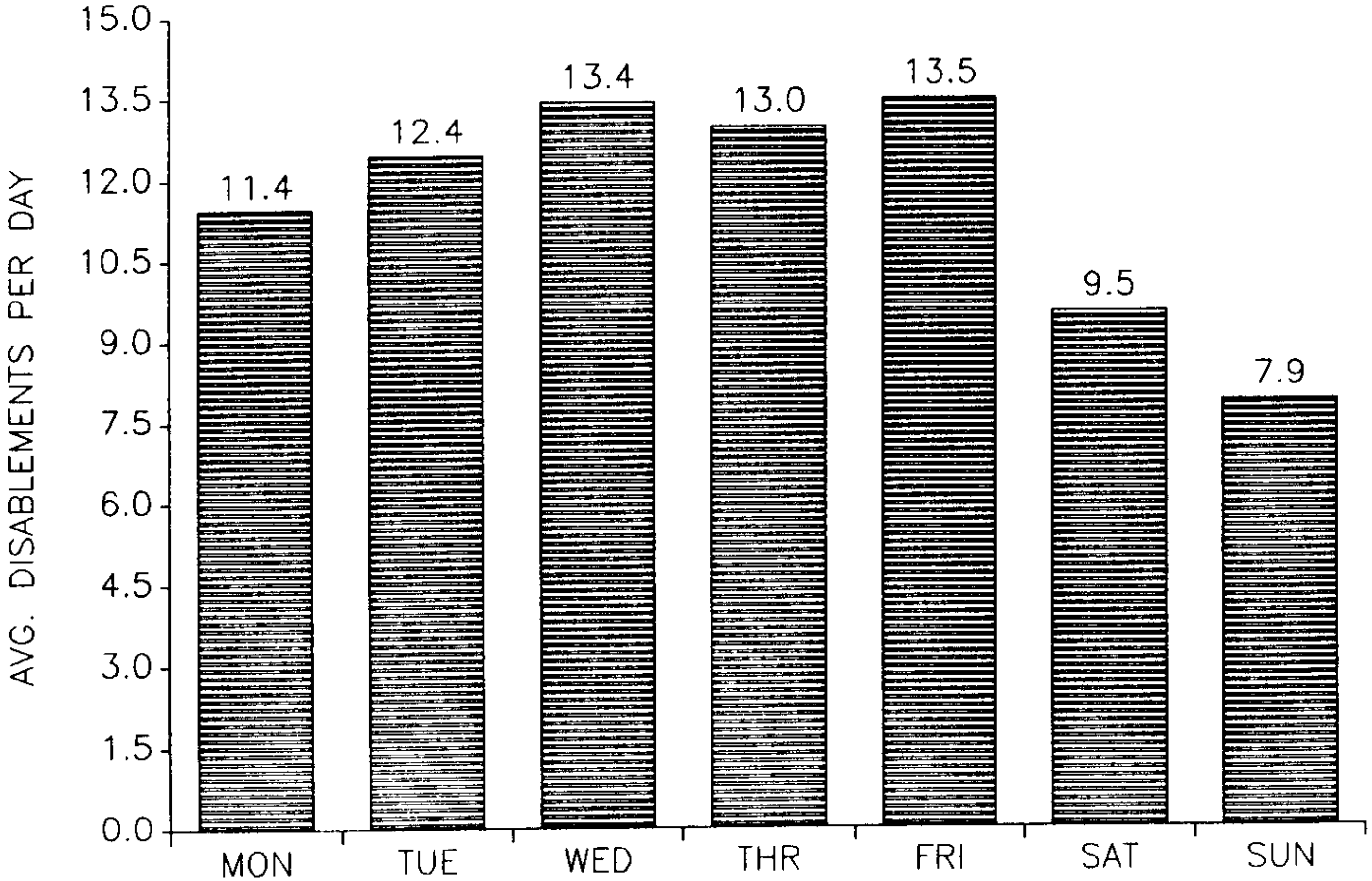

Figure 4.3 Disablements by Day of Week.

SOURCE: WSP DISPATCH DATA

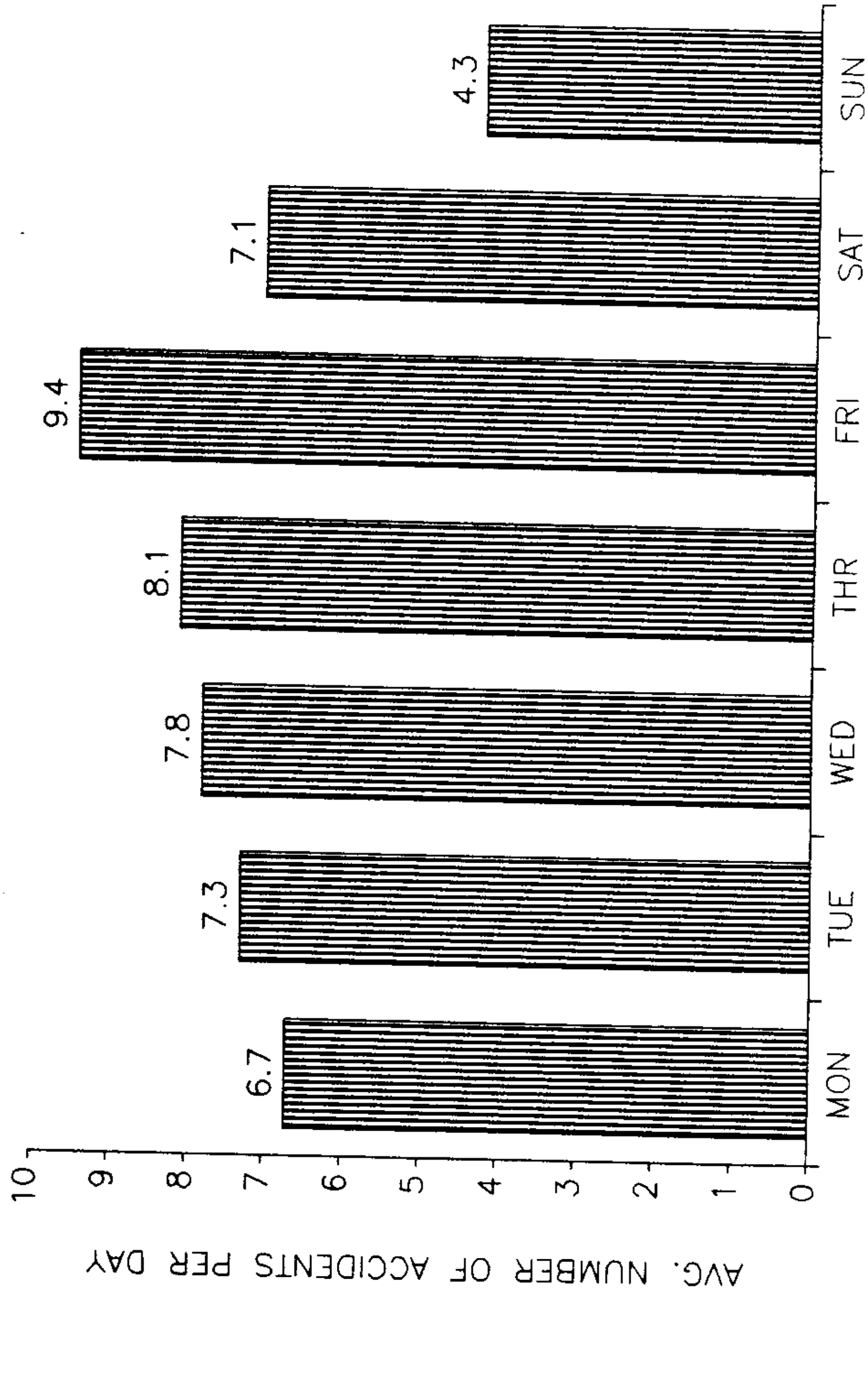

Figure 4.4 Accidents by Day of Week.

SOURCE: WSP ACCIDENT REPORTS

probably due to the greater amount of human error involved in accidents than in disablements.

Without having the actual volume counts, one cannot deduce that driving on the freeways is safer on the weekend. The lower number of accidents on Saturday and Sunday may have been caused by a smaller volume of traffic and may actually have been a higher percentage of the total volume of weekend traffic. One reason for assuming that driving may be safer during the week is that the drivers are commuters who are generally described as more aware of and more familiar with the routes they drive than the weekend driver, who may be driving on a particular route for the first time.

Figure 4.5 shows four interesting facts. One, there were almost as many blocking disablements as non-blocking disablements. Two, there were very few hazards. Three, the total disablements (blocking, non-blocking, and hazards) made up a significantly higher number of incidents than did accidents. Four, the total number of incidents (accidents and disablements) averaged about 20 per day for the study area.

Time of Day

Figure 4.6 gives a brief analysis of the reported start time for all accidents. As would be expected, the plot of number of accidents showed a peak in the morning and afternoon rush hours, 9 to 12 a.m. and 3 to 6 p.m., respectively. The early afternoon hours though, from 12 to 3 p.m., were also very high, creating a broad afternoon peak. This broad peak implies a heavy continual flow of traffic through the study area and a system that is approaching maximum capacity.

CHARACTERISTICS

Driver, vehicle, roadway, and environmental factors often can trigger an incident. An examination of a few of these follows.

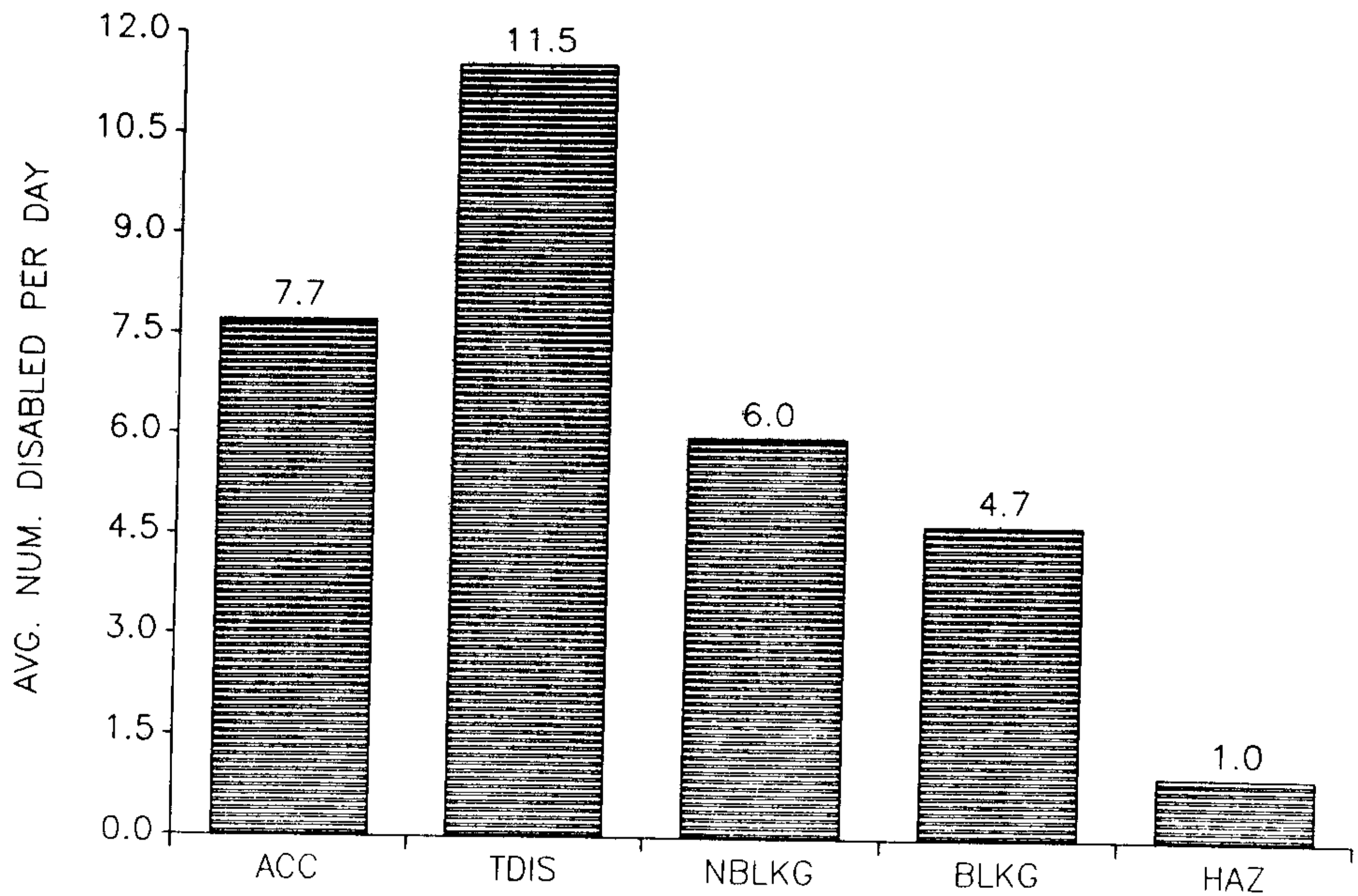

Figure 4.5 Disablements Per Day by Type.

SOURCE: WSP INCIDENT DISPATCH LOGS

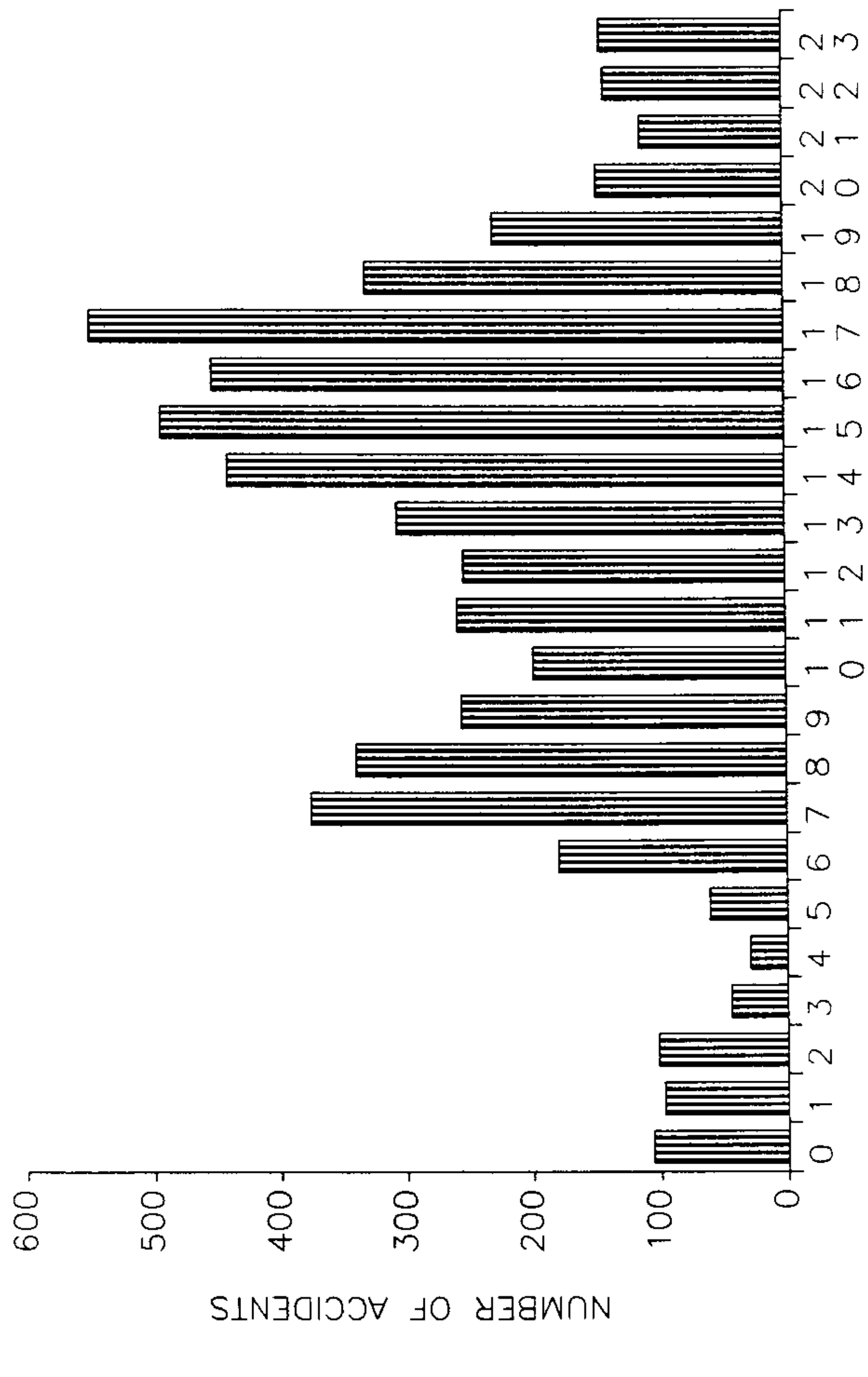

Figure 4.6 Accident Start Time (Hour).

SOURCE: WSDOT ACCIDENT REPORT DATA BASE

Driver Age

Tables 4.2, 4.3, and 4.4 describe different aspects of the driver age variables that were collected. Points of interest regarding each table are separately noted, but numbers that reappear in different tables are as follows. There were 9,958 drivers involved in accidents in the study area over two years. These 9,958 drivers were involved in 5,637 accidents, as reported by the Washington State Patrol. Of the 5,637 accident reports filed, 2,156 accident reports were matched to the dispatch logs used in the analysis of the duration.

Table 4.2 is derived from the WSDOT Accident Data Base, which gives the drivers' ages for up to three drivers involved in an accident. These data were classified by number of vehicles involved in the accident so that the duration statistics were not skewed too much by the severity of the accidents. From the table, note that the average age for one vehicle accidents was 30.61 years, noticeably less than the average for all accidents of 34.55 years. Also note that the average duration for a one vehicle accident was slightly longer than for a two vehicle accident. Three vehicle accidents then resumed to the expected pattern of longer duration.

An interesting note from Table 4.3, which breaks the drivers' ages into ten-year increments, is that just over 62 percent of all accidents involved a driver less than 36 years old. Contrasting older drivers against the rest of the drivers as Table 4.4 revealed a substantially higher average accident duration for older drivers, 60.68 and 54.49 minutes for older drivers and the rest of the drivers, respectively. The standard deviation of the duration for these accidents was also worth noting. A 20-minute difference between the two suggested that a greater variance in accidents occur when older drivers are involved.

Driver Sobriety

Drinking and driving is considered a major problem in the United States. The statistics of Table 4.5 show that drinking and driving was also a very critical

TABLE 4.2

DRIVER AGE CHARACTERISTICS
for All Drivers vs. Number of Vehicles Involved in the Accident

	1 Veh.	2 Veh.	3+ Veh.	All Accidents
Average Age	30.61	35.15	34.43	34.55
Std. Dev.	11.77	13.87	13.23	13.56
Number Obs.	802	5959	3193	9958
Missing Obs.	73	1140	440	1643
Total Records	875	3551	1211	5637
Avg. Duration	53.67	52.47	61.61	54.97
Std. Dev. Dur.	32.76	31.05	27.08	30.63
Number Obs.	395	1139	584	2156
Missing Obs.	33	5	0	3

TABLE 4.3

DRIVER AGE CHARACTERISTICS
by Age Breakdown

Age	Count	Mean	Std. Dev.	% All Indiv.
15-25	3007	21.36	2.57	30.2
26-35	3188	30.29	2.85	32.0
36-45	1883	40.00	2.81	18.9
46-55	918	49.95	2.82	9.2
56-65	591	60.05	2.90	5.9
65+	371	71.68	4.79	3.7
Total	**9958**	**34.55**	**13.56**	**100.0**

TABLE 4.4

DRIVER AGE CHARACTERISTICS
by Old and Young Drivers

Age	**Count**	%	**Avg. Dur.**	**Std. Dev.**	**Dur. Sample**
15-64	5232	92.8	54.49	28.57	1990
65+	405	7.2	60.68	48.72	166
Total	**5637**	**100.0**			**2156**

TABLE 4.5

DRIVER SOBRIETY

	Count	%
Had been drinking	2060	36.5
Had not been drinking	3577	63.5
Total	5637	100.0
Ability impaired	2039	36.2
Toxic test used to identify		
Yes	1799	88.2
No	240	11.8
Total	2039	100.0
Ability not impaired	3439	61.0
Unknown	159	2.8
Total	5637	100.0
Toxic Test Administered		
Yes	1799	31.9
No	3838	68.1
No	3838	68.1
Total	**5637**	**100.0**

TABLE 4.6

DRIVER CHARACTERISTICS for Residence Proximity and Number of Vehicles vs. Duration

	Count	%	Avg. Dur.	Std. Dev. Dur.	Dur. Sample
1 Vehicle					
Not stated	66	7.5	—	—	32
Within 15 miles	698	79.8	52.56	32.35	336
In state	81	9.3	57.60	34.83	40
Out of state	30	3.4	56.75	27.89	20
Subtotal	875	100	—	—	428
2 Vehicles					
Not stated	10	0.3	—	—	5
Within 15 miles	2684	75.6	50.62	23.66	805
In state	590	16.6	54.96	45.96	217
Out of state	267	7.5	55.78	29.60	117
Subtotal	3551	100.0	—	—	1144
3+ Vehicles					
Not stated	2	0.2	—	—	1
Within 15 miles	832	68.7	60.05	25.31	390
In state	250	20.6	60.37	28.02	123
Out of state	127	10.5	72.23	32.57	70
Subtotal	1211	100.0	—	—	584
All Accidents					
Not stated	78	1.4	—	—	38
Within 15 miles	4214	74.8	53.51	26.61	1531
In state	921	16.3	56.99	39.84	380
Out of state	424	7.5	61.43	31.31	207
Total	5637	100.0	—	—	2156

problem in the study area. Table 4.5 shows that 36.5 percent of the accidents reported involved a driver who had been drinking. The validity of the impaired ability of the driver was based on the fact that 88.2 percent of the drunk drivers were tested using the toxic test. The last statistic states that the toxic test was used in 31.9 percent of the accidents reported, substantiating the immense problem of drinking and driving.

Driver Residency

Three features need to be noted from the analysis of mean duration and accident frequency for drivers' proximity to their homes, cross-tabulated by number of vehicles involved, as seen in Table 4.6. The first is the pattern of average duration, which increased the farther the drivers were from their homes. As above, average duration was longer for a one-vehicle accident than a two-vehicle accident. The second feature ties into the statistic shown earlier, that single car accidents had a younger driver on average. Younger drivers might be assumed to be traveling closer to home. Table 4.6 suggests this might be true, since the highest percentage of accidents within 15 miles of the home involved one-car accidents. Interestingly, in-state but not within 15 miles of home, and out-of-state accidents significantly increased durations for two-vehicle accidents and even more for three-vehicle accidents. The third feature of Table 4.6 is the count of 4,214, or 74.8 percent of the accidents within 15 miles of the home. Although this was a high percentage of drivers presumably familiar with the highway facility and traffic patterns, this statistic conforms very nicely to the established estimate that 75 percent of accidents occur within 25 miles of the home.

Vehicle Type

Table 4.7 analyzes the mean duration and accident frequency for vehicle types, cross-tabulated by the number of vehicles involved. Note that a large percentage of accidents, for any number of vehicles involved, included a car, which

Table 4.7

ACCIDENT CHARACTERISTICS
for Number of Vehicles and Vehicle Type

	Count	%	Avg. Dur.	Std. Dev. Dur.	Dur. Sample
1 Vehicle					
Car	669	76.4	50.47	24.86	327
Truck	166	19.0	65.11	53.01	81
Other	40	4.6	48.55	18.61	20
Subtotal	875	100.0	—	—	428
2 Vehicles					
Car and Car	1635	46.0	50.26	22.42	531
Car and Truck	1328	37.4	57.23	46.99	215
Truck and Truck	188	5.3	58.31	35.77	77
Car/Truck and Other	64	1.8	53.67	26.07	24
Missing Data	336	9.5	—	—	297
Subtotal	3551	100.0	52.24	31.18	1144
3+ Vehicles					
Cars Only	468	38.7	61.25	25.21	257
Trucks Only	16	1.3	69.00	35.51	6
Cars and Trucks	492	40.6	63.62	27.33	254
Other	13	1.1	79.60	49.19	11
Missing Data	222	18.3	—	—	56
Subtotal	1211	100.0	61.61	27.08	584

is consistent with the high percentage of cars observed in the traffic flow. Also worthy to note is that the involvement of a truck results in a noticeable increase in incident duration.

Road Surface

The accident report forms reported four types of road surface conditions: dry, wet, snow, and ice. As Figure 4.7 shows, the overwhelming majority of accidents took place on dry surfaces. Of the 5,637 accidents, 4,033 accidents occurred on dry pavements, 1,544 on wet surfaces, 57 on icy surfaces, and only three on snow surfaces. Therefore, only 1,604 accidents took place when the road slickness could be blamed. This low proportion is reflective of Seattle's mild climate and (contrary to popular image) low proportion of rainy days. To establish the true effect of slick roads, a multivariate analysis was needed, as will be undertaken in Chapter 5.

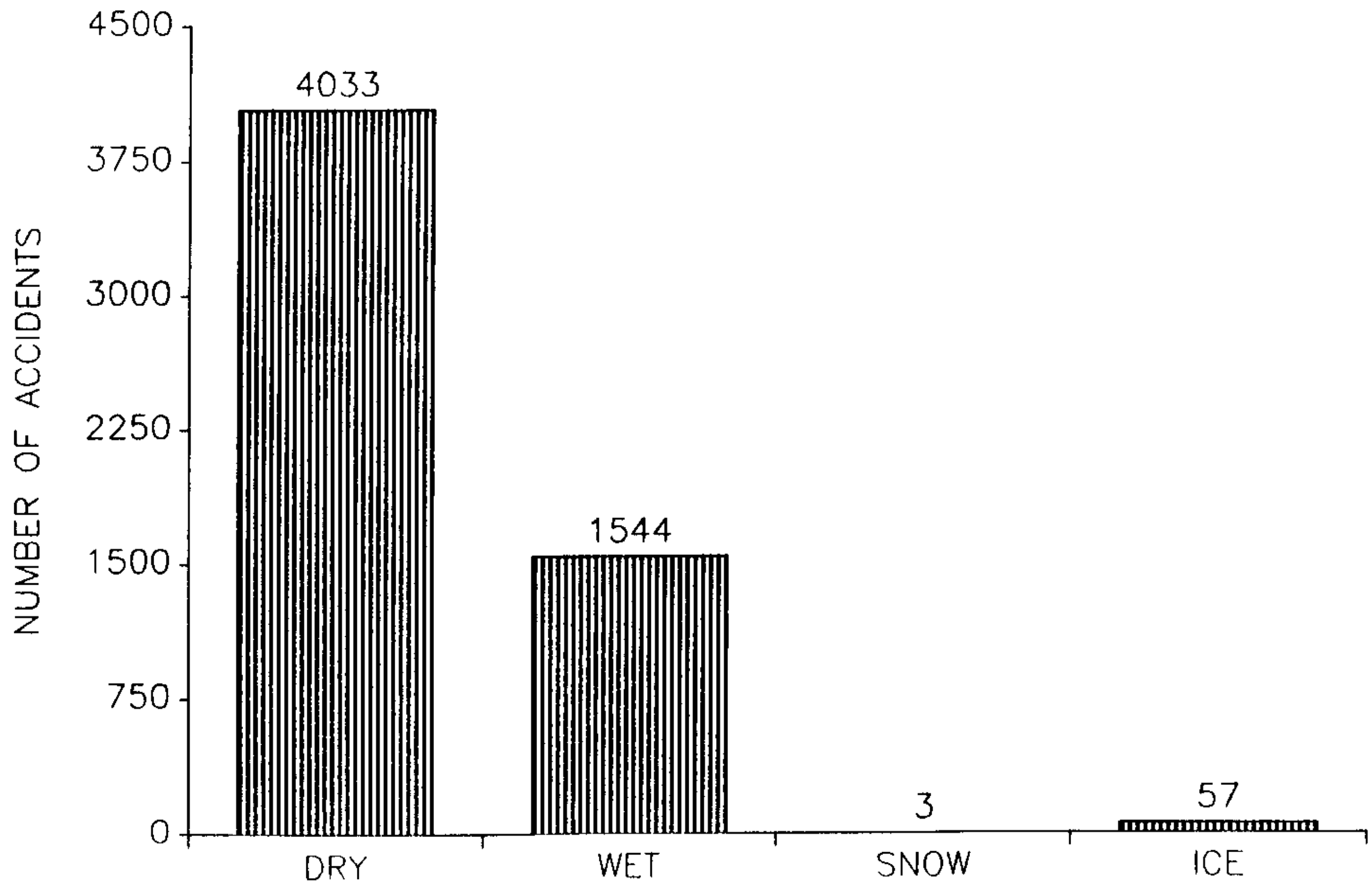

Figure 4.7 Accidents and Road Surface.

SOURCE: WSDOT ACCIDENT REPORT DATA BASE

During icy and snowing weather, only 60 accidents occurred in two years. This low rate was probably caused by the relatively low number of icy and snowy days, or because either many people simply stayed home or the snow and ice reports tended to understate the actual number of accidents that took place during Seattle's brief encounters with winter weather. Many accidents may not have been reported, and many employers, especially larger ones, encouraged their employees to stay home or to use the transit system when driving was dangerous. As noted in the section analyzing the statistics of months, the winter months were low. The low number of winter months disablements may also have corresponded to the low count of accidents in the snow and ice.

Weather

Out of the 5,637 accidents, Figure 4.8 shows that 4,305 accidents occurred during clear or cloudy, but not raining, weather. About one in five accidents took place during inclement weather conditions. This suggests that factors other than the environment may be responsible for many of the accidents.

Location

The average accidents per day for each zone are shown in Figure 4.9, while the average accidents per mile per day are shown in Figure 4.10. Zone 5 had the highest rate of accidents per mile per day, even though it was not one of the largest zones. Zone 5 included the busiest intersection in Washington state, the junction of Interstate 5 and State Route 520. Numerous accidents were reported at this junction. Another high accident rate area within Zone 5 was the Ship Canal Bridge. Zone 1 also had a fairly high accident rate per mile per day because it adjoined Zone 5 at the interchange of I-5 and SR520.

Zone 3 was relatively small but was a "point source" for traffic out of the city in the p.m. and was the destination of much of the highway traffic in the a.m. Relatively few accidents occurred here. Perhaps the highly congested nature of the

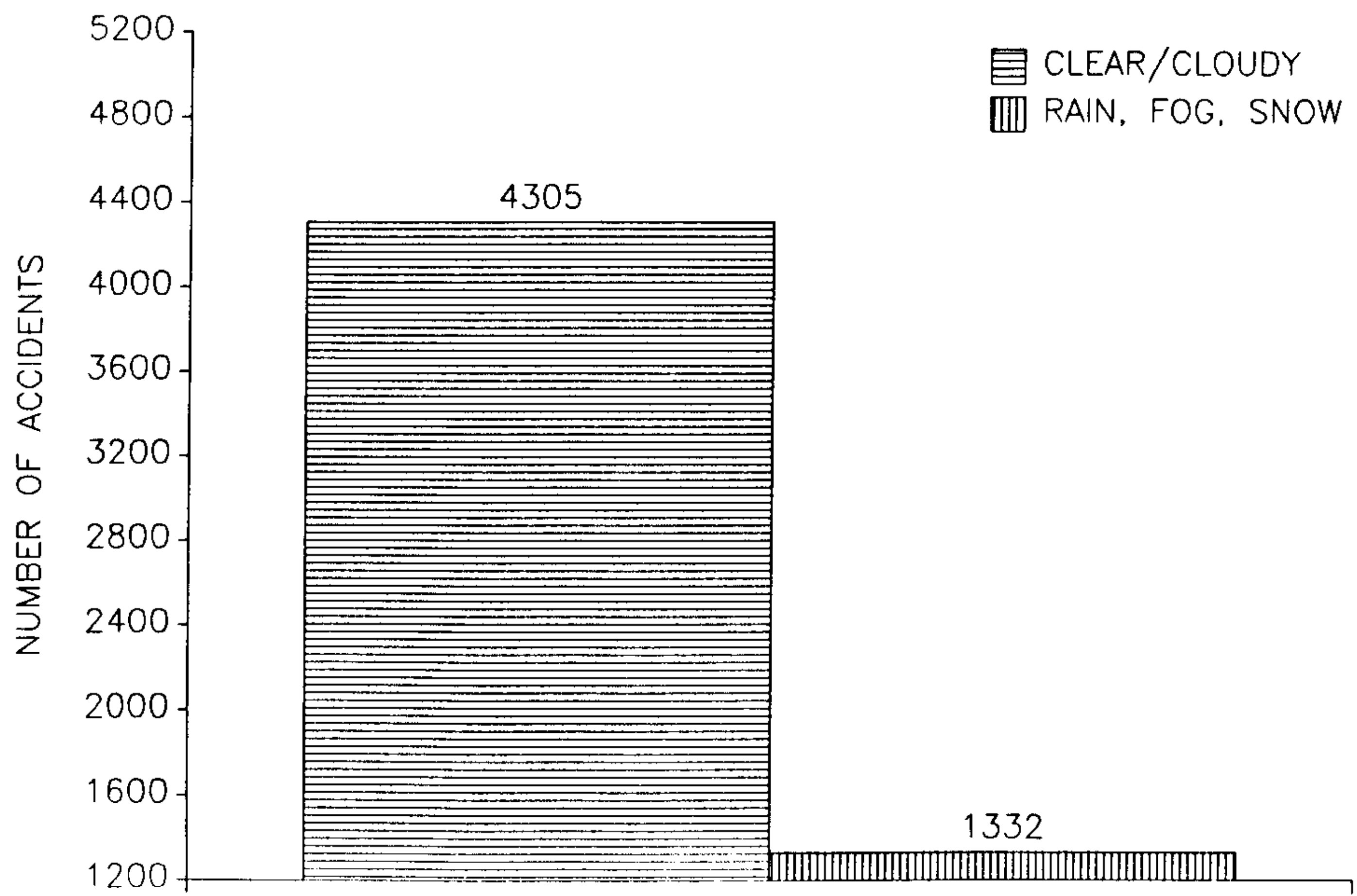

Figure 4.8 Accidents and Weather Type.

SOURCE: WSDOT ACCIDENT REPORT DATA BASE

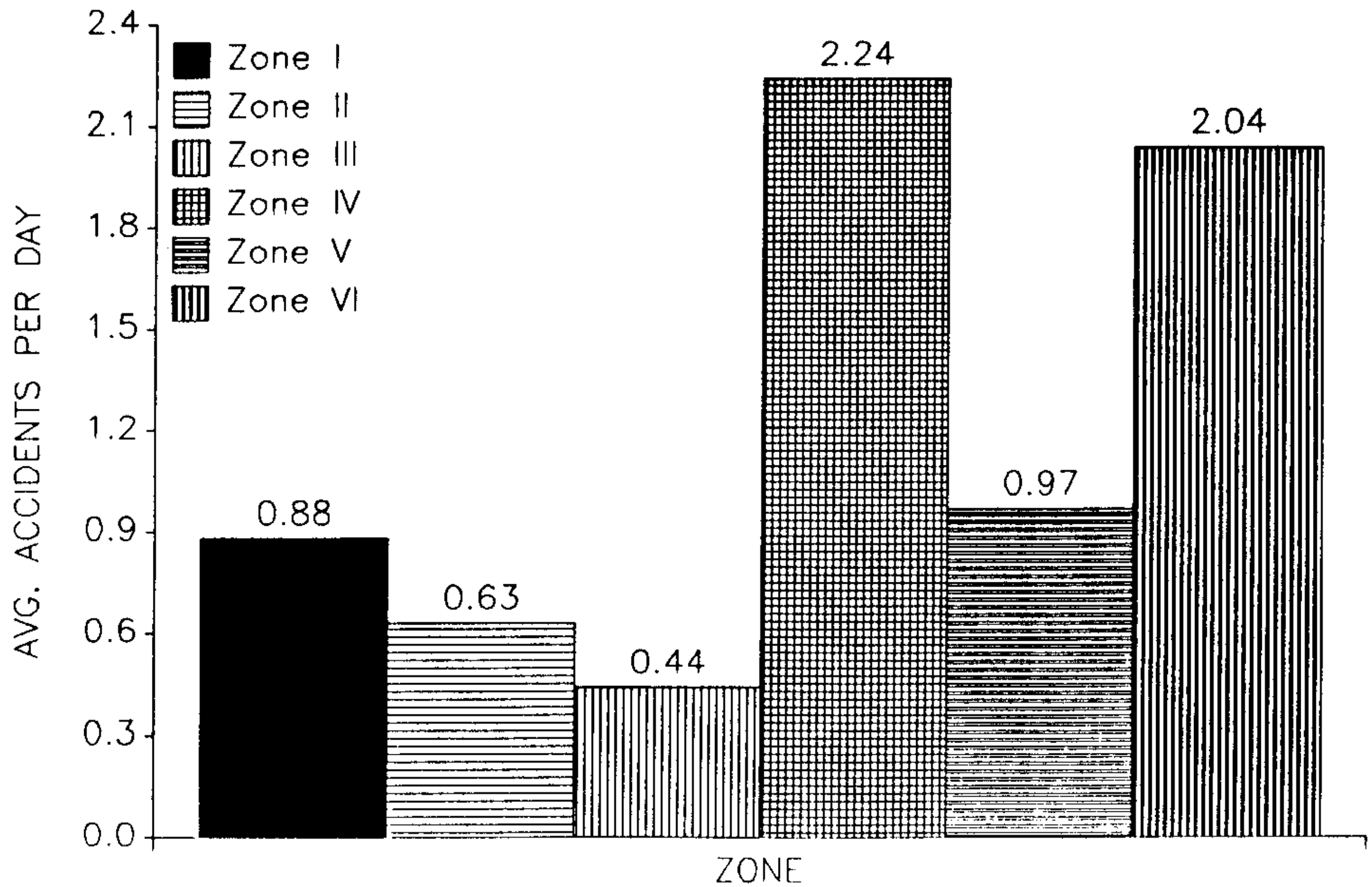

Figure 4.9 Accidents Per Day by Zone.

SOURCE: WSDOT ACCIDENT REPORTS

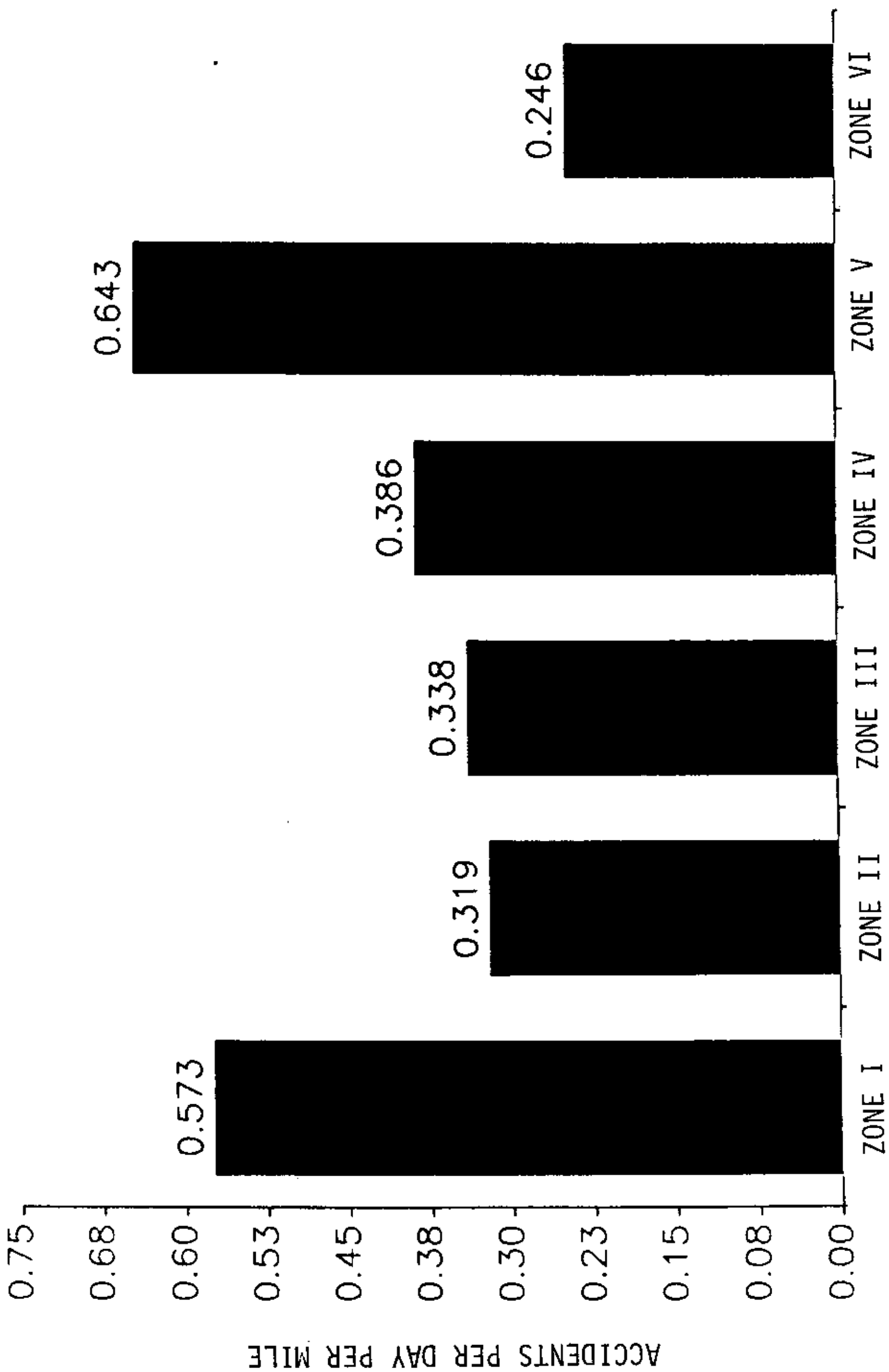

Figure 4.10. Accident Rates Per Day Per Mile.

corridor slowed traffic sufficiently to prevent or to mitigate any damage done in an accident, and therefore reduced the number of reports filed. Although Zone 4 had the highest daily rate of accidents, because of its length, it did not have an accident rate per mile per day much higher than Zone 3. The reason for the somewhat low accident rate per mile per day was probably similar to that of Zone 3.

Although Zone 6 had a fairly high daily accident rate, it also had the longest section of freeway, which gave it the lowest rate of accidents per mile per day.

Figure 4.11 compares accidents by the amount of exposure drivers experienced in different zones. Zone 2, which was the eastern half of the floating bridge, had over five accidents per million miles. Zone 4, where two or more accidents might occur every day, had only a little over half the accidents in terms of vehicle miles over that zone. Zone 2 did have restricted side geometrics and one steep vertical curve, but the number of vehicles per weekday traveling between Seattle and her eastern suburbs on SR 520's two lanes in each direction could not match the number of vehicles traveling on the four lanes each direction of I-5.

Accident Injuries

Injury accidents accounted for 2,206 of the 5,637 accidents in the study area. Tables 4.8 and Table 4.9 show that 60.9 percent of all accidents included no injuries. Of the 39.1 percent accidents that did involve injuries, 71.2 percent caused only one injury, 20.4 percent caused only two injuries, and the remaining 8.4 percent caused three or more injuries, resulting in an average of 1.51 injuries per injury accident or 0.55 injuries per reported accident.

As the number of injuries increased, so did the average duration, except at five or more injuries. The decrease in average duration at five or more injuries was probably due to a small sample size. The average duration for an accident with no injuries was 51.95 minutes. The average duration for an accident with injuries was 58.98 minutes.

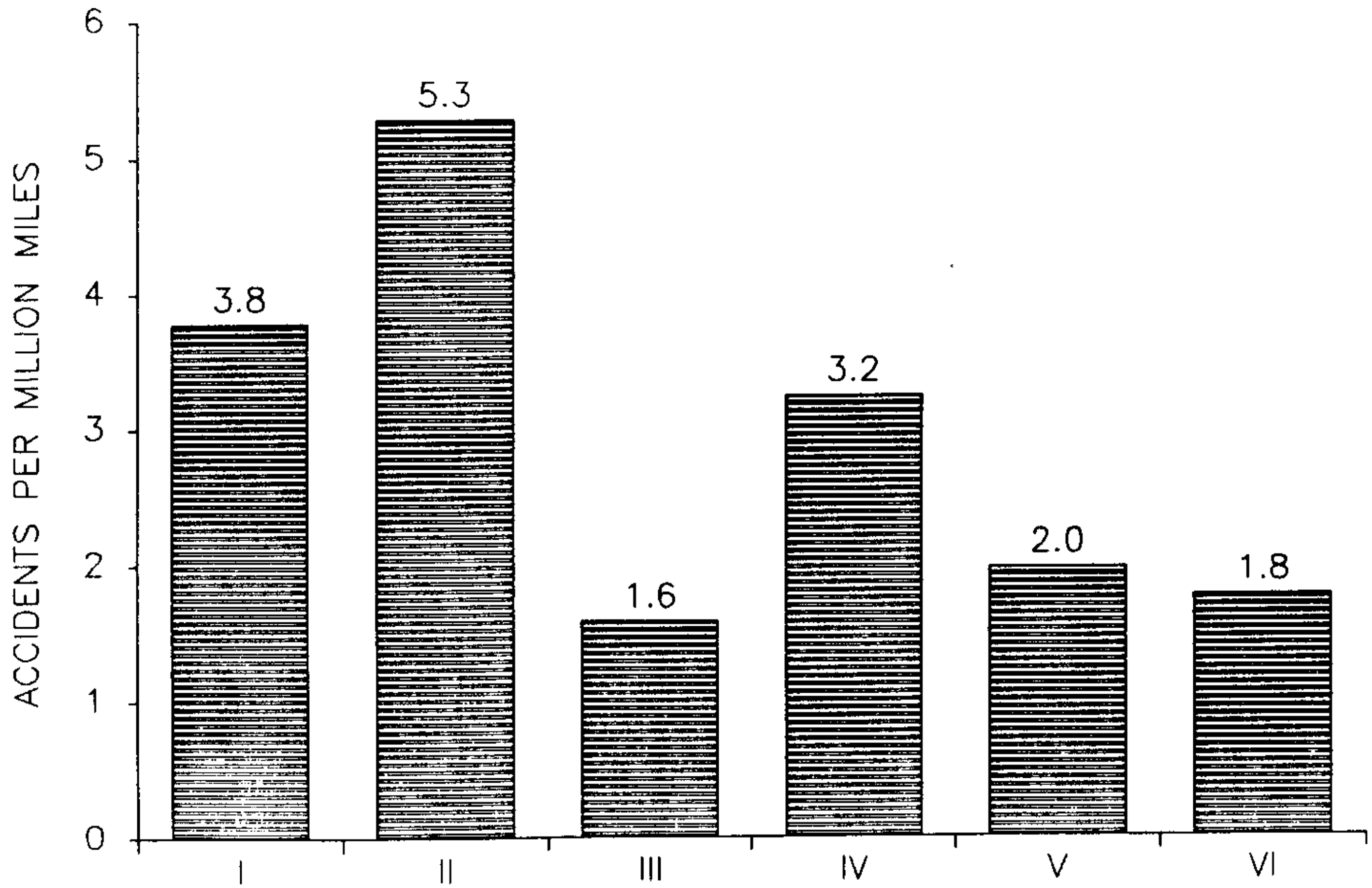

Figure 4.11 Accidents/M.Veh–Mi. by Zone.

SOURCE: WSP ACCIDENT REPORTS
WSDOT RAMP AND ROADWAY VOLUMES

TABLE 4.8

ACCIDENT CHARACTERISTICS
Regarding Number of Injuries

Number Inj.	Cases	% Inj.	% All Acc.	Dur. Sample	% Dur.	Avg. Dur.	Std. Dev.
0	3431	0	60.9	1233	57.2	51.95	31.28
1	1570	71.2	27.9	612	28.4	57.69	29.88
2	449	20.4	8.0	197	9.1	59.71	27.67
3	124	5.6	2.2	77	3.6	64.39	23.08
4	51	2.3	0.8	29	1.3	72.38	33.76
5+	12	0.5	0.2	8	0.4	61.13	28.75
Total	5637	100.0	100.0	2156	100.0	54.97	30.63

Number Injury Accidents=2206
Number Duration Sample=2156
Number Injury Accident Duration Sample=923

TABLE 4.9

ACCIDENT CHARACTERISTICS
Regarding Injuries

Accidents	Avg. Inj.	Avg. Dur.	Std. Dev.	Count	% Count	Dur. Count	% Dur. Sample
Inj.	1.51	58.98	29.18	2206	39.1	926	43
No Inj.	0.00	51.95	31.28	3431	60.9	1230	57
Combined Avg.	0.55	54.97	30.63	5637	100.0	2156	100

TABLE 4.10

ACCIDENT CHARACTERISTICS
by Number of Injuries

Injuries	Count	Total $	% of Total $	Avg. $	Std. Dev. $
None	3431	6,842,352	50.3	1994	2472
1	1498	3,960,581	29.1	2644	2574
2	429	1,766,395	13.0	4117	4063
3	123	673,812	4.9	5478	4579
4	51	284,285	2.1	5574	3522
5+	12	85,425	0.6	7119	5277
Total	**5544**	**13,612,850**	**100.0**	—	—

Accident Property Damage

Property damage was reported in 98.4 percent of the 5,637 accidents, which confirms a bias assumption. The highest damage amount was $82,000, while the average damage amount for the 5,544 accident reports that claimed damage was $2,455.42, resulting in a total damage amount of $13,612,850. Tables 4.10, 4.11, and 4.12 give statistics of property damage on all accident reports claiming property damage by number of injuries, by zone, and by number of vehicles involved in the accident, respectively.

The damages claimed in accidents showed a steady rise, from an average of $1,994 claimed by the 50.3 percent accidents that caused no injuries, to $7,119 for the 0.6 percent accidents that involved five or more injuries.

The average amount of property damage did not vary much across the zones, while the standard deviation varied much more because of the change in percentage of total accidents involved in that zone. Zone 5 had the highest average and standard deviation, with an average of $2,812 and a standard deviation of $4,051. Zone 6 had the highest count of accidents with property damage, but the lowest average ($2,248) and standard deviation ($2,066), while Zone 4 had the highest amount of total property damage.

Larger numbers of vehicles involved also tended to increase the average property damage amount per vehicle, except for single vehicle accidents. The single vehicle average accident damage was somewhat higher, probably because of more severe causes, such as rolling over and driving off the road. Although the average damage increased with the number of vehicles involved, the total damage amount decreased with the number of vehicles involved, again excluding single vehicle accidents.

TABLE 4.11

ACCIDENT CHARACTERISTICS
by Zone

Zone	Count	Total $	% of Total $	Avg. $	Std. Dev. $
1	500	1,306,548	9.6	2613	3188
2	932	2,319,270	17.0	2488	2623
3	254	643,997	4.7	2535	3160
4	1516	3,674,148	27.0	2423	2863
5	716	2,013,714	14.8	2812	4051
6	1626	3,655,173	26.9	2248	2066
Total	5544	13,612,850	100.0	—	—

TABLE 4.12

ACCIDENT CHARACTERISTICS
by Number of Vehicles

Number Vehicles	Count	Total $	% of Total $	Avg. $	Std. Dev. $
1	864	2,055,661	15.2	2379	2789
2	3482	6,712,340	49.3	1928	2364
3	919	3,136,336	23.0	3413	2991
4	211	1,130,793	8.3	5359	3778
5+	68	577,720	4.2	8496	5209
Total	5544	13,612,850	100.0	—	—

TABLE 4.13

PERCENT OF TOTAL ACCIDENTS
by Vehicle vs. Injury Frequency

	Injuries				
Vehicles	0	1	2	3	4
1	10.0	4.7	1.0	0.1	—
2	41.4	16.6	4.0	0.7	0.2
3	7.9	4.9	2.3	0.7	0.4
4	1.2	1.2	0.7	0.4	0.2
5	0.2	0.2	0.2	0.1	—

Frequency and Duration

Table 4.13 shows that two-vehicle accidents occurred most frequently. Accidents of more than five vehicles were very uncommon, as were accidents with more than four injuries.

Table 4.14 shows that the more vehicles and injuries involved in an accident, the longer the duration was, and generally the smaller the standard deviation was. However, there were exceptions to these guidelines.

Figure 4.12 shows that the duration of accidents approximated a normal distribution.

TABLE 4.14

AVERAGE DURATION IN MINUTES
by Vehicle vs. Injury Duration

	Injuries					
Vehicle	0	1	2	3	4	5+
1	50.97	56.75	55.13	66.33	—	—
2	49.43	57.06	56.15	55.26	72.29	70.75
3	61.59	57.98	62.36	64.80	67.15	63.33
4	55.75	59.27	63.00	65.44	71.00	34.00
5	64.11	63.98	76.87	84.60	59.00	—
6	—	66.00	—	99.50	106.00	—
7	—	—	—	—	—	—
8	—	—	—	—	—	90.00

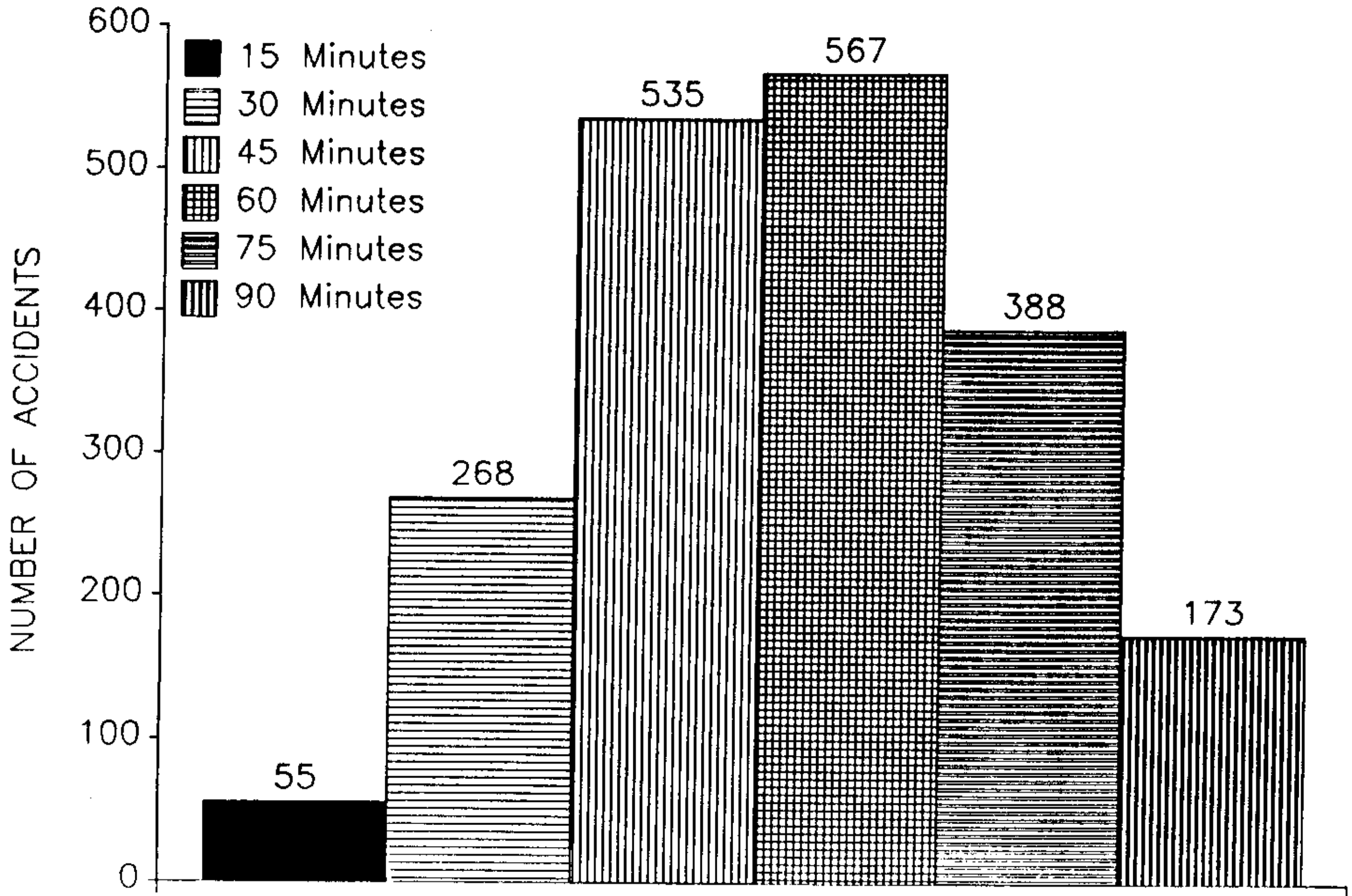

Figure 4.12 Distribution of durations.

SOURCE: WSP DISPATCH DATA

CHAPTER 5

INTERPRETATION: ANALYSIS OF INCIDENT LENGTH AND FREQUENCY

How long do traffic incidents last? How frequent are they? What factors affect the length and frequency of these traffic incidents? The answers to these questions have enormous policy implications. The purpose of this section is to lay the theoretical foundation for an empirical study of incident length and frequency in the Seattle metropolitan area.

The discussion below details the more theoretical aspects of frequency and duration analysis, including modeling techniques for incident frequency, potential problems in duration analysis, the hazard function, and other pertinent issues in frequency and duration analysis.

ANALYSIS OF INCIDENT FREQUENCY

To assess the frequency of incident occurrence, an appropriate statistical modeling technique is needed. Factors such as the weather or the day of the week are expected to affect the frequency of incidents. Within this context, a Poisson distribution is a reasonable description of the number of traffic incidents in a given day. The Poisson regression model, which assumes that the occurrence of the dependent variables follows the Poisson distribution, can effectively overcome the problems caused by discrete and non-negative values of observations that would be found in normal linear regression analysis. (Mannering, 1989) The Poisson distribution has previously been used in such count data applications as trip delay frequency (Mannering and Ahmed, 1988) and beverage choice. (Mullahy, 1986) The Poisson model is as follows:

$$P(n) = (\exp(-\lambda)\lambda^{n})/n! \tag{5.1}$$

where P(n) is the probability of n incidents per day, and

λ is the Poisson parameter, which will be some estimable function of the independent variables.

Such a methodological approach is commonly referred to as a Poisson regression, and it is particularly well suited to the analysis of incident occurrence.

Poisson models can be readily estimated by standard maximum likelihood methods. Herein, the Poisson parameter is defined as follows:

$$\log\lambda = \beta X \tag{5.2}$$

where β is a vector of estimable parameters and X is a vector of commuting and other characteristics for the day. The likelihood function is therefore:

$$L(\beta) = \prod (\exp(-\exp(\beta X))(\exp(\beta X))^{n})/n! \tag{5.3}$$

where the product is over all days. This gives the log-likelihood of :

$$\log L(\beta) = \sum(-\log n! - \exp(\beta X) + n\beta X) \tag{5.4}$$

The object of the Poisson analysis is to estimate the vector, β, thereby providing an estimate of the natural log of the mean number of incidents per day. Finally, note that, unlike standard least squares regression analysis, λ is a deterministic function of X, with the randomness coming from the probability specification for n.

The intensity (the expected value) in the Poisson regression is of conventional form, i.e.,:

$$\lambda = \exp(\beta X) \tag{5.5}$$

for the incident. (Brannas, 1987) The exponential form is adopted to ensure a positive estimator, λ. If $\lambda = \alpha\exp(\beta X)$, where α is gamma distributed with expectation 1 and variance σ^2, the incident frequency is a negative binomial distributed with expectation λ and variance $\lambda(1 + \sigma^2\lambda)$, so that the variance always exceeds the mean. The limit case of $\sigma^2 = 0$ corresponds to the Poisson distribution. (Brannas, 1987)

POISSON STATISTICAL SOFTWARE

The Poisson models for the number of incidents per day for each zone studied were analyzed with the micro computer software package SST. This package had the advantages of being able to analyze a large number of data easily and quickly. Therefore, separate models of the number of incidents in each zone were analyzed, and slightly different results were found for each zone.

DURATIONS AND HAZARD FUNCTIONS

The length of time between when a police officer receives a report of an incident until the incident is cleared is defined as the duration of the incident. Ideally, the time the incident occurs is preferred over the time the report is received, but investigation of the data revealed sporadic and inconsistent reports of occurrence times. Thus, the researchers used the report-received times. A literature has arisen in economics addressing the special problems associated with duration data. This literature has drawn heavily on statistical methods developed largely in industrial engineering, where they are used to describe the useful lives of various machines, and in the biomedical sciences, to describe events such as the survival times of heart transplant recipients. (Kiefer, 1988)

The central concept in the study of duration data is not the unconditional probability of an event taking place (i.e., the probability of an incident lasting exactly 10 minutes), but of its conditional probability (i.e., the probability of an incident ending in the tenth minute given that it has lasted 9 minutes).

The special methods of duration analysis are useful and convenient means of organizing, summarizing, and interpreting data for which a sequence of conditional probabilities is appealing.

Defining a duration precisely requires a time origin (a beginning), a time scale, and an end. The duration of a traffic incident is its length. Typically, durations are the dependent variables under study, and they may be affected by independent variables, such as the condition of the road surface at the incident site.

The Hazard Function

The probability distribution of duration can be specified by the distribution function: (Kiefer, 1988)

$$F(t) = Pr(T<t) \tag{5.6}$$

This equation specifies the probability that the random variable, T, is less than some value, t. In other words, the length of some incident is less than some given value, such as 25 minutes. The corresponding probability density function is as follows:

$$f(t) = dF(t)/dt \tag{5.7}$$

It is also useful to define the survivor function:

$$\begin{aligned} S(t) &= 1 - F(t) \\ &= Pr(T \geq t) \end{aligned} \tag{5.8}$$

The survivor function indicates the probability of an incident lasting longer than 25 minutes. The hazard function is then as follows:

$$h(t) = f(t)/S(t) \tag{5.9}$$

The hazard function indicates the rate at which incidents will be cleared at duration t, given that they have lasted for t minutes.

The hazard function provides a convenient definition of duration dependence. Positive duration dependence exists at the point t* if

$$dh(t)/dt > 0 \text{ at } t=t^* \tag{5.10}$$

Negative duration dependence is then as follows:

$$dh(t)/dt < 0 \text{ at } t=t^* \tag{5.11}$$

A decreasing hazard rate indicates that the probability of the incident ending decreases the longer it has lasted. In other words, the very bad, multiple car incidents are likely to last a very long time. With an increasing hazard rate, the longer the traffic incident lasts, the less likely it is to last longer. (Kiefer, 1988)

Some Statistical Distributions

To help explain the information from the previous section, some examples will be given. The exponential distribution is widely used as a model for duration analysis. It is simple to work with and interpret, and it is often an adequate model for durations that do not exhibit much variation. (Kiefer, 1988) For the exponential distribution with parameter $\gamma > 0$,

$$F(t) = 1 - \exp(-\gamma t) \tag{5.12}$$

$$S(t) = \exp(-\gamma t) \tag{5.13}$$

$$f(t) = \gamma \exp(-\gamma t) \tag{5.14}$$

$$h(t) = \gamma \tag{5.15}$$

A couple simple rules may be used to derive any of the above expressions, given one of the first three. They are as follows:

$$h(t) = f(t)/(1 - F(t)) \; = \; f(t)/S(t) \tag{5.16}$$

$$S(t) = \exp(-\int h(t)dt) \tag{5.17}$$

$$f(t) = h(t)\exp(-\int h(t)dt) \tag{5.18}$$

The exponential distribution is sometimes termed memoryless, because the hazard function is constant, (i.e., not a function of time) and reflects no duration dependence.

The Weibull distribution is only slightly more complicated. It has two parameters (q > 0 and o > 0), with a hazard function:

$$h(t) = \gamma\rho t^{\rho-1} \tag{5.19}$$

The resulting expressions for the distribution function, survivor function, density function, and integrated hazard functions are as follows:

$$F(t) = 1 - \exp(-\gamma t^{\rho}) \tag{5.20}$$

$$S(t) = \exp(-\gamma t^{\rho}) \tag{5.21}$$

$$f(t) = \gamma\rho t^{\rho-1}\exp(-\gamma t^{\rho}) \tag{5.22}$$

$$h(t) = \gamma\rho t^{\rho-1} \tag{5.23}$$

$$dh(t)/dt = (\rho-1)\gamma\rho t^{\rho-2} \tag{5.24}$$

Notice that the hazard function increases with duration if $\rho > 1$, decreases if $\rho < 1$, and remains constant (like the exponential), if $\rho = 1$.

The final distribution to consider is the log-logistic, which is the distribution used in this study. It has parameters $\gamma > 0$, and $\rho > 0$. (Kiefer, 1988) The hazard function is as follows:

$$h(t) = \gamma\rho t^{\rho-1}/(1 + t^{\rho}\gamma) \tag{5.25}$$

The distribution function, survivor function, density function, hazard function, and its derivative are as follows:

$$F(t) = 1 - [1/(1 + t^{\rho}\gamma)] \tag{5.26}$$

$$S(t) = 1/(1 + t^{\rho}\gamma) \tag{5.27}$$

$$f(t) = \gamma\rho t^{\rho-1}/(1 + t^{\rho}\gamma)^2 \tag{5.28}$$

$$h(t) = \gamma\rho t^{\rho-1}/(1 + t^{\rho}\gamma) \tag{5.30}$$

$$dh(t)/dt = (\gamma\rho t^{\rho-2}(\rho - 1 - \gamma t^{\rho}))/(1 + t^{\rho}\gamma)^2 \tag{5.31}$$

For $\rho > 1$ the hazard first increases with duration, then decreases. If $0 < \rho \leq 1$, the hazard function decreases with duration.

ESTIMATION

Non-parametric

Graphical methods of analysis are useful for displaying data on durations and for preliminary analyses, perhaps to suggest functional forms for example. This non-parametric estimation is also called Kaplan-Meier estimation. This type of estimation takes the duration data, divides them into fixed intervals, and performs an analysis, giving empirical plots of the hazard and survivor functions. (SPSS-X User's Guide, 1986) As the name suggests, there are no parameters, and, therefore no exact distribution is specified. In the first period, the probability of failure is given by the following equation:

$$h_1 = n_1/n \tag{5.32}$$

and for the second period it is given as:

$$h_2 = n_2/(n - n_1) \tag{5.33}$$

In this study, Kaplan Meier estimation was used to help select a distribution for the parametric analysis of the duration data.

Parametric Methods

In this estimation method, a particular distribution, with its corresponding parameter(s) (such as exponential or log-logistic distributions), is selected. The specific distribution may be chosen on the basis of a particular theory, convenience, or perhaps some preliminary plotting of data. Other selection criteria include the following (Cox and Oakes, 1984):

1. technical convenience for statistical inference,
2. availability of reasonably simple forms for the survivor function, density, and hazard functions,
3. the ability to admit both over and under-dispersion relative to the exponential distribution,
4. the shape of the empirical hazard function,
5. the behavior of the survivor function, and
6. any connection with a special stochastic model of failure.

In many applications insufficient data will be available to choose among different forms by empirical analysis, and then the choice can legitimately be made on the grounds of convenience.

Parametric methods of estimation frequently use maximum likelihood estimation to evaluate the duration data. This involves writing down a likelihood function and then evaluating it. Such a likelihood function is as follows:

$$L(\pi) = \prod (f(t,\pi)) \tag{5.34}$$

where π is a set of parameters to be estimated, t is the duration of the incident,

$f(t,\pi)$ is a previously specified distribution (e.g., exponential).

The Proportional Hazards Model

Explanatory variables can affect the distribution of durations in many ways. There is no clear-cut starting point for including explanatory variables in duration models. The proportional hazard specification is popular and simple to interpret. The effect of regressors (independent variables) is to multiply the hazard function itself by a scale factor. The accelerated failure time model (the other model of estimation) uses explanatory variables to re-scale the time axis. (Jovanis and Chang, 1988)

The interpretation of the coefficients of the explanatory variables depends on the specification. In general, the coefficient does not have a simple interpretation. The sign of the coefficient indicates the direction of the effect of the explanatory variable on the conditional probability of an incident ending. For example, if "rain" had a positive coefficient, the presence of rain would tend to make the incident last longer.

The proportional hazard model has been widely used in economics and other disciplines. In this model, the hazard function, which depends on a vector of explanatory variables, X, with unknown coefficients B and h0, is factored as follows (Kiefer, 1988):

$$h(t,X\beta,h_0) = y(X\beta)h_0(t) \tag{5.35}$$

where h_0 is a "baseline" hazard corresponding to $y(.) = 1$.

In this specification, the effect of the explanatory variables is to multiply the hazard, h_0, by a factor, y, which does not depend on the duration t. A specification of y in general use is as follows:

$$y(X\beta) = \exp(\beta X) \tag{5.36}$$

In this case, the proportional hazard model can be written as

$$-\ln h_0(t) = t^* = \beta X + \epsilon \tag{5.37}$$

which is a linear model for t* in which the error term has a fully specified distribution (e.g., Weibull, log-logistic, etc.)

The Accelerated Lifetime Model

In the accelerated lifetime model, the effect of the explanatory variables is to re-scale time directly. In this model, the risk components (the x's) alter the rate at which an individual (or length of incident) proceeds along the time axis. As the level of risk components increase, the system hazard function shifts horizontally to reflect accelerated failure time as well as vertically to reflect increased hazard. (Jovanis and Chang, 1988) The hazard function is given as

$$h(t,X\beta) = h_0[ty(X\beta)]y(X\beta) \tag{5.38}$$

Remember that the hazard function for the proportional hazard model was

$$h(t,X\beta) = h_0(t)y(X\beta) \tag{5.39}$$

Using the example of the Weibull distribution, ($S(t) = 1 - \exp(\gamma t^{\rho})$), the proportional hazards hazard rate is as follows:

$$h(t,X\beta) = \exp(\beta X)\rho\gamma t^{\rho-1} \tag{5.40}$$

The accelerated life hazard rate is then as follows:

$$h(t,X\beta) = \exp(\beta X)\rho\gamma(\exp(\beta X)t)^{\rho-1} \tag{5.41}$$

Note that only if $\rho = 1$ does the accelerated life hazard rate equal the proportional hazards hazard rate.

Concerns in Duration Analysis

Duration analysis raises several areas of concern that other forms of regression analysis do not. They are censoring, time-varying explanatory variables, and unobserved heterogeneity. This section will discuss each of these concerns.

It is theoretically possible that an incident may not be observed from beginning to end. If the start of the period is not observed, (i.e., the data begin after the start of the period), the spell, or incident, is said to be left censored. If the data end before the end of the spell is observed, the spell, or incident, is said to be right censored. Several statistical packages (including the SAS statistical package) analyze data with right censoring. In this empirical study on the length of traffic incidents, censoring was not encountered.

The second matter of concern involves time varying explanatory variables. These variables can vary over time, as long as they do not vary over the duration of the incident. These variables must then be non-granger caused (i.e., not endogenous, or varying with the length of the incident) or ancillary with respect to the duration time. Therefore, endogenous explanatory variables (variables that are a function of the length of the duration) cannot be analyzed with the current state of the art.

In this study of traffic incident duration, several variables, including the number of emergency vehicles or number of ambulances, were expected to be endogenous. However, since they were found to be highly correlated with exogenous variables such as the severity of the incident, or the damage in dollars of the incident, they did not have to be included in the model.

The third area of concern is unobserved heterogeneity, or unobserved differences in incident characteristics. The standard procedure to control for population heterogeneity in unobserved variables is the random effect estimator. For single spell data (of which this empirical study is an example), it is the only available estimator. In standard application, the random effect estimator is implemented by assuming a functional form for the structural duration distribution of interest given observed and unobserved variables and a functional form for the distribution of unobservables (frequently normal or log-normal). Maximum likelihood is used to estimate the parameters of the structural duration distribution and the parameters of the distribution of unobservables.

The effects of heterogeneity on apparent duration dependence can be illustrated simply. Consider a random sample of incidents from a mixed distribution consisting of two types of incidents, A and B. Suppose the hazard is higher for incidents of Group A. As time elapses, incidents in Group A will be completed at a higher rate than incidents in Group B. Therefore, as time passes, the fraction of incidents from Group A remaining in the sample falls. Because Group B incidents have a lower hazard function, the decline in the fraction of incidents from Group A shows up as a decline in the hazard function over time. Since one of the criteria for selecting a distribution for the duration times is the shape of the hazard function, the presence of unobserved heterogeneity may bias this decision, resulting in an incorrect specification for the distribution of the duration times. This may then bias the resulting coefficients of the model. For the accelerated life model, the hazard function is a joint distribution (see Kiefer 1988), which is not computationally feasible. Therefore, in the presence of unobserved heterogeneity, the proportional hazards model is required. In this empirical study, careful data analysis was performed to eliminate the presence of unobserved heterogeneity.

Model Selection and Statistical Packages

A major concern in this analysis was selection of the appropriate software. Two software packages appeared to be feasible for the major part of this analysis. They are the LIMDEP and SAS statistical packages. Prior experience indicated that the LIMDEP package suffered from several mechanical problems that made its use much less efficient than the mainframe SAS statistical package. The procedure used is as follows.

As previously discussed, empirical plots of the hazard function (Kaplan-Meier estimation) are an aid in parameterizing the model. The SPSSX statistical package was used to perform this analysis. The resulting analysis indicated that the log-logistic distribution was appropriate.

An example in which the log-normal distribution was selected is given in a recent paper. (Golob, Recker, and Leonard, 1987) The model used the Central Limit Theorem to show the validity of the log-normal distribution in modeling total incident duration. However, this study's data refute the validity of the log-normal in favor of the "wider" probability tails provided by log-logistic distribution. An important note is that the log-logistic is a close approximation of the log-normal. (see Kiefer, 1988.)

The SAS Statistics manual uses the accelerated lifetime approach in the Lifereg procedure to fit parametric models to failure-time data that may be right censored. The class of models includes exponential, Weibull, log-normal, and log-logistic models. The parameters of the model are estimated by maximum likelihood estimation with a Newton-Raphson estimation procedure.

POISSON FREQUENCY MODELS

Two sets of frequency models were estimated, a set of disablement models that used March 1988 through April 1989 data, as discussed in previous chapters, and a set of accident models that used March 1987 through April 1989 data. The estimation results of these models are discussed below.

Disablement Frequency Models

The disablement models include one for blocking disabled vehicles, one for non-blocking disabled vehicles, and one for blocking spills. The researchers attempted to model all disablement types together, but the differences among the types were insufficient to warrant each being modeled separately. Recall that no data for location of disablements were collected except to assure that the disablement occurred in the study area (State Patrol dispatch area 1); therefore, the frequency models for disablements were not based on zones. Below is a discussion of each of the three disablement models: lane blocking, non-lane blocking, and blocking spills, and then the pooled and zonal accident frequency models. First, the significant variables are discussed, followed by coefficient magnitude and the coefficient significance. Lastly, each model's summary statistics are reviewed. The discussion is structured around different variable classes.

The Poisson regression estimation results for blocking, non-blocking, and spill disablements are presented in Table 5.1. A discussion of the explanatory variables follows.

Seasonal Effects

Seasonal effects are evident for the blocking and spill models, as evidenced by the "Long Summer" variable, which is an indicator for all incidents that occurred between June and September. This is a period when many people are moving, which likely results in an increased number of lost loads. Warmer weather also contributes to overheating and, therefore, more blocking disabled vehicles. The non-blocking disabled model did not show a significant effect for an aggregated "long summer" variable, but for June, July, and September the coefficients were positive and strongly significant. August was an anomaly among the summer effects, as it had a negative coefficient in the non-blocking disablement model (the reason that a "long summer" variable could not be used here). There may have been a lower level of travel activity during this favorite vacation season that affected the non-blocking disablements, or older vehicles may have received their annual tune-ups during the first summer months when the non-blocking disablements had been on the rise. An increase in September of non-blocking disablements could be accounted for by an increase in overall traffic volumes caused by the return of vacationers and a large college and university population.

Weekly Trends

The number of non-blocking and spill incidents dropped significantly over the weekend in comparison to weekdays. Blocking disablements had negative coefficients for Saturday and Sunday, indicating significant differences between Saturdays and Sundays, a result that was not found in the non-blocking and spill model.

TABLE 5.1

DISABLEMENT FREQUENCY MODELS
(t-statistics are given in parentheses)

	Model		
Variable	Blocking	Non-Blocking	Spills
Intercept	2.1323 (43.483)	2.4802 (95.385)	0.5986 (10.400)
Long Summer	0.1070 (2.938)	— —	0.3586 (4.685)
Weekend	— —	-0.2529 (-6.573)	-0.4615 (-4.699)
Rain	0.2306 (4.436)	0.0995 (2.091)	— —
Monday	— —	-0.0704 (-1.667)	— —
Tuesday	0.1805 (2.864	— —	— —
Wednesday	0.1973 (3.178)	— —	0.1640 (1.642)
Thursday	0.1363 (2.171)	— —	— —
Friday	0.1841 (2.955)	-0.0753 (-1.667)	— —
Saturday	-0.1804 (-2.626)	— —	— —
Sunday	-0.5732 (-7.468)	— —	— —
June	— —	0.2166 (4.104)	— —
July	— —	0.3065 (6.142)	— —
August	— —	-0.1088 (-1.812)	— —
September	— —	0.1162 (2.120)	— —

Summary Statistics			
Num. Obs.	366	366	366
LL(0)	-5572.93	-7916.47	-752.00
LL(B)	-987.56	-1052.30	-605.35
RHO SQ	0.82	0.87	0.20

Day of Week

Monday and Friday were significant variables for the non-blocking disablement model. The negative coefficient indicated that there were fewer non-blocking incidents than on other weekdays. The usual non-blocking disablements seemed to become blocking disablements during the heavier traffic volumes of Monday and Friday. Spills had a slightly higher coefficient on Wednesdays in comparison to other weekdays. Each day of the week was modeled for blocking disablements (with respect to Monday) and showed a positive effect for all weekdays. Thursday was a little lower than other weekdays, but then the Thursday t-statistic was also somewhat lower than for the other days.

Environmental Effects

Indicator variables for weather type (rain, clear/cloudy,and snowing) and for road surface condition (dry, wet, snow, and ice) were used in the initial disablement models. Only the rain indicator was significant, ($t=4.436$) with a strongly positive coefficient of 0.23 for the blocking disablement model and 0.1 for the non-blocking disablement model. No environmental factor was significant for the spill model. This model indicated a very random process in which only the broadest classification indicator variables seemed to be significant.

Summary Statistics

Since this was a leap year, the number of observations was 366. The rho squared values for the blocking and non-blocking disablement models were very high: 0.82 and 0.87 respectively, indicating that the independent variables were able to explain a good portion of the variance in the data. The rho squared for the spill model (0.20) indicated that a lot of randomness was associated with blocking spills.

Accident Frequency Models

Accidents per day were modeled with a Poisson regression, as described above. Six models and a pooled model were developed to estimate accident frequency and to identify the characteristics peculiar of a particular day that might increase or decrease the number of expected accidents. Below, the pooled model, which included all the variables from any of the other accident frequency models, is presented. Then a comparison of the zonal models is discussed. The independent variables are discussed by variable class: seasonal trends, weekly trends, special events, and environmental conditions. A discussion of the summary statistics follow. Table 5.2 shows each Poisson regression model, its variables, coefficients, t-statistics, and summary statistics.

The Pooled Model

The variables included in the accident frequency models fell into four groups: seasonal, weekly, special events, and environmental. The seasonal variables included indicators for specific months, weekly variables included variables pertaining to the day of the week such as Monday or weekend, special event variables indicated days on which a major sporting event occurred (e.g., a Seahawks football game), and the environmental group included indicators for the weather and road surface conditions.

All variables except the August indicator, the Mariners and Sonics indicators, and the weather indicator were significant in the pooled model. All estimated parameters for the pooled model were positive, indicating an increasing number of accidents per day for each variable; however, the positive parameters were not characteristic of all zones. The day of week indicator variables all had high t-statistics (above 4.0). Wednesday, Thursday, and Friday indicators had the largest coefficients, reflecting an increase in the number of accidents towards the end of the work week. Lower coefficient magnitudes were found for seasonal and special events variables. August had only a 0.08 coefficient, in comparison to 0.17 for March and 0.15 for April. The Sonics and Mariners had a small effect in comparison to the Seahawks and Husky football game-day indicators. Environmental variables included road surface and weather. These were redefined as indicator variables for

TABLE 5.2
COMPARISON OF ACCIDENT FREQUENCY MODELS
(t-statistics are given in parentheses)

Variable	Pooled	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5	Zone 6
Intercept	0.9178 (13.573)	-0.1310 (-6.662)	-1.1410 (-4.923)	-1.6457 (-6.743)	-0.3113 (-2.784)	-1.1320 (-6076)	-0.4225 (-3.644)
Monday	0.4777 (7.788)	0.2805 (1.613)	0.3185 (1.391)	0.5275 (2.169)	0.5777 (5.072)	0.1660 (1.031)	0.6278 (5.629)
Tuesday	0.5568 (9.148)	0.5827 (3.553)	0.5448 (2.489)	0.7263 (3.070)	0.6830 (6.066)	0.2817 (1.778)	0.4970 (4.357)
Wednesday	0.6094 (10.111)	0.6739 (4.172)	0.9635 (4.650)	0.6740 (2.809)	0.7263 (6.494)	0.1678 (1.034)	0.5372 (4.770)
Thursday	0.6094 (10.888)	0.6315 (3.892)	1.0890 (5.380)	0.4946 (2.008)	0.7847 (7.104)	0.2765 (1.747)	0.5895 (5.272)
Friday	0.8236 (14.239)	0.5139 (3.096)	0.6828 (3.216)	0.6865 (2.906)	1.0273 (9.668)	0.7259 (4.990)	0.8221 (7.588)
Saturday	0.4860 (7.663)	0.2887 (1.062)	0.7861 (3.727)	0.4608 (1.861)	0.5542 (4.826)	0.4245 (2.670)	0.4384 (3.790)
February	0.1156 (2.064)	0.4760 (3.469)	0.5606 (3.437)	— —	-0.1720 (-1.649)	— —	0.1277 (1.334)
March	0.1783 (3.377)	0.4602 (3.473)	0.4697 (2.862)	— —	— —	— —	0.1617 (1.784)
April	0.1463 (2.700)	0.3653 (2.652)	0.2267 (1.277)	— —	0.1464 (1.653)	— —	— —
May	0.1412 (2.588)	0.5972 (4.736)	0.5735 (3.703)	— —	-0.1316 (-1.310)	— —	— —
July	0.1264 (2.277)	-0.2801 (-1.547)	— —	— —	— —	0.2973 (2.279)	0.2158 (2.328)
August	0.0811 (1.421)	— —	0.3595 (2.145)	0.3667 (1.941)	0.0913 (0.982)	-0.2385 (-1.455)	— —
October	0.1298 (2.383)	— —	— —	0.3149 (1.663)	0.1561 (1.793)	0.1943 (1.475)	— —
December	0.1819 (3.508)	— —	-0.2435 (-1.125)	0.3977 (2.259)	0.1246 (1.443)	0.1971 (1.500)	0.2382 (2.748)
Mariners	0.0585 (1.535)	0.1498 (1.464	0.1591 (1.309)	— —	— —	0.1577 (1.646)	— —
Sonics	0.0531 (1.139)	— —	0.309 (2.065)	0.2524 (1.463)	— —	— —	0.1051 (1.271)
Seahawks	0.3370 (3.959)	— —	0.4601 (1.519)	0.5630) (1.921)	0.4866 (3.476)	0.3614 (1.671)	— —
Huskies	0.2257 (2.208)	0.3774 (1.270)	— —	— —	— —	0.8391 (4.155)	— —
Road Surface	0.2685 (8.690)	0.2777 (3.741)	-0.2421 (-2.101)	0.1110 (0.982)	0.3317 (7.470)	0.2065 (2.438)	0.2974 (5.349)
Weather	0.0443 (1.308)	— —	— —	— —	— —	0.2341 (2.825)	0.1244 (2.060)
Summary Statistics							
Num. Obs.	731	731	731	731	731	731	731
LL(O)	-34049.5	-1034.4	-943.3	-811.3	-1933.0	-1027.6	-1738.3
LL(B)	-28710.7	-985.5	-842.0	-651.6	-1427.2	-976.1	-1363.7
RHO SQ	0.157	0.047	0.1073	0.197	0.2617	0.0501	0.2155

which weather could be "rain" or "not rain" and road surface was "wet" or "dry." The road surface indicator was much more significant than the "weather" variable in the pooled model.

<u>Comparison of Accident Frequency Models</u>

Each zone was modeled separately for accident frequency. There were significant differences from zone to zone and between the pooled model and each particular zone.

<u>Seasonal Effects</u>. Month indicators for each of the zone models had much higher coefficients and higher t-statistics than the pooled model. There are four negative coefficients within the seasonal group: July in Zone 1, December in Zone 2, February and May in Zone 4, and August in Zone 5. Accidents per day decreased for these cases (subject to the effect of other variables). All other coefficients for month indicators were positive. Spring months had significantly higher coefficients and t-statistics for Zones 1 and 2. This shows an increase in accidents along the State Route 520 section of the study area. This effect is likely the result of spring rain and storms, which are often severe along the Evergreen Point floating bridge, washing waves up onto the bridge deck and splashing over as many as four lanes. The southern CBD section (Zone 3) of the study area did not show any significant effect from the spring months; neither did Zone 5, the section north of the Ship Canal bridge. The Ship Canal bridge section had a negative coefficient for the month of May, perhaps because of longer daylight hours and improved visibility interacting with the unique geometrics of this zone. August had mixed effects from zone to zone. There was no effect in Zones 1 and 6; it had a negative coefficient in Zone 5; and the value of the coefficient ranged from 0.09 in Zone 4 to 0.37 in Zone 3. Three of the four Interstate 5 zones had a significant positive coefficient for October (i.e., increased accident frequency). Zones 4 and 5 had comparable magnitudes of 0.15 and 0.19, but their t-statistics were weak in comparison to Zone 3, in which there was a larger coefficient of 0.3 and a more reliable t-statistic of 1.7.

The December indicator was significant in all but Zone 1. Zone 2 had an unexpected negative value for this variable, indicating a drop in accidents per day for the Portage bay overpass and Montlake areas. Perhaps this reflects the long Christmas vacation at the University of Washington (adjacent to the zone) and flexible work hours at colleges and universities nearby between quarters. Other zones had positive coefficients for December, as would be expected. As with August and October, Zone 3 had the highest coefficient, but it was comparable with Zone 6, which had only a slightly smaller coefficient (0.24) and a higher t-statistic (2.7) than Zone 3. These two zones were sensitive to seasonal changes taking place at the end of the year. In contrast, Zones 1, 2, 4, and 5 seemed to be more sensitive to seasonal changes taking place at the beginning of the year.

Day of Week. Day of Week indicator variables were positive for all days in all models; each day was modeled with respect to Sunday and positive values were to be expected, since Sunday volumes and therefore accident exposure levels were low. In general, the day of week indicators were highly significant and strongly positive. The only exception was in Zone 5, where only Friday and Saturday coefficients rose above 0.3 and the t-statistic above 2.0. For all models, Mondays had a much smaller coefficient than one might have expected from the popular conception relating to the number of morning, rush hour accidents. They ranged from 0.11 in Zone 5 to 0.62 in Zone 6. Wednesday and Thursday both had much higher coefficients for Zones 1, 2, 3, 4, and 6; they also were higher than the Friday indicator in two of the zones (e.g., Zone 4 was 1.02 for Thursday and 0.55 for Friday). Weekday t-statistics were all above 2.0 except in Zone 5, as mentioned above. The Saturday indicator was strongest in the zones away from the CBD, an indication of a lower level of traffic exposure in a largely abandoned work area (Zones 3 and 4). Zones 1 and 3 were situated closest to the CBD and had low magnitude and significance.

Special Events. The effect of planned special events on accident frequency was explored by including indicator variables for scheduled game days of local sports teams: the Mariners baseball team, the Sonics basketball team, the Seahawks football team, and the Husky football team. Lower than average t-statistics might have been expected for the football game days, since all Huskies games were played on a Saturday and all Seahawks games were played on Sunday. A Pearson's correlation coefficient analysis did not show that Saturday and Sunday were highly correlated to Huskies and Seahawks games, and the high t-statistics for both variables in zones near the game sites did bear this out. As a class, the special events variables were less uniform in their magnitude and significance than the Seasonal or the Day of Week variables. Proximity to the game site and inclusion of important weaving sections seemed to have had the greatest effect. The Seahawks played their games in the Kingdome, a county operated facility only a few blocks away from Interstate 5 and approximately 1-1/2 miles from the intersection of State Route 520. On Interstate 5 Zones 3, 4, and 5 showed significant positive effects of Seahawk games on accident frequency. Zone 6 was too far removed from the facility to show any significant effect. Interestingly, Zone 2 on State Route 520, not Zone 1, had a significant t-statistic for Seahawk games. Zone 2 was farther from the game area than Zone 1. However, the 84th and the SR 520 High Occupancy Vehicle lanes merged at the westbound extremity of Zone 2; traffic bound for the Seahawk game did not have any other required weaving sections until it reached Interstate 5 in Zone 3. Here the magnitude of the coefficient was highest for all zones with respect to the Seahawk indicator, since all game attenders had to merge into southbound lanes.

Husky games had significant effect on Zones 1 and 5. Zone 1 lay adjacent to the game site. As parking was extremely difficult to find near the stadium, many Husky fans parked far north of the campus at park-and-ride lots, which accounted for the increased accident frequency in Zone 5.

Considering the low attendance at the Mariners games, the increase in expected number of accidents for those game days in Zones 1 and 2 (the SR 520 section) and in Zone 5 was surprising. Each of these zones had comparable magnitudes for their coefficients (Zone 1 was 0.15, Zone 2 was 0.16, Zone 5 was 0.16) and the t-statistics were much the same (Zone 1 was 1.5, Zone 2 was 1.3 and Zone 5 was 1.6). What is puzzling was that Zones 3 and 4 did not show any significant effects. Most likely, much of the game traffic was diverted to arterials and other surface streets, since many of the games were scheduled for weekdays near the evening rush hour.

The Seattle Sonics played their games in the Seattle Center Coliseum. Their games were scheduled for both weekdays and for weekends. Access to the coliseum for northbound traffic included a left side off-ramp to Mercer Street. For southbound traffic (including traffic from SR 520) the "Mercer mess" had to be negotiated. Traffic merging from SR 520 had to cross from the left side on ramp to Interstate 5 and move four lanes right to the Mercer Street exit within just a mile. Zones 2 and 3 had significant coefficients for Sonics game days. Zone 2 had the most significant t-statistic, probably for the same reason given above for the 84th Street merge westbound. Zone 3 had a lower t-statistic, and a lower effect, but this may have been because of a higher overall capacity for Interstate 5. Zone 6 had no significant Special events indicators except for the Sonics, and this one had a low t-statistic (1.3) and coefficient (0.11) in comparison to the other zone models.

Environmental Factors. Two environmental factors were included in the model: road surface character ("wet surface") and weather (raining versus all other weather types). The rain indicator showed surprisingly little effect on accident frequency. Only Zones 5 and 6 had significant results for rain: the pooled model had a very small coefficient (0.04) and t-statistic (1.3), indicating that the system as a whole did not respond much to rainy weather. Zone 5 had a much higher coefficient (0.23) and t-statistic (2.8) than Zone 6 (0.12 and 2.1, respectively). Zone 5 included the northbound merge into the fast lanes of the mainline from the express lanes; the speed differentials there and the merge and weave maneuvers required to reach the Northgate Shopping Mall increased the number of factors a driver had to deal with and could have contributed to an increase in accidents.

The road surface indicator was significant for all zones. The variable was derived from the accident reports from which a daily average was determined. There was a very low amount of variation within each day's reported road surface for accidents. The lowest coefficient (0.11) and t-statistic (0.98) for this factor was found in Zone 3. Much of Zone 4 was covered by structures such as the Freeway Park, the Convention Center, and numerous overpasses. The changing illumination levels and the effect of a wet road surface contributed to higher accident frequencies. Other zones had coefficient magnitudes ranging from 0.21 in Zone 5 to 0.33 in Zone 4. All these t-statistics were above 2.0; the highest was for Zone 6 (coefficient was 0.297, t-statistic was 5.3). In general, precipitation fell heavier north and east of Seattle than in the CBD area itself (Jones, 1987); it is not surprising, then, that rain and its related variable, wet road, should be more significant in zones further from the city.

<u>Model Statistics</u>. The total number of observations for each model was 731, one observation for each day of the two-year study period. Rho squareds are given at the bottom of Table 5.2. Zones 3, 4, and 6 had high rho squares, indicating that much of the dependent variable variation (i.e., accidents per day for each zone) has been explained. The Zone 2 model performed less well. Zones 1 and 5 had a lot of variance that remains unexplained. Still, significant variables for accident frequency have been identified and their relative magnitudes gave an idea of how accident frequency was affected. Since the number of observations for all models was the same, likelihood ratio tests were not appropriate to compare the pooled and sub models.

DURATION

The following is a discussion of the estimated duration models, the sign and magnitude of the model coefficients, the significance of the t-statistics, and the effect of the variables on duration. Each model is shown in Table 5.3. A pooled model is not presented because of specification error, which renders the results of a pooled model meaningless. Specification errors are the result of the large differences in the variables found to be significant in each zone's duration model.

Accident duration, the dependent variable, was measured from the time the State Patrol dispatcher was notified to the time the officer responding to the accident left the scene. The classes of variables used were narrower than for the frequency models, mostly because the data used pertained to specific events. New classes included driver and accident characteristics. The Day of Week class was dropped and replaced by Daily Variations, which included the rush hour variable, accidents per day, dark, and funtime variables. The rush hour variable indicated whether an accident occurred during the morning or evening rush hours (6-9 a.m. and 3-6 p.m.) in the corresponding rush-hour direction. The accidents per day variable was a proxy measure of the State Patrol workload in the study area. Average number of disablements per day was used in preliminary models for the same purpose, but it did not produce significant results in any of the zones. A variable indicating accidents going northbound on Interstate 5 was included. The dark variable indicated accidents that occurred after sunset; the funtime variable indicated accidents that occurred between 7:00 p.m. and midnight on Fridays and Saturdays, traditional recreational periods.

Seasonal, day of week, time of day, driver characteristics, special events, and environmental variables are discussed in turn and compared by zone. Note that Zone 3 was left out of the duration analysis because of an insufficient sample size of accident records matched up to dispatch data. Unlike with the frequency models, the relative magnitude of variable coefficients could not be readily compared across

TABLE 5.3
COMPARISON OF DURATION MODELS FOR EACH ZONE
LOG LOGISTIC MODELS
(t-statistics are given in parentheses)

Variable	Zone 1	Zone 2	Zone 4	Zone 5	Zone 6
Intercept	3.7367	3.6403	3.6170	3.8742	3.7068
	(14.95)	(33.06)	(45.78)	(28.68)	(28.51)
April	—	—	—	0.1302	—
	—	—	—	(1.56)	—
July	—	—	—	0.1717	—
	—	—	—	(2.21)	—
August	0.1953	—	—	—	—
	(1.74)	—	—	—	—
September	—	—	—	—	0.0942
	—	—	—	—	(1.64)
October	—	—	—	—	-0.1765
	—	—	—	—	(-2.90)
November	0.2663	—	—	0.0849	—
	(1.93)	—	—	(1.20)	—
Accidents per Day	0.0106	—	—	—	—
	(1.77)	—	—	—	—
Rush	-0.2281	-0.0810	-0.0825	—	-0.1860
	(-2.58)	(-1.12)	(-1.62)	—	(-4.73)
Husky	-0.2730	—	—	—	—
	(-1.30)	—	—	—	—
Number of Injuries	0.0953	—	—	—	—
	(2.58)	—	—	—	—
Number of Vehicles	—	0.0688	0.0412	0.1337	—
	—	(2.25)	(1.83)	(5.74)	—
Alcohol	—	-0.1155	-0.0772	-0.1321	—
	—	(-1.52)	(-1.50)	(-2.31)	—
Damage (Thou. $)	—	0.0213	0.0164	—	0.0273
	—	(2.130)	(1.64)	—	(3.90)
Dark	—	—	0.1343	0.1663	—
	—	—	(3.13)	(3.44)	—
Northbound	—	—	0.0500	—	—
	—	—	(1.36)	—	—
Number of Lanes Blocked	—	—	0.0869	—	0.0610
	—	—	(3.02)	—	(2.33)
Truck or Bus	—	—	0.2318	0.2102	0.1785
	—	—	(3.64)	(3.21)	(3.32)
Funtime	—	—	0.0795	—	-0.1553
	—	—	(1.25)	—	(-2.75)
Property Damage Only	—	—	—	-0.1856	—
	—	—	—	(-4.15)	—
Injury Accident	—	—	—	—	0.1003
	—	—	—	—	(1.78)
Young Driver	—	—	—	—	-0.0721
	—	—	—	—	(-1.14)
Vehicles Only	—	—	—	—	0.0796
	—	—	—	—	(1.94)
Scale Parameter	0.2313	0.0688	0.0261	0.2343	0.2631
	(16.17)	(2.25)	(2.93)	(22.10)	(31.70)
Summary Statistics					
Num. Obs.	183	267	622	346	716
LL(B)	-98.2	-162.5	-420.5	-194.2	-487.6

models (see Kiefer, 1988), except to note common significant variables with variable classes. Relative coefficient values could, of course, be compared to other variables within the same zone model.

Also note that accident durations were highly variable within any of several subgroups, as discussed in Chapter 3. The overall average accident duration was 54.97 minutes, with a standard deviation of 30.63 minutes. A much smaller amount of variation was found in the frequency models discussed above, and it follows that while the duration model parameters often had lower t-statistics, the information gained by estimation of a parameter with a t-statistic over 1.0 was important.

Seasonal Effects

The duration of accidents was not consistently affected by seasonal or day of week variations to the extent that frequency of accidents were. Some months did have significant effects for particular zones. For example, April, July, and November accidents had longer durations in Zone 5. Zone 5 also tended to have more frequent accidents in July, which may have contributed to increased duration as the same number of troopers responded to a growing number of accidents. The duration from Zone 1 had increased in August and November. Zone 1 had numerous closely spaced off-ramps, and several lane changes were required to stay on the mainline. The fact that August is a favorite vacation month and Thanksgiving travel in November brings many unfamiliar drivers through Seattle's CBD may explain the increase in duration. Zone 6 had an increase in duration during September, followed by decrease for October. Zone 6 was the longest zone in this study, and its larger sample of accidents may reflect the initial rise in accident rates for September as vacations were finished, school started again, and business as usual resumed.

Daily Variations

Accidents per day were significant only in Zone 1. The Portage Bay and Montlake areas included all the eastbound ramps to State Route 520. During the p.m. peak hours, traffic was very dense, only two lanes were available for traffic, and shoulders were very narrow or non-existent except for a small section just east of Montlake. Additional accidents would have created more merging sections in this area, making it much more difficult for accident investigation and vehicle removal, thus increasing overall duration here. Zone 2 was adjacent to Zone 1, but it did not include the merges in Zone 1 and so no duration increase resulting from accident frequency was indicated.

The rush hour indicator showed a decrease in accident duration during the peak hours in Zones 1, 2, 4, and 6. In Zones 1 and 2 the State DOT had tow trucks stationed at each end of the Evergreen Point floating bridge. This service, an advantage to detection and vehicle clearing, was unavailable during the rest of the day. Then local towing companies provided accident clearing services on a rotating basis but were not required to dedicate their vehicles to any specific location for faster service. The peak hours were generally recognized as problem times for traffic, and public agencies allocated an increased incident response supply at this time. The durations were reduced because of this anticipated response.

After dark accidents were longer in Zones 4 and 5. Poorer sight distances, difficult investigation environments, and perhaps the prior activities of drivers at night (e.g., parties, stopping at the local tavern) contributed to longer clearing times. Night accidents also tended to be more severe, since drivers may not have recognized hazards until too late. Moreover, Zones 4 and 5 gave access to dense residential sections of Seattle. These sections would have been traversed by persons returning home from parties or other recreational activities. In contrast, Zones 1 and 2 had more restricted geometrics that allowed better illumination of the highway, and Zone 6 serviced a less dense residential area. These facts may have accounted for the insignificant result of the Dark variable in these zones.

As expected, the northbound accident indicator variable was significant in Zone 4 (Zones 1 and 2 ran east and west). Zone 4 was close to the city and received much of the outbound p.m. peak traffic. There were few places where an accident would not have interfered with access or egress from ramps in this area and thus contributed to longer durations. During weekday peak hours, headways were very small, impeding response by emergency vehicles. Zone 5 and, especially, Zone 6 had traffic volumes that decreased as distance from the CBD increased, explaining why the northbound variable was not significant in those zones.

The "Funtime" indicator was positive for Zone 4 and negative for Zone 6. Zone 4 included most of the access to Seattle's major public facilities in the Seattle Center and the CBD. The magnitude of the coefficient was small in comparison to the other variables for the Zone 4 model, but there was definitely an effect in duration increases.

Special Events

Only the Husky indicator was found significant in any of the zones. A highly organized, cooperative effort of local police, the University of Washington, and transit agencies in Special Event Incident Management produced a negative coefficient in Zone 1, the section of highway adjacent to the game site. More accidents were expected on game days, as seen in the frequency models, but the level of police presence and excellent response drove down the time necessary to clear accidents and complete investigations. This anticipated response effect demonstrates the effectiveness of an incident management program.

Driver and Vehicle Characteristics

Approximately one third of the drivers involved in an accident in the study area had been drinking alcoholic beverages. This includes drivers whose ability was not judged to be impaired. The impact of these drivers on accident duration was not an increase but a decrease in the length of time to clear the accident. Although a great many fatal and severe accidents are linked to drinking, the study seemed to indicate that many of these accidents were minor and could be dealt with quickly.

Drivers under 65 years old tended to have shorter accident durations for Zone 6. The other zones were relatively small compared to Zone 6, and the high traffic volumes put everyone at about the same risk. Farther from the CBD the effect of individual driver characteristics became more pronounced. The magnitude of the coefficient was comparable to the other variables in the Zone 6 model, just a little low; the t-statistic was also weaker than for the other variables.

Truck or bus accidents had strongly significant, relatively high magnitude coefficients for the Interstate 5 zones. State Route 520 (Zones 1 and 2) did not show a significant effect, which was likely a result of the fact that most Interstate truckers used Interstate 90 to the south instead of SR 520 when possible. In contrast, Interstate 5 carried much of the north and south truck traffic. The exposure to possible truck accidents was therefore much higher in Zones 4, 5, and 6. The magnitude of the coefficients was one of the highest in each of the models where it was significant. The effect of a major tractor-trailer type accident, especially one involving anything suspected of being a hazardous material, can close the freeway in both directions while the fire department makes certain that the truck and cargo are innocuous. This standard operating procedure has only a remote chance of preventing further injury, and the cost in terms of delay cannot be exceeded by any other known incident management method.

Accidents involving only vehicles were longer than accidents involving other creatures, such as dogs or ducks. Zone 6 had more greenbelt than other zones, so accidents involving animals could be adequately analyzed (i.e., significant vehicle only variable). The other zones had an insufficient number of animal accidents.

The number of vehicles involved in an accident raised the duration of the accident significantly in Zones 2, 4, and 5. The poor geometrics of these zones (narrow shoulders, weaving areas, and so on) caused a significant increase in duration as additional vehicles became involved in the accident.

Accident Results

Property damage, measured in thousands of dollars, was significant in indicating longer accidents for Zones 2, 4, and 6. Property Damage as an indicator variable was significant in Zone 5 (see below). Zone 1 did not show a significant effect attributable to property damage, perhaps because of the low variance of damage in this zone. The t-statistics were relatively high for this variable, and the effect of the parameter coefficient was considerable, since the average property damage per accident was $2,415.

Property damage only was an indicator variable that separated property damage from injury and fatal accidents. Zone 5 did not show a significant effect on duration by dollars of property damage, but it did indicate that accidents involving only property damage were expected to have a shorter duration than accidents involving injuries and fatalities. This variable was not significant in any other zone. Perhaps the significance of this variable was attributable to the fact that this zone had a higher mix of injury, property, and fatalities than other zones.

Injury accidents were significant for Zone 6. The number of persons injured was significant for Zone 1, indicating an increasing duration as the number of injuries increased. Three different measures of accident severity were significant for different zones. For Zones 2, 3, and 5, severity was indicated significantly by the number of vehicles involved in the accident; in Zone 1, the significant severity measure was number of persons injured (see below); for Zone 6 the significant severity measure was property damage only.

Number of lanes blocked had a very significant coefficient for Zones 4 and 6 (Zone 4=3.02 and Zone 6=2.33); the magnitude of the coefficient was large compared to other variables in each model.

Finally, for all models, the log-logistic scale parameter was less than 1, indicating that the hazard function was decreasing throughout. This means that as accident durations became larger, the likelihood that they would end soon became smaller. This was an expected result that attestested to the effect of accident severity on accident impact duration.

SUMMARY

This chapter has presented frequency models for disablement and accidents and duration models for accidents. The model estimation results showed that a wide range of factors affected the frequency and duration of incidents in the Seattle area. Moreover, there were significant differences within the Seattle area, as evidenced by the different coefficient estimates across zones. The results in this regard have significant implications for incident management in the Seattle area. These implications will be discussed in detail in Volume IV of this report.

REFERENCES

Brannas, Kurt, "Swedish Travel Frequencies," working paper, Dept. of Statistics, University of Umea, Umea, Sweden, 1987.

Cox, D.R., and D. Oakes, "Analysis of Survival Data," Chapman and Hall, New York, NY, 1984.

Giuliano, Genevieve, "Incident Characteristics, Frequency, and Duration on a High Volume Urban Freeway," Institute of Transportation Studies, University of California, Irvine, 1988.

Golob, Thomas F., Recker, Wilfred W., and Leonard, John D., "An Analysis of the Severity and Incident Duration of Truck-Involved Freeway Accidents," Accident Analysis and Prevention, Vol. 19, 1987.

Jovanis, Paul P., and Chang, Hsin-Li, "Modeling the Relationship of Accidents to Miles Traveled," Transportation Research Record 1068, TRB, National Research Council, Washington, D.C., 1987.

Jovanis, Paul P., and Chang, Hsin-Li, "Disaggregate Model of Highway Accident Occurrence Using Survival Theory," working paper, 1988.

Kiefer, Nicholas M., "Economic Duration Data and Hazard Functions," Journal of Economic Literature, Vol. 26, 1988.

Mannering, Fred L., "Poisson Analysis of Commuter Flexibility in Changing Routes and Departure Times," Transportation Research Board, Vol. 23B, 1989.

Mannering, Fred L., and Ahmed, Mohammed, "Occurrence, Frequency, and Duration of Commuters' Work-to-Home Departure Delay," forthcoming in Transportation Research, part B.

Mullahy, John, "Specification and Testing of Some Modified Count Data Models," Journal of Econometrics, Vol. 33, 1986.

SPSS-X User's Guide, SPSS Inc., Chicago, Il., 1986.

SAS User's Guide: Statistics, SAS Institute Inc., Cary, N.C., 1985.