Statistical Models of At-Grade Intersection Accidents—Addendum

PUBLICATION NO. FHWA-RD-99-094 MARCH 2000

U.S. Department of Transportation
Federal Highway Administration

Research, Development, and Technology
Turner-Fairbank Highway Research Center
6300 Georgetown Pike
McLean, VA 22101-2296

FOREWORD

This report is an addendum to the published Report, FHWA-RD-96-125, titled *Statistical Models of At-Grade Intersection Accidents*. The objective of both reports is to develop accident prediction models relating total intersection accidents to highway design elements. The design elements considered include functional class, traffic flow, channelization, traffic control type, median, access control, terrain type, number of lanes, lane width, shoulder width, and lighting. While the previously published report is only for multi-vehicle accidents, this report includes all collision types (single and multi-vehicle). The result is a preliminary effort in the development of the Interactive Highway Safety Design Model (IHSDM), which is a Federal Highway Administration (FHWA) objective to develop a highway safety evaluation tool.

Based on retrospective analysis, several statistical modeling techniques were tried. Besides using innovative statistical techniques (e.g., Poisson and negative binomial regression models), five preliminary accident models were developed for at-grade intersections: (1) Rural, four-leg, STOP-controlled; (2) Rural, three-leg, STOP-controlled; (3) Urban, four-leg, STOP-controlled; (4) Urban, three-leg, STOP-controlled; and (5) Urban, four-leg, signalized.

Michael F. Trentacoste,
Director, Office of Safety and Traffic Operations
Research and Development

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA-RD-99-094		
4. Title and Subtitle Statistical Models of At-Grade Intersection Accidents—Addendum		5. Report Date
		8. Performing Organization Report No. 5231
7. Author(s) K.M. Bauer and D.W. Harwood		
9. Performing Organization Name and Address Midwest Research Institute 425 Volker Boulevard Kansas City, Missouri 64110-2299		10. Work Unit No. (TRAIS) 3A5A
		11. Contract or Grant No.
12. Sponsoring Agency Name and Address Federal Highway Administration Office of Safety and Traffic Operations Research and Development 6300 Georgetown Pike McLean, Virginia 22101-2296		13. Type of Report and Period Covered Technical Report June 1998–October 1998
		14. Sponsoring Agency Code

15. Supplementary Notes

Contracting Officer's Technical Representative (COTR): Joe Bared, HSR-20

16. Abstract

This report is an addendum to the work published in FHWA-RD-96-125 titled "Statistical Models of At-Grade Intersection Accidents." The objective of both research studies was to develop statistical models of the relationship between traffic accidents and highway geometric elements for at-grade intersections. While the previously published report used only multiple-vehicle accidents in developing predictive models, this addendum presents models based on all collision types (including both multiple-vehicle and single-vehicle accidents).

The statistical modeling approaches used in the research included lognormal, Poisson, and negative binomial regression analyses. The models for all collision types are similar to those developed in the previous report for multiple-vehicle accidents. The regression models of the relationships between accidents and intersection geometric design, traffic control, and traffic volume variables were found to explain between 16 and 39 percent of the variability in the accident data. However, most of that variability was explained by the traffic volume variables considered (major road and crossroad ADTs). Geometric design variables accounted for only a small additional portion of the variability.

Generally, negative binomial regression models were developed to fit the accident data at rural, three- and four-leg, STOP-controlled intersections and urban, three-leg, STOP-controlled intersections. On the other hand, lognormal regression models were found more appropriate for modeling accidents at urban, four-leg, STOP-controlled and urban, four-leg, signalized intersections. The decision to use negative binomial or lognormal regression analysis was based on evaluation of the accident frequency distribution for the specific categories of intersections.

17. Key Words		18. Distribution Statement
Accident Modeling Traffic Accidents Geometric Design At-Grade Intersections	Poisson Regression Negative Binomial Regression Lognormal Regression	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.

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

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

[This page intentionally left blank]

TABLE OF CONTENTS

[This page intentionally left blank]

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

This technical report presents the results of statistical analyses of accident experience for at-grade intersections. This report is an addendum to the previously published FHWA Report No. FHWA-RD-96-125 titled "Statistical Models of At-Grade Intersection Accidents."[(1)] In that study, only multiple-vehicle accidents were modeled. The analyses reported here include all collision types (i.e., both multiple- and single-vehicle accidents) using 3-year accident frequencies (1990 to 1992) and geometric design, traffic control, and traffic volume data from a database provided by the California Department of Transportation (Caltrans). The data used for the analyses reported in this addendum are in all respects identical to those used for Report No. FHWA-RD-96-125, except that all collision types were included in the accident frequencies used as the dependent variable in modeling.

This report is organized similarly to Section 5 of the previous report. To minimize repetition between the two reports, the general discussion of the statistical methods has been abbreviated here (the reader is referred to the previous report for details).[(1)] Statistical modeling results for five specific types of intersections are discussed in this report:

- Rural, four-leg, STOP-controlled intersections
- Rural, three-leg, STOP-controlled intersections
- Urban, four-leg, STOP-controlled intersections
- Urban, three-leg, STOP-controlled intersections
- Urban, four-leg, signalized intersections

Section 2 of this report provides detailed accident frequency distributions for the five types of intersections. Section 3 presents the results from various statistical models that were developed with negative binomial and lognormal regressions. Section 4 presents the conclusions of the study. Detailed data on accident type and severity distribution are presented in an appendix.

[This page intentionally left blank]

2. ACCIDENT FREQUENCY DISTRIBUTIONS

This section of the report summarizes the accident statistics for all collision types for the intersections that were used in developing the models presented in this report. The purpose of these frequency distributions is to illustrate the reasons for selecting the statistical modeling approach used for each of the five intersection types. More detailed distributions that show the accident data broken down by accident type and accident severity are presented in the appendix.

Prior to beginning the statistical modeling activities, the general shape of each accident frequency distribution was assessed for each of the five intersection types of interest. This was done visually by plotting the data for the 3-year totals and by calculating basic statistics. For each of the five intersection categories, Table 1 shows yearly and 3-year total accident statistics (minimum, median, mean, maximum) for total accidents. Total accident counts, as well as fatal and injury accident counts, are shown separately. Table 1 also shows the total number of accidents of each type in each given year. For each type of intersection, 3-year totals are also shown for multiple-vehicle accidents only. The percentages of total accidents that were multiple-vehicle and single-vehicle accidents were computed for all five intersection categories for total accidents. Generally, multiple-vehicle accidents represent a large proportion of all accidents, except for those occurring at rural, three-leg, STOP-controlled intersections.

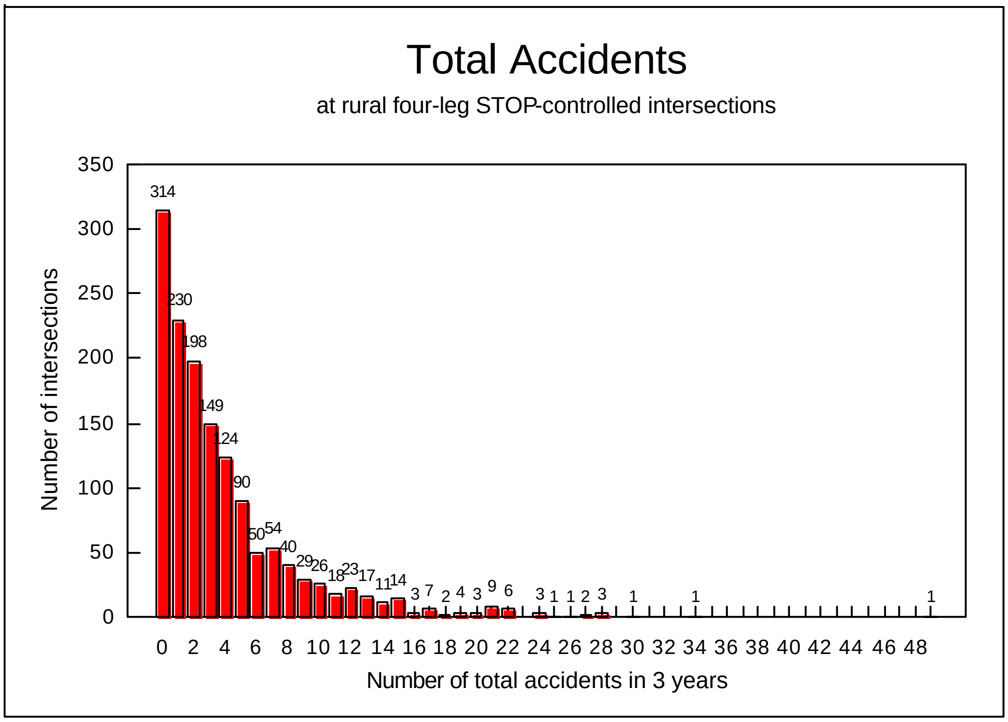
Next, the frequency data are plotted separately for each type of accident and each type of intersection in Figures 1 through 5. The plots shown in Figures 1 through 5 highlight the different shapes of accident frequencies. With large numbers of intersections with no or low accident experience, the distribution tends to follow the shape of a Poisson distribution. This observation clearly applies to rural, four- and three-leg, STOP-controlled intersections and to urban, three-leg, STOP-controlled intersections. When the number of intersections with no or low accident experience is relatively small, the distribution tends to follow the shape of a lognormal distribution. This is clearly seen in the case of urban, four-leg, signalized intersections and also in the case of urban, four-leg, STOP-controlled intersections.

Table 1. Annual Accident Statistics, All Collision Types, 1990–1992

	Total accidents					Fatal and injury accidents				
Year	Minimum	Median	Mean	Maximum	Total	Minimum	Median	Mean	Maximum	Total
Rural, Four-Leg, STOP-Controlled—1,434 Intersections										
1990	0	1	1.39	15	1,996	0	0	0.68	9	980
1991	0	1	1.26	16	1,802	0	0	0.61	10	880
1992	0	1	1.28	19	1,833	0	0	0.63	8	899
1990–92	0	2	3.93	49	5,631	0	1	1.92	27	2,759
(86 percent multiple-vehicle accidents, 14 percent single-vehicle accidents)										
Rural, Three-Leg, STOP-Controlled—2,692 Intersections										
1990	0	0	0.84	21	2,250	0	0	0.38	12	1,028
1991	0	0	0.76	15	2,053	0	0	0.34	8	928
1992	0	0	0.78	26	2,096	0	0	0.35	11	949
1990–92	0	1	2.38	59	6,399	0	0	1.08	25	2,905
(76 percent multiple-vehicle accidents, 24 percent single-vehicle accidents)										
Urban, Four-Leg, STOP-Controlled—1,342 Intersections										
1990	0	2	2.71	26	3,641	0	1	1.13	12	1,521
1991	0	2	2.35	19	3,158	0	1	1.07	9	1,434
1992	0	2	2.38	21	3,190	0	1	1.07	12	1,440
1990–92	0	5	7.44	57	9,989	0	2	3.27	24	4,395
(88 percent multiple-vehicle accidents, 12 percent single-vehicle accidents)										
Urban, Three-Leg, STOP-Controlled—3,057 Intersections										
1990	0	1	1.62	33	4,939	0	0	0.65	12	1,984
1991	0	1	1.50	31	4,571	0	0	0.64	10	1,951
1992	0	1	1.43	31	4,359	0	0	0.61	18	1,867
1990–92	0	3	4.54	81	13,869	0	1	1.90	28	5,802
(86 percent multiple-vehicle accidents, 14 percent single-vehicle accidents)										

Table 1 (Continued)

	Total accidents					Fatal and injury accidents				
Year	Minimum	Median	Mean	Maximum	Total	Minimum	Median	Mean	Maximum	Total
Urban, Four-Leg, Signalized—1,306 Intersections										
1990	0	7	8.49	49	11,087	0	3	3.32	23	4,330
1991	0	6	7.77	54	10,152	0	3	3.19	20	4,167
1992	0	6	7.17	58	9,364	0	2	3.06	20	4,002
1990–92	0	19	23.43	151	30,603	0	8	9.57	51	12,499
(93 percent multiple-vehicle accidents, 7 percent single-vehicle accidents)										



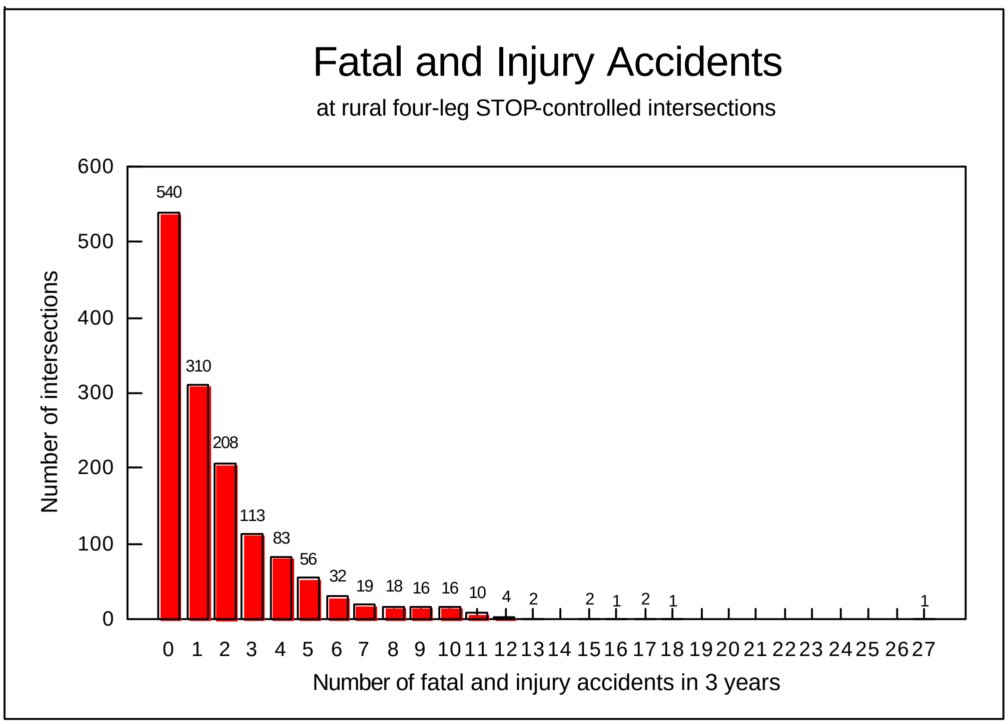
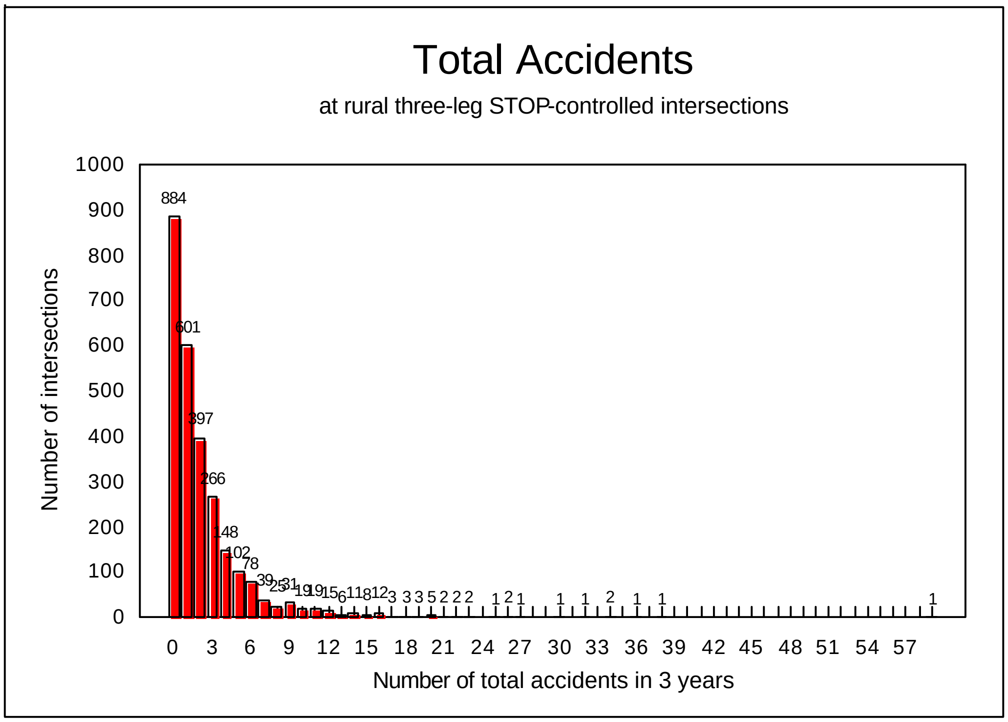


Figure 1. Accident Frequency Distributions at Rural, Four-Leg, STOP-Controlled Intersections



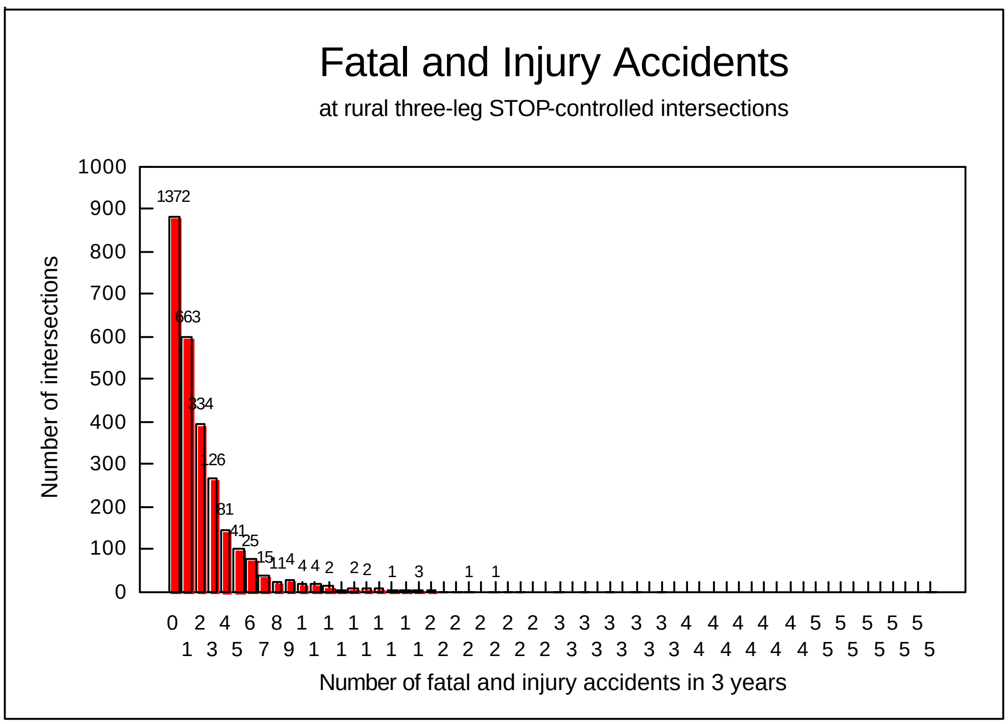
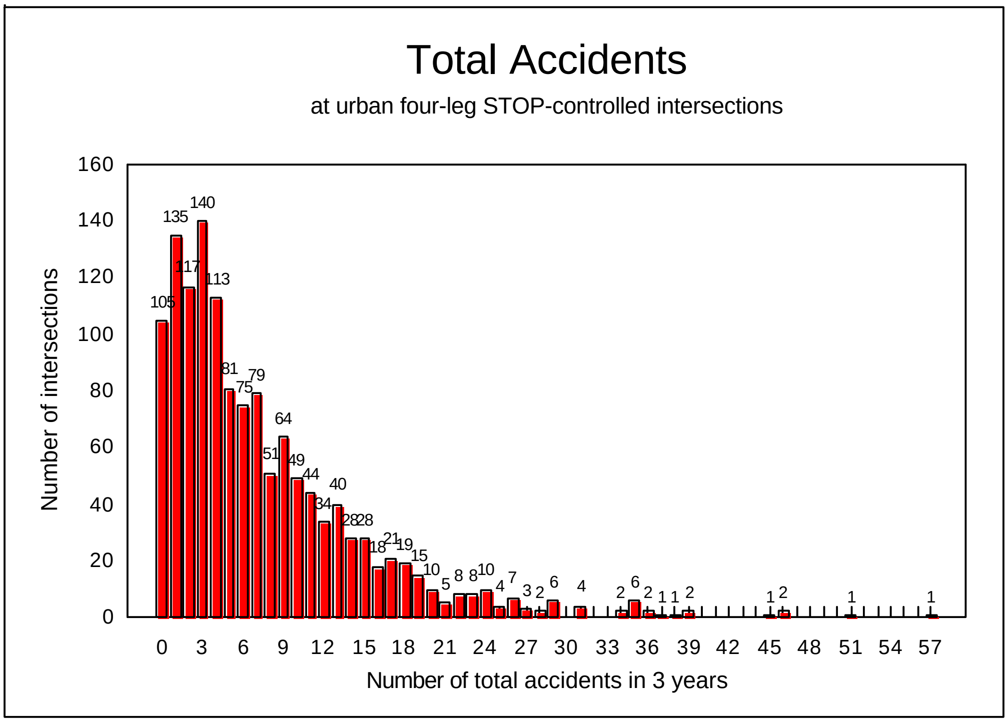


Figure 2. Accident Frequency Distributions at Rural, Three-Leg, STOP-Controlled Intersections



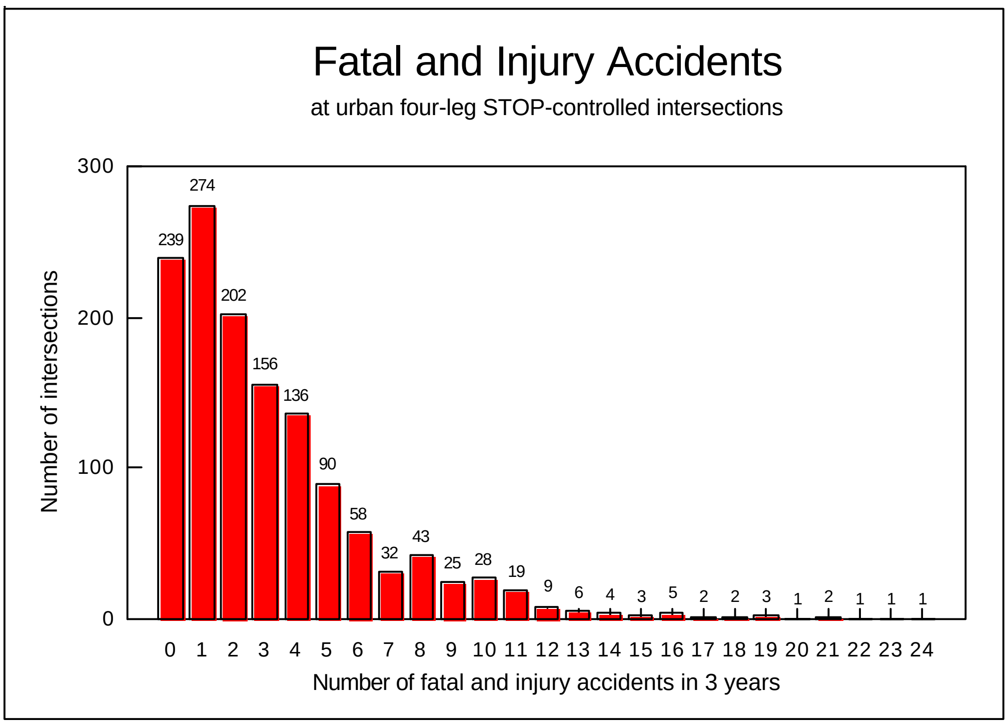
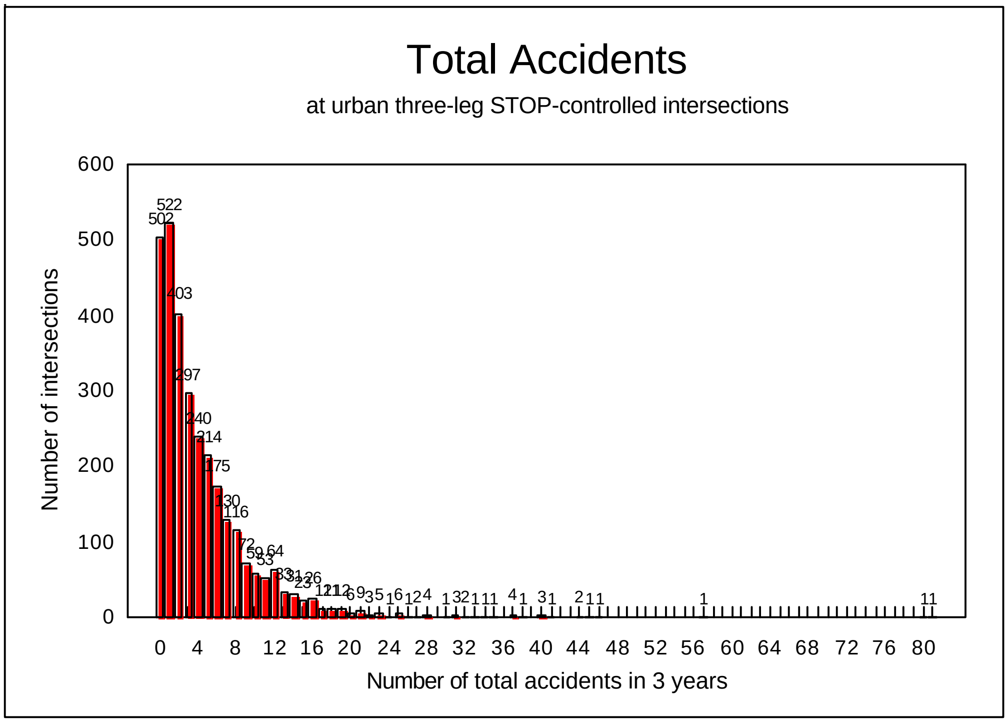


Figure 3. Accident Frequency Distributions at Urban, Four-Leg, STOP-Controlled Intersections



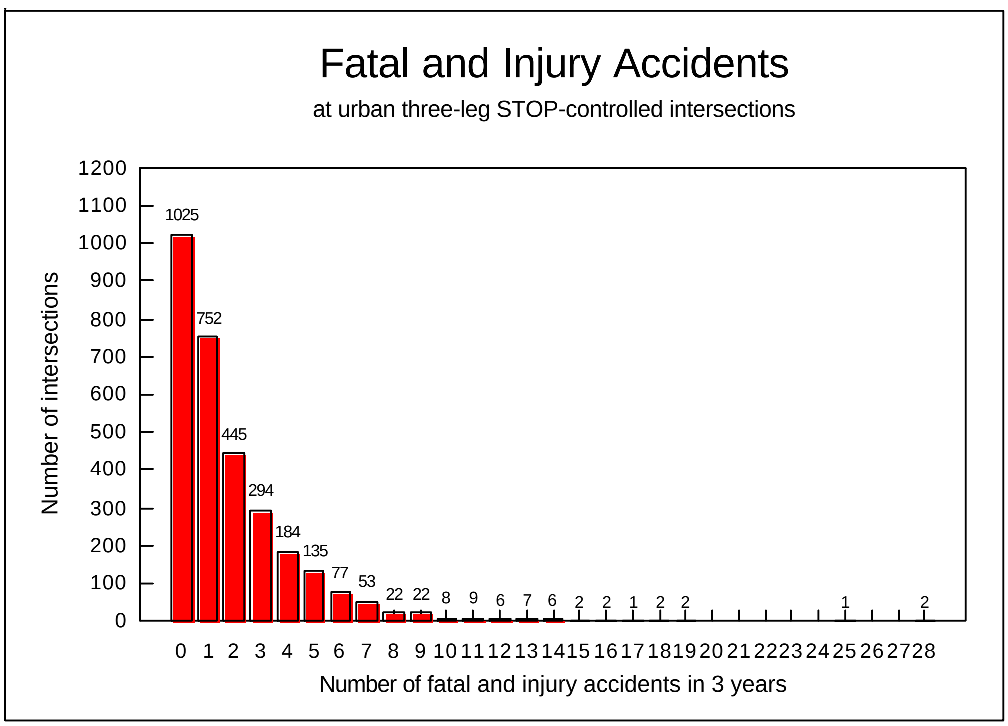
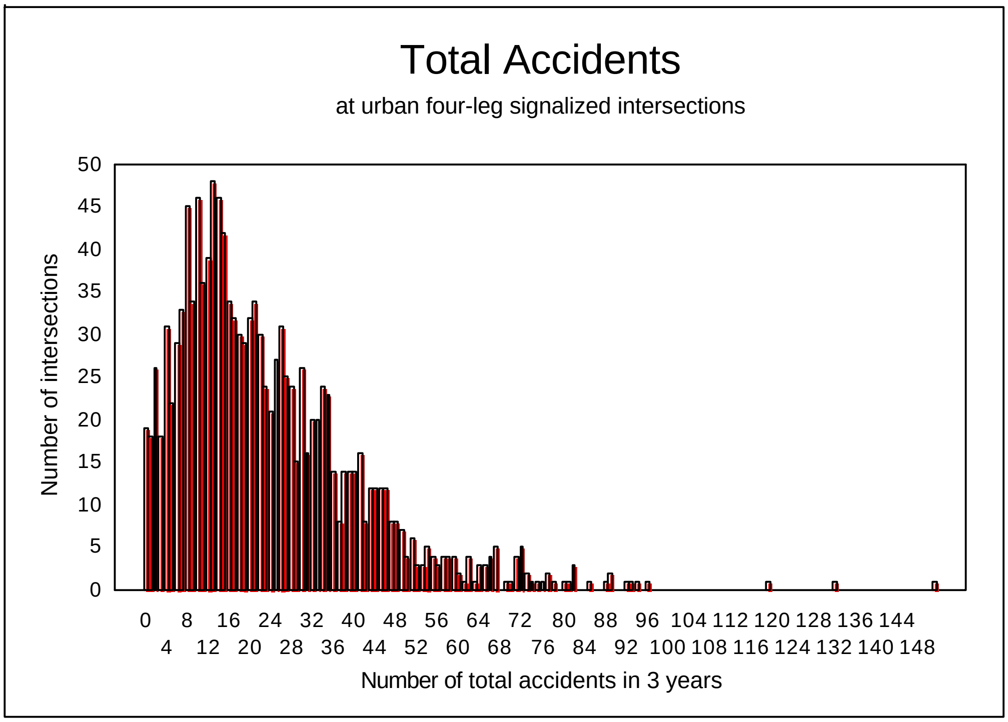


Figure 4. Accident Frequency Distributions at Urban, Three-Leg, STOP-Controlled Intersections



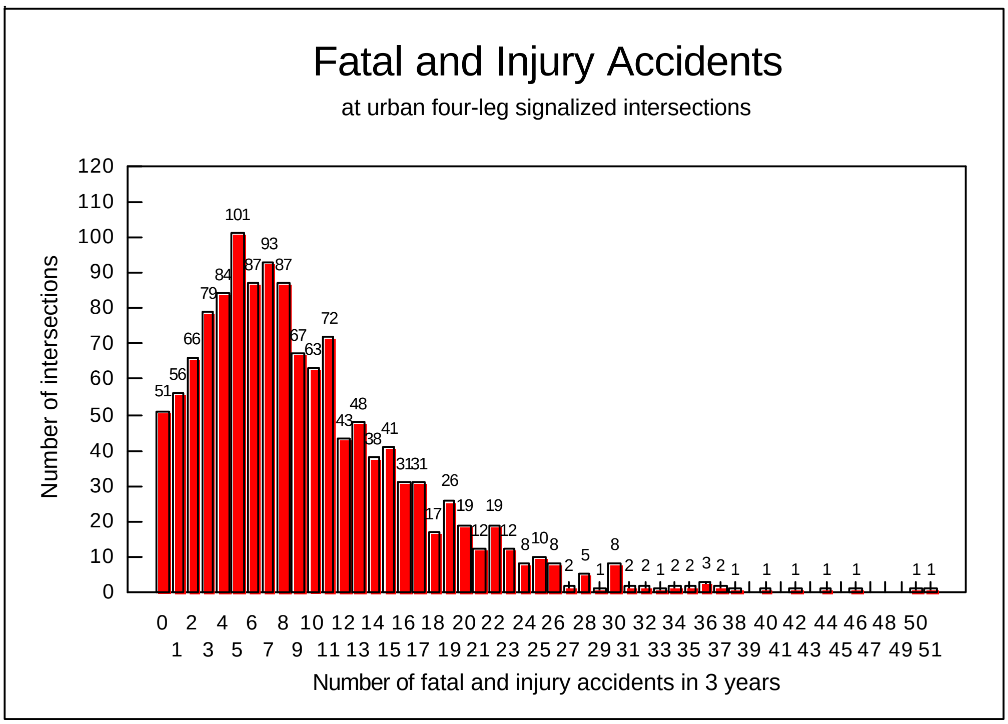


Figure 5. Accident Frequency Distributions at Urban, Four-Leg, Signalized Intersections

3. STATISTICAL MODELING

Lognormal and Loglinear Regression Models

Several candidate analysis methods were investigated for application to the accident frequencies in the five types of at-grade intersections in this study. The analysis approach was driven by both the actual distribution of accident frequencies and by recommendations and practices in the field of accident data analysis. The frequency distributions of total fatal and injury accidents in the 3-year study period are shown in Figures 1 through 5 above. The percentages of intersections with zero or one accident in the 3-year period are:

Intersection type	Percent of intersections with 0 or 1 total accident in 3-year period	Percent of intersections with 0 or 1 fatal and injury accident in 3-year period
Rural, four-leg, STOP-controlled	37.9	59.3
Rural, three-leg, STOP-controlled	55.2	75.6
Urban, four-leg, STOP-controlled	17.9	38.2
Urban, three-leg, STOP-controlled	33.5	58.1
Urban, four-leg, signalized	2.8	8.2

For most types of intersections, a large proportion of the intersections experienced at most one accident over the three-year period. This observation is not true for urban, four-leg, signalized intersections, where about 50 percent of the intersections experienced 19 or more total accidents over the 3-year period, and only 10 percent of the intersections had 5 or fewer accidents. Also, the pattern for urban, four-leg, STOP-controlled intersections differs somewhat from the two extreme situations in that only eight percent of the intersections experienced no accidents and about half of the intersections experienced at least four accidents in the three-year period. These same patterns also apply to fatal and injury accidents.

As was done in the previous study, two general types of statistical models were applied to the accident data in this study: (1) a lognormal regression model for all urban, four-leg intersections (both STOP-controlled and signalized); and (2) a loglinear regression model—a Poisson regression followed by a negative binomial regression model—for all rural STOP-controlled (three- and four-leg) and urban, three-leg, STOP-controlled intersections. These models are briefly described next.

Consider a set of n intersections of a given class (e.g., rural, four-leg, STOP-controlled intersections). Associated with each intersection i is a set of q parameters (X_{i1}, X_{i2}, ..., X_{iq}) describing the geometric design, traffic control, traffic volume, and other related characteristics of that intersection. Let the number of accidents occurring at the ith intersection during a 3-year period be denoted by Y_i, where i = 1, ..., n. Next, denote by y_i the actual observation of Y_i during the 3-year period, that is, $y_i = 0, 1, 2, ...$ and i = 1, ..., n.

The objective of a statistical model is to provide a relationship between a function of the expected number of accidents, $E(Y_i) = \mu_i$, at the ith intersection and the q intersection parameters, $X_{i1}, X_{i2}, ..., X_{iq}$. This relationship can be formulated through a general linear model of the form:

$$\text{Function}(\mu_i) = \beta_0 + \beta_1 X_{i1} + ... + \beta_q X_{iq} \qquad (1)$$

where the regression coefficients, $\beta_0, \beta_1, \beta_2, ... \beta_q$, are to be estimated from the data. The estimation procedure used to obtain the regression coefficients is dependent on the assumption made about the distribution of the Y_i.

Note: Throughout this report, all logarithms are natural logarithms and are denoted by *log* in all equations.

Lognormal Regression Models

Lognormal regression models are based on the assumption that the natural logarithm of Y_i follows a normal distribution with mean μ_i and variance σ^2. In other words, it is assumed that Y_i follows a lognormal distribution, a reasonable choice whenever the data are inherently non-negative, suggesting that a model with positive skewness is needed and the mean is relatively large. This model also ensures that μ_i, the expected number of accidents, remains positive.

In this case, the relationship between the expected number of accidents at the ith intersection and the q predictor variables, $X_1,...,X_q$, can be written as:

$$\log(\mu_i) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + ... + \beta_q X_{iq} \qquad (2)$$

or alternatively, in the multiplicative form, as

$$\mu_i = \exp(\beta_0) \exp(\beta_1 X_{i1}) \exp(\beta_2 X_{i2}) ... \exp(\beta_q X_{iq}) \qquad (3)$$

where the log(number of accidents) is assumed to follow a normal distribution with mean μ_i and variance σ^2. The coefficients, $\beta_0, \beta_1, \beta_2, ... \beta_q$, are the linear regression coefficients to be estimated by the ordinary least-squares method.

Loglinear Regression Models

As was done in the previous study, two loglinear models were considered for application to at-grade accident frequencies: the Poisson and the negative binomial models. Their general forms are described below.

Poisson Model. When the average number of accidents at an intersection becomes small, the assumption of a lognormal distribution is no longer valid. The Poisson model then becomes a natural choice as it models the occurrence of rare discrete events well. The relationship between the expected number of accidents occurring at the ith intersection and the q intersection parameters, X_{i1}, X_{i2}, ..., X_{iq}, is assumed to be of the same form as shown in Equation (2):

$$\log(\mu_i) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_q X_{iq} \quad (4)$$

However, the assumption is now made that the number of accidents, Y_i, follows a Poisson distribution with mean μ_i.

Negative Binomial Model. A limitation of the Poisson distribution is that the mean equals the variance of the distribution. Previous work in the field of accident research has shown that this is not always the case. Suppose a Poisson model is used for modeling accidents and the variance, or dispersion, of the data exceeds the estimated mean of the accident data distribution. The data are then said to be overdispersed, and the underlying assumption of the variance being equal to the mean for the Poisson distribution is violated. The negative binomial, which is a discrete distribution, provides an alternative model to deal with overdispersion in count data such as accident frequencies.

As for the Poisson model above, the relationship between the expected number of accidents occurring at the ith intersection and the q intersection parameters, X_{i1}, X_{i2}, ..., X_{iq}, is still taken to be:

$$\log(\mu_i) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_q X_{iq} \quad (5)$$

However, the assumption is now made that the number of accidents, Y_i, follows a negative binomial distribution with parameters α and k (with $0 \leq \alpha \leq 1$ and $k \geq 0$).

For both the Poisson and negative binomial models, the regression coefficients, β_0, β_1, β_2, ... β_q, are estimated by the method of maximum likelihood. The statistical analysis system (SAS) provides a procedure, PROC GENMOD (a generalized linear model procedure), that can be used to estimate the regression coefficients of Equations (4) and (5) by specifying the appropriate type of distribution.

Treatment of ADT Variables in Regression Models

In all models in this study, the natural logarithm of the major-road and crossroad average daily traffic (ADT) variables was used. This parallels the approach taken by other researchers where accident counts rather than accident rates are modeled. On the log-scale, the ratio of accident counts over ADT becomes the difference between log(accident counts) and log(ADT). The difference here is that it is assumed that the coefficient of log(ADT) is not equal to one, but rather is a coefficient to be estimated through analysis. Thus, in the lognormal and Poisson and negative binomial models described above, X_1 and X_2 generally represent $\log(ADT_{major\ road})$ and $\log(ADT_{crossroad})$, respectively. The multiplicative model relating the expected accident counts and independent variables can thus be rewritten as:

$$\mu_i = \exp(\beta_0)\left(ADT_{major\ road}\right)^{\beta 1}\left(ADT_{crossroad}\right)^{\beta 2} \exp(\beta_3 X_{i3}) \bullet \ ... \ \bullet \exp(\beta_q X_{iq}) \qquad (6)$$

Note: For further details, the reader is referred to Section 5 of the previous report, which also contains a list of statistical references.[(1)]

Accident Modeling Results

The following sections present the modeling results separately for each of the five selected types of intersections:

- Rural, four-leg, STOP-controlled
- Rural, three-leg, STOP-controlled
- Urban, four-leg, STOP-controlled
- Urban, three-leg, STOP-controlled
- Urban, four-leg, signalized

Rural, Four-Leg, STOP-Controlled Intersections

The first step in the analysis of rural, four-leg, STOP-controlled intersections was to select candidate independent variables for that particular group of intersections. Both engineering judgment and sample size requirements for the levels of each candidate variable were involved in the decision as to whether a particular variable was included in the modeling effort. A small number of independent categorical variables were not included in the full model because either all or nearly all intersections fell into one level of that variable. The variables originally considered that were not included for these reasons were:

- Major-road left-turn prohibition (all intersections had permitted left turns)
- Crossroad left-turn channelization (none of the intersections had left-turn lanes)
- Crossroad left-turn prohibition (none of the intersections had left turns prohibited)
- Number of lanes on the crossroad (99.8 percent had two lanes)

Table 2 identifies the variables that were selected for modeling of rural, four-leg, STOP-controlled intersections. This table also provides descriptive statistics for three types of variables: (1) total and fatal and injury accident frequencies in the 3-year study period (i.e., the dependent variables for the modeling effort); (2) all independent continuous variables considered; and (3) all independent categorical variables considered. Minimum, mean, median, and maximum values are given for the first two types of variables. For categorical variables, the percent of intersections within each level is given.

Next, using all the continuous and categorical variables shown in Table 2, a negative binomial regression model was fitted separately to the data for total and fatal and injury accidents. The analyses were performed using the SAS GENMOD procedure with the negative binomial distribution and the appropriate deviance functions and variance adjustment factor, k.[1,2,3] The significance of each regression coefficient was examined. If a coefficient was not significant at the 10 percent level, the corresponding variable was deleted from the model, and the negative binomial regression was rerun. In addition, while it was not considered appropriate to include independent variables with significance levels above 10 percent in the models presented in this report, the text of the report identifies those independent variables that were found to have significance levels between 10 percent and 20 percent. This significance level, α, is indicated for each such variable.

Table 2. Descriptive Statistics for Rural, Four-Leg, STOP-Controlled Intersections

Parameter	Level	Percent of intersections	Minimum	Mean	Median	Maximum
Total accidents; 1990 through 1992 combined			0	3.9	2	49
Fatal and injury accidents; 1990 through 1992 combined			0	1.9	1	27
Major-road ADT (veh/day)			400	8,262	6,646	72,000
Crossroad ADT (veh/day)			100	630	351	9,585
Design speed of major road (mi/h)			25	56	60	70
Outside shoulder width on major road (ft)			0	6.7	8	15
Average lane width on major road (ft)			8	12.0	12	15
Terrain	Flat	64				
	Rolling	26				
	Mountainous	11				
Functional class of major road	Principal arterial	27				
	Minor arterial	61				
	Major collector	12				
Lighting	No	62				
	Yes	39				
Major-road left-turn channelization	No left-turn lane	63				
	Painted left-turn lane	32				
	Curbed left-turn lane	4.5				
Major-road right-turn channelization	No free right turns	90				
	Provision for free right turns	10				
Number of lanes on major road	3 or less	83				
	4 or more	17				
Crossroad right-turn channelization	No free right turns	95				
	Provision for free right turns	4.8				
Presence of median on major road	Divided	19				
	Undivided	81				
Access control on major road	None	81				
	Partial	19				

1,434 intersections

Conversion: 1 km/h = 0.621 mi/h; 1 m = 3.28 ft

The final model diagnostics for the reduced negative binomial regression model are shown in Table 3. The model diagnostics, which are shown separately for each type of accident considered, include both basic statistics and goodness-of-fit criteria. The following model statistics are shown:

Model Statistic	Explanation
Basic Statistics	
Number of intersections, n	Total sample size in that category of intersections
Number of parameters in model	Total number of independent variables, both categorical and continuous
Parameters degrees of freedom, p	Each continuous independent variable has one degree of freedom; the number of degrees of freedom associated with each categorical variable equals the number of levels minus one. The intercept has one degree of freedom. The sum of these degrees of freedom is denoted as p.
k factor	Only applicable to the negative binomial distribution. The use of this factor results in a ratio of the deviance to its degrees of freedom of approximately one.
Criteria for Assessing Goodness of Fit	
Deviance/(n − p)	The deviance of the model containing all the parameters (including the intercept) divided by its degrees of freedom, n − p. This statistic (mean deviance) provides a test for overdispersion and a measure of fit of the model. Asymptotically, this value tends toward one.
Pearson chi-square/(n − p)	The Pearson chi-square statistic divided by its degrees of freedom, n − p. This statistic provides another measure of fit of the model. Asymptotically, this value tends toward one. This statistic is referred to as the Pearson chi-square ratio in subsequent sections.
R^2	A goodness-of-fit parameter based on the ordinary multiple correlation coefficient.
R^2_{FT}	A goodness-of-fit parameter based on the Freeman-Tukey variance stabilizing transformation of variables discussed in Fridstrøm, et al.[(4)]

Two goodness-of-fit measures, the mean deviance and the Pearson chi-square ratio (the Pearson chi-square value divided by its degrees of freedom), were used to assess the fit of the model. Generally, if the Pearson chi-square ratio is between 0.8 and 1.2, this is an indication that the model can be assumed to be appropriate in modeling the data.

Table 3. Model Diagnostics for Total and Fatal and Injury Accidents at Rural, Four-Leg, STOP-Controlled Intersections

	Negative binomial regression (reduced model)
Total Accidents (3-year period)	
Number of intersections (n)	1,434
Total number of parameters considered	14
Number of parameters in reduced model	9
Parameters degrees of freedom[a] (p)	12
k factor	0.65
Deviance/(n - p)	1.00
Pearson chi-square/(n - p)	0.95
R^2 (percent)	37.26
R^2_{FT} (percent)	39.32
Fatal and Injury Accidents (3-year period)	
Number of intersections (n)	1,434
Total number of parameters considered	14
Number of parameters in reduced model	7
Parameters degrees of freedom[a] (p)	10
k factor	0.70
Deviance/(n - p)	0.99
Pearson chi-square/(n - p)	1.00
R^2 (percent)	31.14
R^2_{FT} (percent)	30.47

[a]Includes one degree of freedom for the intercept.

Of the 14 original independent variables considered, only 9 remain statistically significant in the final negative binomial model for total accidents. A variance stabilizing factor, k, of 0.65 was needed to achieve a mean deviance of approximately one, indicating that the data are neither overdispersed nor underdispersed relative to the model. The Pearson chi-square ratio equals approximately 0.95, a value within the acceptable range of 0.8 to 1.2. These two goodness-of-fit results indicate that the choice of the negative binomial model appears appropriate. The two additional measures of goodness of fit, R^2 and R^2_{FT}, are approximately 37 percent and 39 percent, respectively, for total accidents.

Of the five independent variables considered in the full, but not in the final, reduced negative binomial model, only one variable—major-road left-turn channelization (α = 0.16)—was not significant at the 10 percent level, but would have been significant at the 20 percent level.

The use of the negative binomial model had a similar impact on the results for fatal and injury accidents. Of the 14 original independent variables considered, only 7 remain statistically significant in the final negative binomial model for fatal and injury accidents. A variance stabilizing factor, k, of 0.70 was needed to achieve a mean deviance of approximately one. The Pearson chi-square ratio equals approximately one, a value within the acceptable range of 0.8 to 1.2. Again, the two goodness-of-fit results provide an indication that the choice of the negative binomial model appears appropriate. The two additional measures of goodness of fit, R^2 and R^2_{FT}, are approximately 31 percent and 30 percent, respectively, for fatal and injury accidents.

Of the seven independent variables considered in the full, but not in the final, reduced negative binomial model, only two variables—access control on major road ($\alpha = 0.12$) and major-road left-turn channelization ($\alpha = 0.12$)—were not significant at 10 percent, but would have been significant at the 20 percent level.

Tables 4 and 5 summarize the regression results for the final negative binomial model for total accidents and fatal and injury accidents, respectively. Each table identifies the:

- Statistically significant variables remaining in the final model
- Chi-square statistic for each remaining variable; all of these chi-square statistics are statistically significant at the 10 percent significance level or better
- Levels of each statistically significant categorical variable
- Direction of the effect if the effect was inverse to the expected direction
- Value of the regression coefficient for each continuous variable or each level of each categorical variable in the model
- Relative effect of a unit change in each variable on the expected accident frequency in a 3-year period (this is simply e^{β}, where β is the coefficient given in the table)
- Lower and upper 90 percent confidence limits of the regression coefficient

In each table, the independent variables are listed in decreasing order of their ability to explain the variations in intersection accident frequencies as indicated by the chi-square value, which represents the strength of the relationship of each variable to accident frequency, taking into account all other variables in the model.

Table 4. Negative Binomial Regression Results for Total Accidents at Rural, Four-Leg, STOP-Controlled Intersections

Independent variable[a]	Chi square statistic[b]	Variable level	Direction of effect[c]	Coefficient	Relative effect[d]	90 percent confidence limits[e]	
						Lower	Upper
Intercept				−10.025		−10.846	−9.211
Crossroad ADT (log)	361.52		—	0.532	1.70	0.484	0.579
Major-road ADT (log)	347.76		—	0.758	2.13	0.689	0.827
Number of lanes on major road	15.25	3 or less	—	0.321	1.38	0.186	0.455
		4 or more	—	0			
Design speed on major road	11.19			0.009	1.01	0.005	0.014
Access control on major road	5.78	None	—	0.200	1.22	0.063	0.336
		Partial	—	0			
Functional class of major road	7.26	Principal arterial	—	0			
		Minor arterial	—	0.181	1.20	0.070	0.292
		Major collector	—	0.173	1.19	0.0005	0.347
Lighting	3.68	No		0.122	1.13	0.017	0.226
		Yes		0			
Terrain	4.81	Flat	I	0.053	1.05	0.055	0.160
		Rolling	I	0			
		Mountainous	I	−0.159	0.85	−0.328	0.011
Major-road right-turn channelization	2.78	No free right-turn lane	—	0.157	1.17	0.002	0.310
		Provision for free right-turns	—	0			

Note: This analysis is based on the set of 1,434 intersections for which summary statistics are shown in Table 2.

[a] All variables significant at the 90 percent confidence level or higher

[b] Chi-square likelihood ratio statistic for testing the significance of the effect of the variable; with 1 degree of freedom for continuous variables; with (p-1) degrees of freedom for categorical variables with p levels

[c] Direction of effect: I = Inverse of expected direction

[d] Relative effect of unit change in the variable on the expected number of accidents, equals exp(coefficient)

[e] 90 percent lower and upper confidence limits of the estimated coefficient

Table 5. Negative Binomial Regression Results for Fatal and Injury Accidents at Rural, Four-Leg, STOP-Controlled Intersections

Independent variable[a]	Chi square statistic[b]	Variable level	Direction of effect[c]	Coefficient	Relative effect[d]	90 percent confidence limits[e]	
						Lower	Upper
Intercept				−10.294		−11.244	−9.356
Crossroad ADT (log)	282.17		—	0.546	1.73	0.491	0.602
Major-road ADT (log)	201.92		—	0.680	1.97	0.599	0.762
Number of lanes on major road	18.17	3 or less	—	0.385	1.47	0.237	0.533
		4 or more	—	0			
Design speed on major road	16.60		—	0.013	1.01	0.008	0.019
Terrain	15.98	Flat	I	0.183	1.20	0.060	0.306
		Rolling	I	0			
		Mountainous	I	−0.234	0.79	−0.442	−0.026
Functional class of major road	11.87	Principal arterial	—	0			
		Minor arterial	—	0.261	1.30	0.136	0.386
		Major collector	—	0.170	1.18	-0.030	0.370
Lighting	9.03	No	—	0.219	1.24	0.099	0.339
		Yes	—	0			

Note: This analysis is based on the set of 1,434 intersections for which summary statistics are shown in Table 2.

[a] All variables significant at the 90 percent confidence level or higher

[b] Chi-square likelihood ratio statistic for testing the significance of the effect of the variable; with 1 degree of freedom for continuous variables; with (p-1) degrees of freedom for categorical variables with p levels

[c] Direction of effect: I = Inverse of expected direction

[d] Relative effect of unit change in the variable on the expected number of accidents, equals exp(coefficient)

[e] 90 percent lower and upper confidence limits of the estimated coefficient

To predict the average accident frequency at rural, four-leg, STOP-controlled intersections, one replaces the regression coefficients, β_0, β_1, β_2, ... β_q, with the estimated values found in the table, and the variables X_1, X_2, ..., X_q, with their appropriate values or levels. For example, the expected 3-year total accident frequency can be estimated using the model presented in Table 4 as:

$$Y = e^{-10.025}\ (X_1)^{0.532}\ (X_2)^{0.758}\ \exp(0.321X_3)\ \exp(0.009X_4)\ \exp(0.200X_5)\ \exp(0.181X_6)\ \exp(0.173X_7)\ \exp(0.122X_8)\ \exp(-0.053X_9)\ \exp(0.159X_{10})\ \exp(0.157X_{11}) \quad (7)$$

where:

Y = expected number of total accidents in a 3-year period
X_1 = ADT of the crossroad (veh/day)
X_2 = ADT of the major road (veh/day)
X_3 = 1 if the major road has 3 or fewer lanes in both direction of travel combined; 0 if 4 or more
X_4 = design speed on major road (mi/h)
X_5 = 1 if the major road has no access control; 0 if access control is partial
X_6 = 1 if the major road is a minor arterial; 0 otherwise
X_7 = 1 if the major road is a major collector; 0 otherwise
X_8 = 1 if the intersection is lighted; 0 otherwise
X_9 = 1 if terrain is flat; 0 otherwise
X_{10} = 1 if terrain is mountainous; 0 otherwise
X_{11} = 1 if no right-turn lane is present on the major road; 0 otherwise

Note that when the level of a categorical variable is 0, the multiplicative term in Equation (7) becomes $e^0 = 1$ and is therefore omitted from the model.

The relative effect of each variable, all other variables being held constant, can be calculated by simply taking the exponent of the corresponding coefficient. For example, the relative effect of having a major road with three lanes or fewer as opposed to four lanes or more is exp(0.321) = 1.38. In other words, decreasing the number of lanes on the major road from 4 or more to 3 or less, with all other factors being held constant, would increase the expected number of accidents by a factor of 1.38 or by 38 percent. Similarly, intersections on major roads without access control were found to have 22 percent more accidents than intersections on major roads with partial access control.

The results of the negative binomial regression modeling shown in Tables 4 and 5 show that most of the variables of interest have effects in the direction expected. However, the observed effect of terrain is not in the expected direction since the results imply that intersections in flat terrain have more accidents than intersections in rolling terrain which, in turn, have more accidents than intersections in mountainous terrain. Such effects that are opposite the direction expected can represent situations in which a variable, for which data are available, is correlated with and serves as a surrogate for another variable for which data are not available.

Figure 6 illustrates the influence of the major-road and crossroad ADT on the annual number of accidents at rural, four-leg, STOP-controlled intersections with the typical geometrics identified in the box below the figure. Each curve in the figure represents combinations of major-road and crossroad ADT that would be expected to result in a specific annual number of accidents, ranging from 0.5 to 5 accidents per year.

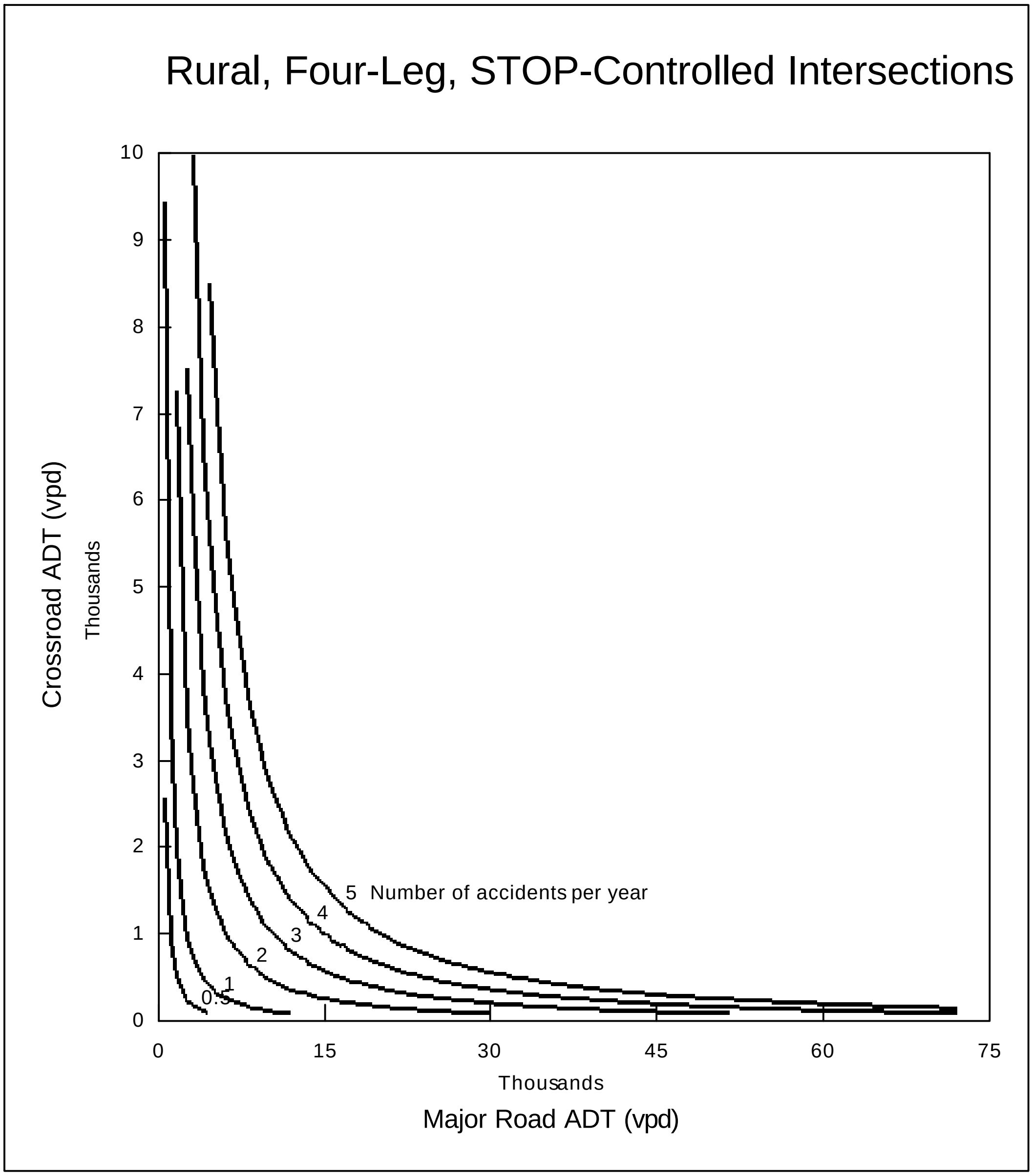
Rural, Three-Leg, STOP-Controlled Intersections

The statistical analysis approach used for rural, three-leg, STOP-controlled intersections was identical to that used for rural, four-leg, STOP-controlled intersections. The median number of total accidents at any single intersection was 1 accident in the 3-year study period, with a maximum of 59 accidents in the 3-year period. As shown in Figure 2, approximately 55 percent of all 2,692 rural, three-leg, STOP-controlled intersections in the study experienced either no accidents or 1 accident in the 3-year period. Thus, the negative binomial model appeared to be a logical choice for analysis of this data set.

The selection of independent variables was done in a fashion similar to that described earlier in Section 3. Table 6 identifies the variables that were selected for modeling accidents at rural, three-leg, STOP-controlled intersections. As before, a small number of independent categorical variables were not included in the full model because either all or nearly all intersections fell into one level of that variable. The variables originally considered that were not included for this reason were:

- Major-road left-turn prohibition (no intersections had left turns prohibited)
- Crossroad left-turn channelization (none of the intersections had left-turn lanes)
- Crossroad left-turn prohibition (none of the intersections had left turns prohibited)
- Number of lanes on the crossroad (99.7 percent had two lanes)

As shown in the table, 14 independent variables, both continuous and categorical, were considered in the full negative binomial regression model. Of these 14 variables, 7 were found to have a statistically significant effect on accidents (both total and fatal and injury multiple-vehicle) at the 10 percent significance level. A reduced negative binomial model was then rerun using only the seven statistically significant variables.



Major road characteristics:	**Other:**
‣ 3 lanes or less ‣ Design speed of 89 km/h (55 mi/h) ‣ No access control ‣ Minor arterial ‣ No free right turns	‣ No lighting ‣ Flat terrain

Figure 6. Number of Accidents per Year as a Function of Traffic Volumes for Typical Rural, Four-Leg, STOP-Controlled Intersections

Table 6. Descriptive Statistics for Rural, Three-Leg, STOP-Controlled Intersections

Parameter	Level	Percent of intersections	Minimum	Mean	Median	Maximum
Total accidents: 1990 through 1992 combined			0	2.4	1	59
Fatal and injury accidents: 1990 through 1992 combined			0	1.1	0	25
Major-road ADT (veh/day)			400	8,288	6,138	72,000
Crossroad ADT (veh/day)			100	487	210	10,001
Design speed of major road (mi/h)			25	53	55	70
Outside shoulder width on major road (ft)			0	5.7	6	15
Average lane width on major road (ft)			8	11.8	12	15
Terrain	Flat	40				
	Rolling	35				
	Mountainous	25				
Functional class of major road	Principal arterial	20		**2,692 intersections**		
	Minor arterial	68				
	Major collector	12				
Lighting	No	70				
	Yes	30				
Major-road left-turn channelization	No left-turn lane	68				
	Painted left-turn lane	30				
	Curbed left-turn lane	2.1				
Major-road right-turn channelization	No free right turns	92				
	Provision for free right turns	7.7				
Number of lanes on major road	3 or less	88				
	4 or more	12				
Crossroad right-turn channelization	No free right turns	97				
	Provision for free right turns	2.6				
Presence of median on major road	Divided	16				
	Undivided	84				
Access control on major road	None	91				
	Partial	9.4				

Conversion: 1 km/h = 0.621 mi/h; 1 m = 3.28 ft

A final, reduced negative binomial regression model was used with a variance adjustment factor, k, of 0.70 for total and 0.55 for fatal and injury accidents, respectively. The model diagnostics are summarized in Table 7. In each case, this approach resulted in a mean deviance of approximately one, and a Pearson chi-square ratio of 1.07 for total accidents and 1.15 for fatal and injury accidents. Of the original 14 independent variables considered for modeling, only 7 remained statistically significant at the 90 percent confidence level for total accidents. A slightly different set of 7 variables remained statistically significant at the 90 percent confidence level for fatal and injury accidents.

Table 7. Model Diagnostics for Total and Fatal and Injury Accidents at Rural, Three-Leg, STOP-Controlled Intersections

	Negative binomial regression (reduced model)
Total Accidents (3-year period)	
Number of intersections (n)	2,692
Total number of parameters considered	14
Number of parameters in reduced model	7
Parameters degrees of freedom[a] (p)	11
k factor	0.70
Deviance/(n - p)	1.01
Pearson chi-square/(n - p)	1.07
R^2 (percent)	34.00
R^2_{FT} (percent)	34.09
Fatal and Injury Accidents (3-year period)	
Number of intersections (n)	2,692
Total number of parameters considered	14
Number of parameters in reduced model	7
Parameters degrees of freedom[a] (p)	10
k factor	0.55
Deviance/(n - p)	1.00
Pearson chi-square/(n - p)	1.15
R^2 (percent)	26.46
R^2_{FT} (percent)	23.83

[a]Includes one degree of freedom for the intercept.

Of the 7 independent variables considered in the full, but not in the final, reduced negative binomial model for total accidents, only 3 variables—average lane width on major road (α = 0.15), lighting (α = 0.18), and number of lanes on major road (α = 0.20)—were not significant at the 10 percent level, but would have been at the 20 percent level. Of the 7 independent variables considered in the full, but not in the final, reduced negative binomial model for fatal and injury accidents, 4 variables—crossroad right-turn channelization (α =

0.11), design speed ($\alpha = 0.12$), average lane width or major road ($\alpha = 0.15$), and terrain ($\alpha = 0.18$)—were not significant at the 10 percent level, but would have been at the 20 percent level.

Tables 8 and 9 summarize the regression results for the final negative binomial model for total accidents and fatal and injury accidents, respectively. The tables show that for rural, three-leg, STOP-controlled intersections, all of the independent variables evaluated in the total accident models had effects in the direction expected. The variable for crossroad right-turn channelization in the model for fatal and injury accidents has an effect opposite the direction expected since the model implies that an intersection without free right-turn channelization would experience fewer accidents than an intersection with free right-turn channelization.

Figure 7 illustrates the variations of the annual number of accidents with major-road and crossroad ADT for rural, three-leg, STOP-controlled intersections of the type specified at the upper right of the figure.

Urban, Four-Leg, STOP-Controlled Intersections

The statistical analysis approach used for urban, four-leg, STOP-controlled intersections was at first identical to that used for rural, four- and three-leg, STOP-controlled intersections. The median number of total accidents at any intersection was 5 accidents in the 3-year study period with a maximum of 57 accidents in the 3-year period. As shown in Figure 3, only approximately 8 percent of all 1,342 urban, four-leg, STOP-controlled intersections in the study experienced no accidents in the 3-year period.

The selection of independent variables was done in a fashion similar to that described earlier. Table 10 identifies the variables that were selected for modeling accidents at urban, four-leg, STOP-controlled intersections. As before, a small number of independent categorical variables were not included in the full model because either all or nearly all intersections fell into one level of that variable. The variables originally considered that were not included for this reason were:

- Crossroad left-turn prohibition (only 4 percent of the intersections had left turns prohibited; in addition, this variable showed a high negative correlation of −0.73 with the equivalent variable on the major road)
- Number of lanes on the crossroad (99.9 percent had 2 lanes)

Table 8. Negative Binomial Regression Results for Total Accidents at Rural, Three-Leg, STOP-Controlled Intersections

Independent variable[a]	Chi square statistic[b]	Variable level	Direction of effect[c]	Coefficient	Relative effect[d]	90 percent confidence limits[e]	
						Lower	Upper
Intercept				−9.178		−9.746	−8.616
Major-road ADT (log)	772.08		—	0.830	2.29	0.778	0.882
Crossroad road ADT (log)	301.33		—	0.383	1.47	0.345	0.420
Major-road left-turn channelization	15.73	No left-turn lane	—	0.213	1.24	0.124	0.301
		Painted left-turn lane	—	0			
		Curbed left-turn lane	—	0.124	1.13	−0.120	0.374
Access control on major road	7.31	None	—	0.225	1.25	0.089	0.362
		Partial	—	0			
Functional class of major road	7.27	Principal arterial	—	0			
		Minor arterial	—	0.145	1.16	0.048	0.242
		Major collector	—	0.211	1.23	0.059	0.363
Outside shoulder width on major road	4.63			−0.017	0.98	−0.030	−0.004
Terrain	5.32	Flat	—	−0.045	0.96	−0.130	0.041
		Rolling	—	0			
		Mountainous	—	0.095	1.10	−0.004	0.195

Note: This analysis is based on the set of 2,692 intersections for which summary statistics are shown in Table 6. [a] All variables significant at the 90 percent confidence level or higher

[b] Chi-square likelihood ratio statistic for testing the significance of the effect of the variable; with 1 degree of freedom for continuous variables; with (p-1) degrees of freedom for categorical variables with p levels

[c] Direction of effect: I = Inverse of expected direction

[d] Relative effect of unit change in the variable on the expected number of accidents, equals exp(coefficient)

[e] 90 percent lower and upper confidence limits of the estimated coefficient

Table 9. Negative Binomial Regression Results for Fatal and Injury Accidents at Rural, Three-Leg, STOP-Controlled Intersections

Independent variable[a]	Chi square statistic[b]	Variable level	Direction of effect[c]	Coefficient	Relative effect[d]	90 percent confidence limits[e]	
						Lower	Upper
Intercept				−9.141		−9.841	−8.447
Major-road ADT (log)	517.56		—	0.781	2.18	0.722	0.841
Crossroad ADT (log)	217.64		—	0.384	1.47	0.341	0.428
Outside shoulder width on major road	11.61		—	−0.030	0.97	−0.044	−0.015
Lighting	9.35	No	—	0.169	1.18	0.078	0.261
		Yes	—	0			
Major-road left-turn channelization	8.89	No left-turn lane	—	0.180	1.20	0.080	0.279
		Painted left-turn lane	—	0			
		Curbed left-turn lane	—	0.062	1.06	−0.194	0.318
Functional class of major road	6.90	Principal arterial	—	0			
		Minor arterial	—	0.164	1.18	0.058	0.270
		Major collector	—	0.192	1.21	0.021	0.362
Crossroad right-turn channelization	2.74	No free right turns	I	−0.219	0.80	−0.437	−0.001
		Provision for free right turns	I	0			

Note: This analysis is based on the set of 2,692 intersections for which summary statistics are shown in Table 6.

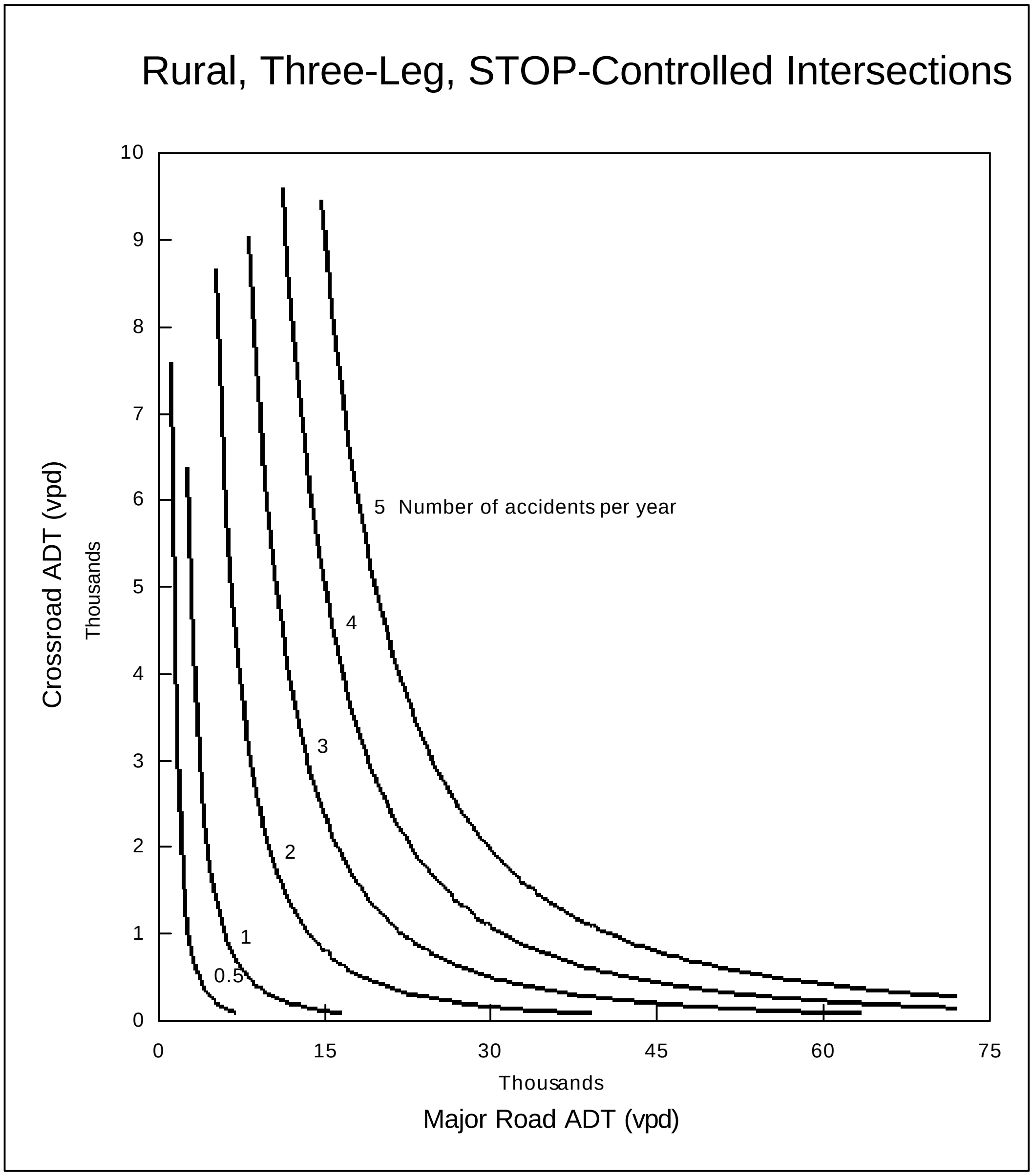
[a] All variables significant at the 90 percent confidence level or higher

[b] Chi-square likelihood ratio statistic for testing the significance of the effect of the variable; with 1 degree of freedom for continuous variables; with (p-1) degrees of freedom for categorical variables with p levels

[c] Direction of effect: I = Inverse of expected direction

[d] Relative effect of unit change in the variable on the expected number of accidents, equals exp(coefficient)

[e] 90 percent lower and upper confidence limits of the estimated coefficient



Major road characteristics:	**Other:**
‣ No left-turn lane ‣ No access control ‣ Minor arterial ‣ Outside shoulder width of 1.8 m (6 ft)	‣ Flat terrain

Figure 7. Number of Accidents per Year as a Function of Traffic Volumes for Typical Rural, Three-Leg, STOP-Controlled Intersections

Table 10. Descriptive Statistics for Urban, Four-Leg, STOP-Controlled Intersections

Parameter	Level	Percent of intersections	Mininum	Mean	Median	Maximum
Total accidents; 1990 through 1992 combined			0	7.4	5	57
Fatal and injury accidents; 1990 through 1992 combined			0	3.3	2	24
Major-road ADT (veh/day)			1,100	23,240	21,217	79,000
Crossroad ADT (veh/day)			100	1,255	900	16,940
Design speed of major road (mi/h)			25	50	50	70
Outside shoulder width on major road (ft)			0	7.2	8	15
Average lane width on major road (ft)			8	12.1	12	15
Terrain	Flat	75				
	Rolling or mountainous	25				
Functional class of major road	Principal arterial	91		**1,342 intersections**		
	Minor arterial	7.7				
	Major collector	1.6				
Lighting	No	15				
	Yes	85				
Major-road left-turn channelization	No left-turn lane	44				
	Painted left-turn lane	40				
	Curbed left-turn lane	16				
Major-road right-turn channelization	No free right turns	96				
	Provision for free right turns	4.2				
Major-road left-turn prohibition	Left turns permitted	97				
	Left turns prohibited	3.1				
Number of lanes on major road	3 or less	31				
	4 or 5	61				
	6 or more	7.9				
Crossroad left-turn channelization	No left-turn lane	98				
	Painted left-turn lane	2.0				
Crossroad right-turn channelization	No free right turns	97				
	Provision for free right turns	3.3				
Presence of median on major road	Divided	55				
	Undivided	45				
Access control on major road	None	96				
	Partial	4.0				

As is shown in Table 10, 16 independent variables, both continuous and categorical, were considered in the full regression model. Based on the shape of the accident data distributions (see Figure 3), a lognormal rather than a negative binomial regression model was used to model accidents at this type of intersection. The natural logarithm of the accident counts was modeled using the full set of 16 independent variables. All modeling was performed using the SAS stepwise regression procedure.

The model diagnostics are summarized in Table 11. In this case, the root mean squared error (RMSE) has been added as a measure of fit of the model to the data. This statistic provides an estimate of the standard deviation of the error term (on the log scale).

Table 11. Model Diagnostics for Total and Fatal and Injury Accidents at Urban, Four-Leg, STOP-Controlled Intersections

	Lognormal Regression (reduced model)
Total Accidents (3-year period)	
Number of intersections (n)	1,342
Total number of parameters considered	16
Number of parameters in reduced model	9
Parameters degrees of freedom[a] (p)	12
Deviance/(n - p)	0.96
Pearson chi-square/(n - p)	0.96
R^2 (percent)	21.12
R^2_{FT} (percent)	na
Root mean square error	0.96
Fatal and Injury Accidents (3-year period)	
Number of intersections (n)	1,342
Total number of parameters considered	16
Number of parameters in reduced model	9
Parameters degrees of freedom[a] (p)	12
Deviance/(n - p)	0.80
Pearson chi-square/(n - p)	0.80
R^2 (percent)	18.70
R^2_{FT} (percent)	na
Root mean square error	0.80

[a] Includes one degree of freedom for the intercept

The mean deviance, the Pearson chi-square ratio, and the RMSE are all identical, with values of 0.96 and 0.80 for total and fatal and injury accidents, respectively. The R^2 values are approximately 21 percent for total accidents and approximately 19 percent for fatal and injury accidents. These measures of fit, however, are relatively poor compared to those obtained for the previous types of intersections.

Of the original 16 independent variables considered for modeling, only 9 remained statistically significant at the 90 percent confidence level for total accidents. A slightly different set of 9 variables remained statistically significant at the 90 percent confidence level for fatal and injury accidents.

Of the 7 independent variables considered in the full, but not in the final, reduced lognormal model for total accidents, only 2 variables—outside shoulder width ($\alpha = 0.13$) and terrain ($\alpha = 0.19$)—were not significant at the 10 percent level, but would have been at the 20 percent level. Of the 7 independent variables considered in the full, but not in the final, reduced lognormal model for fatal and injury accidents, only 1 variable—crossroad left-turn channelization ($\alpha = 0.15$)—was not significant at the 10 percent level, but would have been at the 20 percent level.

Tables 12 and 13 summarize the regression results for the final lognormal model for total accidents and fatal and injury accidents, respectively. The tables indicate that no statistically significant effect on accidents was found for major-road left-turn channelization. The average lane width on the major road was found to have an effect on intersection accidents in the expected direction (i.e., for each decrease of 0.3 m (1 ft) in lane width on the major-road approaches, intersection-related accidents increased by 9.1 percent). Three of the variables evaluated had effects that were inverse to the direction expected: access control on the major road, crossroad right-turn channelization, and intersection lighting.

Figure 8 illustrates the variation of the annual number of intersection accidents with major-road and crossroad ADT for urban, four-leg, STOP-controlled intersections with the typical conditions specified in the figure.

Urban, Three-Leg, STOP-Controlled Intersections

The statistical analysis approach used for urban, three-leg, STOP-controlled intersections was identical to that used for rural, four- and three-leg, STOP-controlled intersections. The median number of total accidents at any 1 intersection was 3 accidents in the 3-year study period with a maximum of 81 accidents in the 3-year period. As shown in Figure 4, approximately 47 percent of all 3,057 urban, three-leg, STOP-controlled intersections in the study experienced 2 or fewer accidents in the 3-year period. Thus, the negative binomial model appeared to be a logical choice for analysis of this data set.

The selection of independent variables was done in a fashion similar to that described in earlier sections. Table 14 identifies the variables that were selected for modeling accidents at urban, three-leg, STOP-controlled intersections. None of the variables considered for modeling was deleted due to small sample sizes. Although the percentage of intersections in some levels is relatively small (e.g, 0.7 percent of intersections had 4 or more lanes), the large number of intersections (3,057) in this category justified the inclusion of these variables and their levels in the analysis.

Table 12. Lognormal Regression Results for Total Accidents at Urban, Four-Leg, STOP-Controlled Intersections

Independent variable[a]	Chi square statistic[b]	Variable level	Direction of effect[c]	Coefficient	Relative effect[d]	90 percent confidence limits[e]	
						Lower	Upper
Intercept				−4.664		−5.783	−3.545
Major-road ADT (log)	132.18		—	0.620	1.86	0.532	0.709
Crossroad ADT (log)	88.36		—	0.281	1.32	0.232	0.330
Major-road left-turn prohibition	31.64	Left turns prohibited	—	−0.941	0.39	−1.216	−0.666
		Left turns permitted	—	0			
Average lane width on major road	16.28		—	−0.097	0.91	−0.136	−0.057
Number of lanes on major road	14.99	3 or less	—	0.401	1.49	0.179	0.623
		4 or 5	—	0.120	1.13	−0.064	0.304
		6 or more	—	0			
Access control on major road	8.49	None	I	−0.437	0.65	−0.683	−0.190
		Partial	I	0			
Crossroad right-turn channelization	6.01	No free right turns	I	−0.384	0.68	−0.641	−0.126
		Provision for free right turns	I	0			
Lighting	3.48	No	I	−0.160	0.85	−0.302	−0.019
		Yes	I	0			
Functional class of major road	2.95	Principal arterial	—	0			
		Minor arterial	—	−0.153	0.86	−0.332	0.025
		Major collector	—	−0.229	0.80	−0.593	0.135

Note: This analysis is based on the set of 1,342 intersections for which summary statistics are shown in Table 10.

[a] All variables significant at the 90 percent confidence level or higher

[b] Chi-square likelihood ratio statistic for testing the significance of the effect of the variable; with 1 degree of freedom for continuous variables; with (p-1) degrees of freedom for categorical variables with p levels

[c] Direction of effect: I = Inverse of expected direction

[d] Relative effect of unit change in the variable on the expected number of accidents, equals exp(coefficient)

[e] 90 percent lower and upper confidence limits of the estimated coefficient

Table 13. Lognormal Regression Results for Fatal and Injury Accidents at Urban, Four-Leg, STOP-Controlled Intersections

Independent variable[a]	Chi square statistic[b]	Variable level	Direction of effect[c]	Coefficient	Relative effect[d]	90 percent confidence limits[e]	
						Lower	Upper
Intercept				−4.693		−5.790	−3.597
Major-road ADT (log)	117.44		—	0.584	1.79	0.495	0.672
Crossroad ADT (log)	48.45		—	0.206	1.23	0.157	0.255
Major-road left-turn prohibition	19.95	Left turns prohibited	—	−0.747	0.47	−1.022	−0.472
		Left turns permitted	—	0			
Average lane width on major road	11.36		—	−0.081	0.92	−0.121	−0.042
Access control on major road	7.31	None	I	−0.382	0.68	−0.615	−0.150
		Partial	I	0			
Number of lanes on major road	9.88	3 or less	—	0.282	1.33	0.059	0.504
		4 or 5	—	0.049	1.05	−0.137	0.234
		6 or more	—	0			
Outside shoulder width on major road	5.94		—	−0.020	0.98	−0.033	−0.006
Crossroad right-turn channelization	3.66	No free right turns	I	−0.300	0.74	−0.557	−0.042
		Provision for free right turns	I	0			
Functional class of major road	3.71	Principal arterial	—	0			
		Minor arterial	—	−0.079	0.92	−0.257	0.100
		Major collector	—	−0.401	0.67	−0.766	−0.037

Note: This analysis is based on the set of 1,342 intersections for which summary statistics are shown in Table 10.

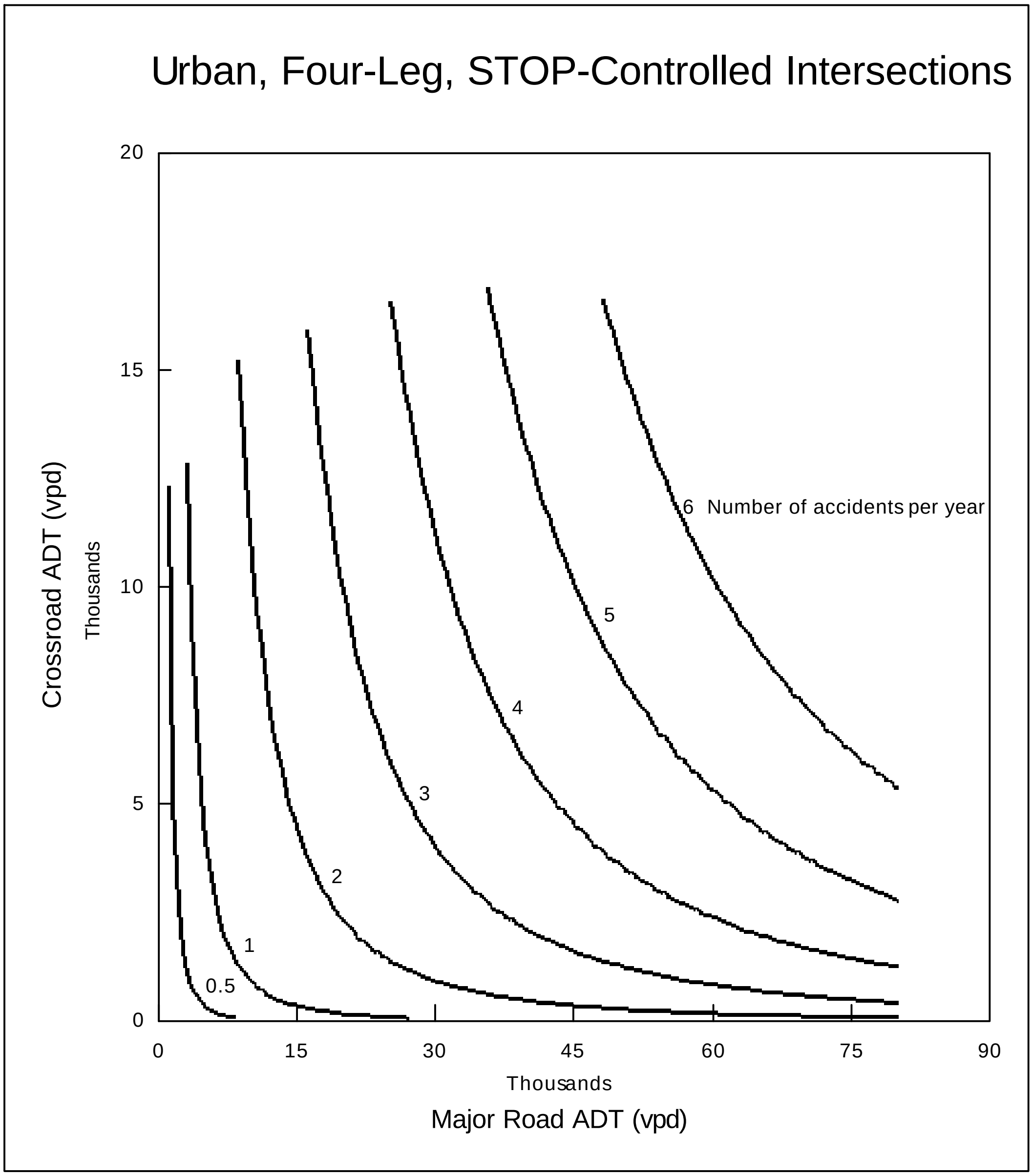
[a] All variables significant at the 90 percent confidence level or higher

[b] Chi-square likelihood ratio statistic for testing the significance of the effect of the variable; with 1 degree of freedom for continuous variables; with (p-1) degrees of freedom for categorical variables with p levels

[c] Direction of effect: I = Inverse of expected direction

[d] Relative effect of unit change in the variable on the expected number of accidents, equals exp(coefficient)

[e] 90 percent lower and upper confidence limits of the estimated coefficient



Major road characteristics:	Crossroad characteristics:	Other:
▸ Left turns permitted ▸ Average lane width of 3.6 m (12 ft) ▸ 4 or 5 lanes ▸ No access control ▸ Principal arterial	▸ No free right turns	▸ Lighting

Figure 8. Number of Accidents per Year as a Function of Traffic Volumes for Typical Urban, Four-Leg, STOP-Controlled Intersections

Table 14. Descriptive Statistics for Urban, Three-Leg, STOP-Controlled Intersections

Parameter	Level	Percent of intersections	Minimum	Mean	Median	Maximum
Total accidents; 1990 through 1992 combined			0	4.5	3	81
Fatal and injury accidents; 1990 through 1992 combined			0	1.9	1	28
Major-road ADT (veh/day)			520	25,557	23,400	97,000
Crossroad ADT (veh/day)			100	808	501	21,800
Design speed of major road (mi/h)			25	50	50	70
Outside shoulder width on major road (ft)			0	7.0	8	15
Average lane width on major road (ft)			8	12.0	12	15
Terrain	Flat	72				
	Rolling or mountainous	29				
Functional class of major road	Principal arterial	90		**3,057 intersections**		
	Minor arterial	10				
Lighting	No	23				
	Yes	77				
Major-road left-turn channelization	No left-turn lane	50				
	Painted left-turn lane	36				
	Curbed left-turn lane	14				
Major-road right-turn channelization	No free right turns	97				
	Provision for free right turns	3.5				
Major-road left-turn prohibition	Left turns permitted	87				
	Left turns prohibited	13				
Number of lanes on major road	3 or less	32				
	4 or 5	56				
	6 or more	12				
Crossroad left-turn channelization	No left-turn lane	99				
	Painted or curbed left-turn lane	1.4				
Crossroad right-turn channelization	No free right turns	97				
	Provision for free right turns	3.0				
Crossroad left-turn prohibition	Left turns permitted	86				
	Left turns prohibited	13				
Number of lanes on crossroad	3 or less	99				
	4 or more	0.7				
Presence of median on major road	Divided	61				
	Undivided	39				
Access control on major road	None	98				
	Partial	2.4				

As shown in the table, 18 independent variables, both continuous and categorical, were considered in the full negative binomial regression model. Of these 18 variables, 8 were found to have a statistically significant effect on either total or fatal and injury accidents at the 10 percent significance level. A negative binomial regression model was then used with a variance adjustment factor, k, of 0.86 for total and 0.79 for fatal and injury accidents. The final model diagnostics are summarized in Table 15. In each case, this approach resulted in a mean deviance of approximately 1, and a Pearson chi-square ratio of 1.08 for total accidents and 1.00 for fatal and injury accidents. For both types of accidents, the models produced relatively poor results based on the two R^2 values, all in the range of 16 to 18 percent.

Table 15. Model Diagnostics for Total and Fatal and Injury Accidents at Urban, Three-Leg, STOP-Controlled Intersections

	Negative binomial regression (reduced model)
Total Accidents (3-year period)	
Number of intersections (n)	3,057
Total number of parameters considered	18
Number of parameters in reduced model	8
Parameters degrees of freedom[a] (p)	10
k factor	0.86
Deviance/(n - p)	1.00
Pearson chi-square/(n - p)	1.08
R^2 (percent)	16.66
R^2_{FT} (percent)	17.67
Fatal and Injury Accidents (3-year period)	
Number of intersections (n)	3,057
Total number of parameters considered	18
Number of parameters in reduced model	8
Parameters degrees of freedom[a] (p)	10
k factor	0.79
Deviance/(n - p)	1.00
Pearson chi-square/(n - p)	1.00
R^2 (percent)	16.21
R^2_{FT} (percent)	15.66

[a] Includes one degree of freedom for the intercept.

Of the original 18 independent variables considered for modeling, only 8 remained statistically significant at the 90 percent confidence level for total accidents. A slightly different set of 8 variables remained statistically significant at the 90 percent confidence level for fatal and injury accidents.

Of the 10 independent variables considered in the full, but not in the final, reduced negative binomial model for total accidents, none that was not significant at the 10 percent level would have been significant at the 20 percent level. Of the 10 independent variables considered in the full, but not in the final, reduced negative binomial model for fatal and injury accidents, only 3 variables were not significant at the 10 percent level, but would have been at the 20 percent level. These variables were lighting ($\alpha = 0.12$); number of lanes on crossroad ($\alpha = 0.19$); and access control on major road ($\alpha = 0.20$).

Tables 16 and 17 summarize the regression results for the final negative binomial model for total accidents and fatal and injury accidents, respectively. The tables show an effect in the expected direction for the presence of a median on the major road; intersections on divided highways appear to have 14 percent fewer accidents than intersections on undivided highways. A concern with the models developed is that the effect on safety of major-road left-turn and right-turn channelization is opposite the direction expected.

Figure 9 shows the variation of the annual number of intersection accidents with major-road and crossroad ADT for urban, three-leg, STOP-controlled intersections with the typical conditions specified in the figure.

Urban, Four-Leg, Signalized Intersections

Accident frequencies at urban, four-leg, signalized intersections are shown in Figure 5, for both total and fatal and injury accidents. A clear departure from a Poisson distribution is visible in these distribution plots. Only 19 out of 1,306 intersections (or 1.5 percent) experienced no accidents in the 3-year study period. Approximately half of all intersections in this category experienced 19 accidents or more in the 3-year period, with a maximum of 151 total accidents. Given these high accident frequencies and the shape of the distribution for both types of accidents, a lognormal regression model presented a logical choice. Thus, the statistical analysis approach used for urban, four-leg, signalized intersections was identical to that used for urban, four-leg, STOP-controlled intersections.

The selection of independent variables was done in a fashion similar to that described earlier in Section 3. Table 18 identifies the variables that were selected for modeling accidents at urban, four-leg, signalized intersections. As before, a small number of independent categorical variables was not included in the full model because either all or nearly all intersections fell into one level of that variable. The variables originally considered that were not included for this reason were:

- Lighting (all intersections were lighted)
- Presence of major-road signal mast arm (a mast arm was present on all intersections)
- Major-road left-turn prohibition (no intersections had left turns prohibited)

Table 16. Negative Binomial Regression Results for Total Accidents at Urban, Three-Leg, STOP-Controlled Intersections

Independent variable[a]	Chi square statistic[b]	Variable level	Direction of effect[c]	Coefficient	Relative effect[d]	90 percent confidence limits[e]	
						Lower	Upper
Intercept				−5.557		−6.281	−4.834
Major-road ADT (log)	325.80		—	0.683	1.98	0.620	0.745
Crossroad ADT (log)	147.31		—	0.245	1.28	0.211	0.278
Crossroad right-turn channelization	29.28	No free right turns	I	−0.559	0.57	−0.742	−0.383
		Provision for free right turns	I	0			
Major-road left-turn prohibition	27.51	Left turns prohibited	—	−0.402	0.67	−0.528	−0.276
		Left turns permitted	—	0			
Major-road left-turn channelization	12.60	No left-turn lane	—	0.019	1.02	−0.075	0.113
		Painted left-turn lane	—	0			
		Curbed left-turn lane	I	0.210	1.23	0.112	0.310
Design speed of major road	8.15		—	−0.006	0.99	−0.009	−0.002
Presence of median on major road	6.10	Divided	—	−0.147	0.86	−0.245	−0.049
		Undivided	—	0			
Average lane width on major road	4.60		—	−0.037	0.96	−0.066	−0.009

Note: This analysis is based on the set of 3,057 intersections for which summary statistics are shown in Table 14.

[a] All variables significant at the 90 percent confidence level or higher

[b] Chi-square likelihood ratio statistic for testing the significance of the effect of the variable; with 1 degree of freedom for continuous variables; with (p-1) degrees of freedom for categorical variables with p levels

[c] Direction of effect: I = Inverse of expected direction

[d] Relative effect of unit change in the variable on the expected number of accidents, equals exp(coefficient)

[e] 90 percent lower and upper confidence limits of the estimated coefficient

Table 17. Negative Binomial Regression Results for Fatal and Injury Accidents at Urban, Three-Leg, STOP-Controlled Intersections

Independent variable[a]	Chi square statistic[b]	Variable level	Direction of effect[c]	Coefficient	Relative effect[d]	90 percent confidence limits[e]	
						Lower	Upper
Intercept				−6.618		−7.449	−5.792
Major-road ADT (log)	257.57		—	0.696	2.01	0.623	0.769
Crossroad ADT (log)	112.32		—	0.238	1.27	0.201	0.275
Crossroad right-turn channelization	26.95	No free right turns	I	−0.581	0.56	−0.773	−0.394
		Provision for free right turns	I	0			
Major-road left-turn prohibition	21.19	Left turns prohibited	—	−0.393	0.68	−0.533	−0.253
		Left turns permitted	—	0			
Major-road left-turn channelization	13.06	No left-turn lane	I	−0.057	0.94	−0.161	0.047
		Painted left-turn lane	I	0			
		Curbed left-turn lane	I	0.209	1.23	0.103	0.316
Presence of median on major road	7.53	Divided	—	−0.182	0.83	−0.292	−0.073
		Undivided	—	0			
Average lane width on major road	6.01		—	−0.048	0.95	−0.080	−0.016
Lighting	3.01	No	—	0.094	1.10	0.005	0.184
		Yes	—	0			

Note: This analysis is based on the set of 3,057 intersections for which summary statistics are shown in Table 14.

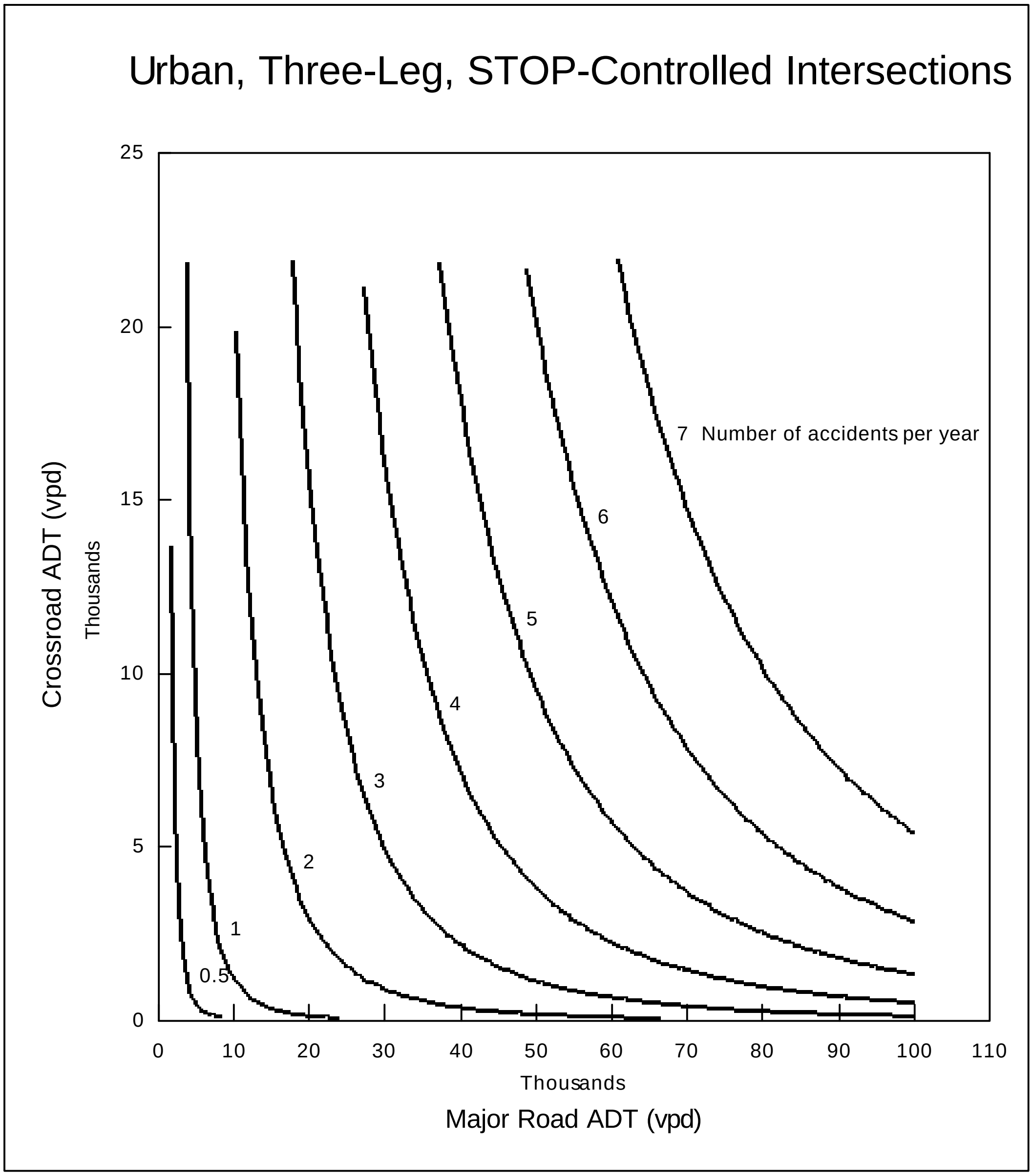
[a] All variables significant at the 90 percent confidence level or higher

[b] Chi-square likelihood ratio statistic for testing the significance of the effect of the variable; with 1 degree of freedom for continuous variables; with (p-1) degrees of freedom for categorical variables with p levels

[c] Direction of effect: I = Inverse of expected direction

[d] Relative effect of unit change in the variable on the expected number of accidents, equals exp(coefficient)

[e] 90 percent lower and upper confidence limits of the estimated coefficient



Major road characteristics:	Crossroad characteristics:
‣ Left turns permitted ‣ No left-turn lane ‣ Design speed of 64 km/h (40 mi/h) ‣ Divided ‣ Average lane width of 3.6 m (12 ft)	‣ No free right turns

Figure 9. Number of Accidents per Year as a Function of Traffic Volumes for Typical Urban, Three-Leg, STOP-Controlled Intersections

Table 18. Descriptive Statistics for Urban, Four-Leg, Signalized Intersections

Parameter	Level	Percent of intersections	Minimum	Mean	Median	Maximum
Total accidents; 1990 through 1992 combined			0	23.4	19	151
Fatal and injury accidents; 1990 through 1992 combined			0	9.6	8	51
Major-road ADT (veh/day)			2,400	31,995	31,000	79,000
Crossroad ADT (veh/day)			101	8,061	5,501	48,000
Design speed of major road (mi/h)			25	51	50	70
Outside shoulder width on major road (ft)			0	7.0	8	15
Average lane width on major road (ft)			8	12.0	12	15
Terrain	Flat	80				
	Rolling or mountainous	20				
Functional class of major road	Principal arterial	96		**1,306 intersections**		
	Minor arterial	4.1				
Signal timing	Pretimed	2.5				
	Semiactuated	13				
	Fully actuated	85				
Signal phasing	Two-phase	21				
	Multiphase	79				
Major-road left-turn channelization	No left-turn lane	4.7				
	Painted left-turn lane	40				
	Curbed left-turn lane	56				
Major-road right-turn channelization	No free right turns	74				
	Provision for free right turns	26				
Number of lanes on major road	3 or less	7.9				
	4 or 5	72				
	6 or more	20				
Presence of crossroad signal mast arm	Mast arm not present	27				
	Mast arm present	73				

Table 18. Descriptive Statistics for Urban, Four-Leg, Signalized Intersections (Continued)

Parameter	Level	Percent of intersections	Minimum	Mean	Median	Maximum
Crossroad left-turn channelization	No left-turn lane	45				
	Painted left-turn lane	36				
	Curbed left-turn lane	19				
Crossroad right-turn channelization	No free right turns	70				
	Provision for free right turns	30				
Crossroad left-turn prohibition	Left turns prohibited	0.5				
	Left turns permitted	97				
Number of lanes on crossroad	3 or less	59				
	4 or 5	38				
	6 or more	3.0				
Presence of median on major road	Divided	83				
	Undivided	18				
Access control on major road	None	94				
	Partial	5.9				

The natural logarithm of the accident counts was modeled using the full set of 19 independent variables listed in Table 18. All modeling was performed using the SAS stepwise regression procedure.

Of the original 19 independent variables considered for modeling, only 9 remained statistically significant at the 90 percent confidence level for total accidents. A slightly different set of 8 variables remained statistically significant at the 90 percent confidence level for fatal and injury accidents. The lognormal model was rerun using only the statistically significant variables to obtain the regression coefficients, their 90 percent confidence intervals, and other regression statistics.

Table 19 presents the model statistics for the final, reduced lognormal model. Although the lognormal models are significant at the 90 percent confidence level for both types of accidents, the percent variance explained by the model is relatively low, with an R^2 value of approximately 25 percent for total accidents and approximately 24 percent for fatal and injury accidents. Also, the Pearson chi-square ratios (0.69 and 0.68, respectively, for both types of accidents) are below the 0.8 to 1.2 range, indicating that the lognormal model might not provide the best fit.

Of the 10 independent variables considered in the full, but not in the final, reduced lognormal model for total accidents, only 1 variable—presence of crossroad signal mast arm ($\alpha = 0.11$)—was not significant at the 10 percent level, but would have been at the 20 percent level. Of the 11 independent variables considered in the full, but not in the final, reduced lognormal model for fatal and injury accidents, only 1 variable—average lane width on major road ($\alpha = 0.15$)—was not significant at the 10 percent level, but would have been at the 20 percent level.

Tables 20 and 21 summarize the regression results for the final lognormal model for total accidents and fatal and injury accidents, respectively. No statistically significant effect on accidents was found for either the major-road or crossroad left-turn channelization variable. However, it should be noted that major-road left-turn channelization could not be evaluated effectively because only 5 percent of the intersections had no left-turn lanes on the major road approaches. The only variables that appeared to have effects in the direction opposite to that expected were access control on the major road (for both models) and major-road right-turn channelization (for the total accident model only).

Table 19. Model Diagnostics for Total and Fatal and Injury Accidents at Urban, Four-Leg, Signalized Intersections

	Lognormal regression (reduced model)
Total Accidents (3-year period)	
Number of intersections (n)	1,306
Total number of parameters considered	19
Number of parameters in reduced model	9
Parameters degrees of freedom[a] (p) in final model	12
Deviance/(n - p)	0.69
Pearson chi-square/(n - p)	0.69
R^2 (percent)	24.77
R^2_{FT} (percent)	na
Root mean squared error	0.69
Fatal and Injury Accidents (3-year period)	
Number of intersections (n)	1,306
Total number of parameters considered	19
Number of parameters in reduced model	8
Parameters degrees of freedom[a] (p) in final model	11
Deviance/(n - p)	0.68
Pearson chi-square/(n - p)	0.68
R^2 (percent)	24.40
R^2_{FT} (percent)	na
Root mean squared error	0.68

[a] Includes one degree of freedom for the intercept.

Figure 10 shows the variation of the annual number of intersection accidents with major-road and crossroad ADT for urban, four-leg, signalized intersections with the typical conditions specified in the figure. As shown in the figure, typical accident experience at these intersections, in the range of data for which model predictions appear valid, extends up to 12 accidents per year.

Table 20. Lognormal Regression Results for Total Accidents at Urban, Four-Leg, Signalized Intersections

Independent variable[a]	Chi square statistic[b]	Variable level	Direction of effect[c]	Coefficient	Relative effect[d]	90 percent confidence limits[e]	
						Lower	Upper
Intercept				−3.428		−4.871	−1.985
Crossroad ADT (log)	57.08		—	0.224	1.25	0.175	0.273
Signal timing	56.06	Pretimed	—	0.063	1.06	−0.259	0.384
		Semiactuated	—	0			
		Fully actuated	—	0.622	1.86	0.475	0.770
Major-road ADT (log)	50.31		—	0.503	1.65	0.387	0.620
Signal phasing	7.13	Two-phase	—	0			
		Multiphase	—	−0.200	0.82	−0.324	−0.077
Access control on major road	6.64	None	I	−0.310	0.73	−0.508	−0.112
		Partial	I	0			
Number of lanes on crossroad	3.92	3 or less	—	−0.130	0.88	−0.238	−0.022
		4 or more	—	0			
Average lane width on major road	3.46		—	−0.053	0.95	−0.100	−0.006
Major-road right-turn channelization	3.01	No free right turns	I	−0.115	0.89	−0.225	−0.006
		Provision for free right turns	I	0			
Number of lanes on major road	3.44	3 or less	—	−0.225	0.80	−0.457	−0.006
		4 or 5	—	−0.130	0.88	−0.257	−0.003
		6 or more	—	0			

Note: This analysis is based on the set of 1,306 intersections for which summary statistics are shown in Table 18.
[a] All variables significant at the 90 percent confidence level or higher
[b] Chi-square likelihood ratio statistic for testing the significance of the effect of the variable; with 1 degree of freedom for continuous variables; with (p-1) degrees of freedom for categorical variables with p levels
[c] Direction of effect: I = Inverse of expected direction
[d] Relative effect of unit change in the variable on the expected number of accidents, equals exp(coefficient)
[e] 90 percent lower and upper confidence limits of the estimated coefficient

Table 21. Lognormal Regression Results for Fatal and Injury Accidents at Urban, Four-Leg, Signalized Intersections

Independent variable[a]	Chi square statistic[b]	Variable level	Direction of effect[c]	Coefficient	Relative effect[d]	90 percent confidence limits[e]	
						Lower	Upper
Intercept				−5.745		−7.100	−4.389
Major-road ADT (log)	65.12		—	0.574	1.78	0.457	0.691
Crossroad ADT (log)	53.13		—	0.215	1.24	0.167	0.264
Signal timing	25.67	Pretimed	—	−0.051	0.95	−0.373	0.271
		Semiactuated	—	0			
		Fully actuated	—	0.400	1.49	0.253	0.547
Signal phasing	10.40	Two-phase	—	0			
		Multiphase	—	−0.240	0.79	−0.363	−0.118
Access control on major road	5.78	None	I	−0.290	0.75	−0.488	−0.092
		Partial	I	0			
Number of lanes on crossroad	5.54	3 or less	—	−0.155	0.86	−0.262	−0.047
		4 or more	—	0			
Number of lanes on major road	3.94	3 or less	—	−0.163	0.85	−0.392	0.067
		4 or 5	—	−0.151	0.86	−0.277	−0.026
		6 or more	—	0			
Design speed on major road	3.33		—	0.005	1.01	0.001	0.010

Note: This analysis is based on the set of 1,306 intersections for which summary statistics are shown in Table 18.

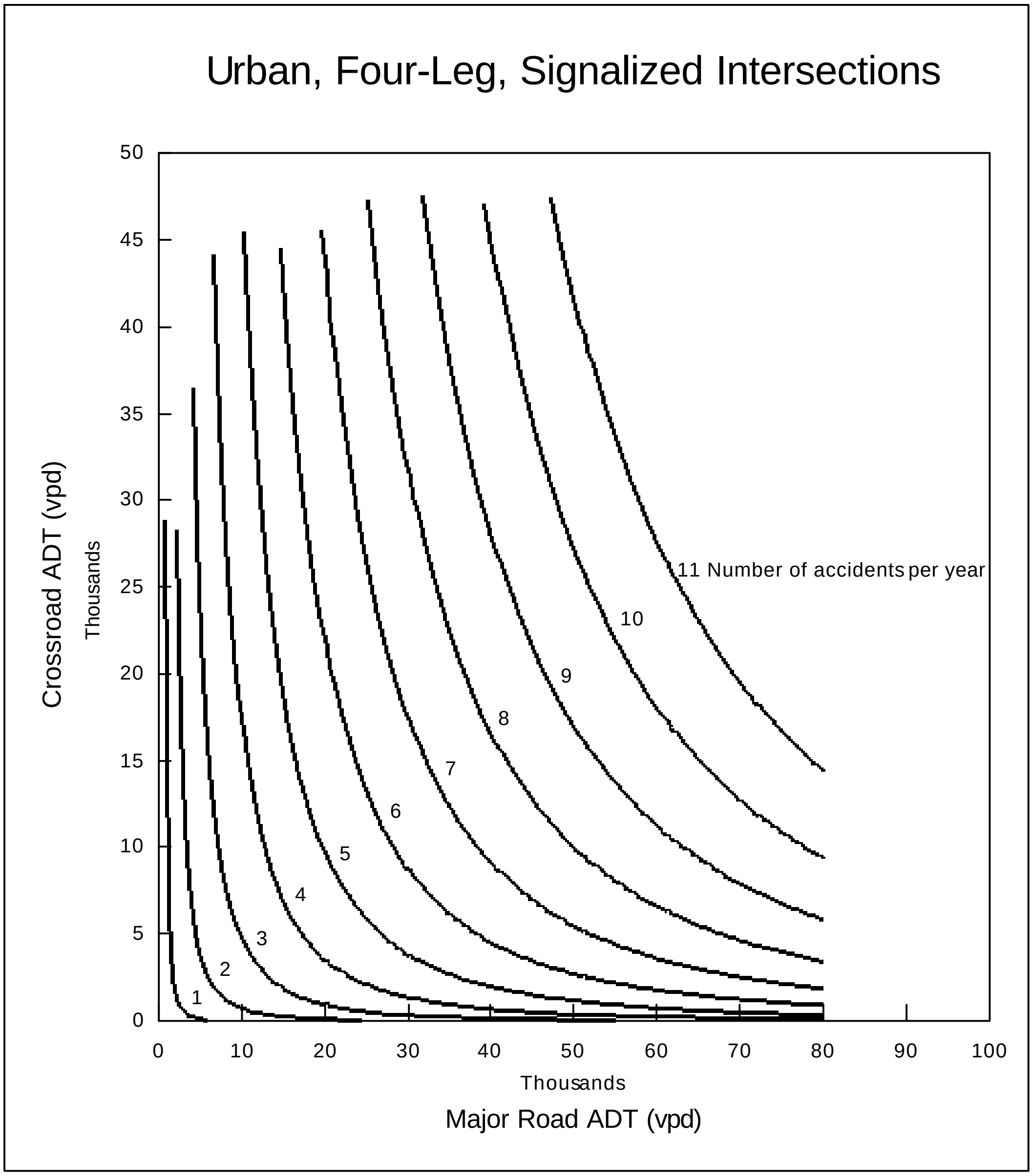
[a] All variables significant at the 90 percent confidence level or higher

[b] Chi-square likelihood ratio statistic for testing the significance of the effect of the variable; with 1 degree of freedom for continuous variables; with (p-1) degrees of freedom for categorical variables with p levels

[c] Direction of effect: I = Inverse of expected direction

[d] Relative effect of unit change in the variable on the expected number of accidents, equals exp(coefficient)

[e] 90 percent lower and upper confidence limits of the estimated coefficient



Major road characteristics:	**Crossroad characteristics**:	**Other:**
‣ No access control ‣ Average lane width of 3.6 m (12 ft) ‣ No free right turns ‣ 4 or 5 lanes	‣ 3 lanes or less	‣ Fully-actuated signal timing ‣ Multiphase signal phasing

Figure 10. Number of Accidents per Year as a Function of Traffic Volumes for Typical Urban, Four-Leg, Signalized Intersections

[This page intentionally left blank]

4. CONCLUSIONS

The following conclusions were reached as a result of the statistical analysis of relationships between traffic accidents and geometric design features of at-grade intersections conducted in this research.

1. The modeling results for accidents if all collision types are combined are similar to those that were found for multiple-vehicle accidents only, published in FHWA-RD-96-125, "Statistical Models of At-Grade Intersection Accidents."(1)
2. The negative binomial and lognormal distributions appear to be better suited to modeling of accident relationships than the normal distribution. As found in the previous study, the form of the statistical distribution selected for modeling any particular type of intersection should be chosen based on a review of the accident frequency distribution for that type of intersection. Due to overdispersion observed in the accident data, the negative binomial distribution was preferred over the Poisson distribution when using a loglinear model.
3. The lognormal and negative binomial regression models developed to represent relationships between accidents of all collision types and intersection geometric design, traffic control, and traffic volume variables explained between 16 and 39 percent of the variability in the accident data.
4. Models developed to predict accidents for all severity levels combined generally performed slightly better than did models for fatal and injury accidents.
5. In all regression models, the major-road ADT and crossroad ADT variables accounted for most of the variability in accident data that was explained by the models. Generally, geometric design variables accounted for only a small additional portion of the variability.
6. For rural, three- and four-leg, STOP-controlled intersections and urban, three-leg, STOP-controlled intersections, the geometric design features of at-grade intersections whose effects on safety were statistically significant in the negative binomial regression models included number of lanes on major road, presence of major-road left-turn prohibition, type of access control on major road, width of major-road outside shoulder, presence of major-road and crossroad right-turn channelization, design speed of major road, presence of a median, presence of major-road left-turn channelization, and average lane width on major road. The type of the terrain, functional class of major road, and presence of lighting at the intersection were also found to be statistically significant. In some cases, however, the observed effects on accidents of the geometric design and other variables identified above were in the opposite direction to that expected.
7. For urban, four-leg, signalized and STOP-controlled intersections, the lognormal distribution was found to be an appropriate choice for modeling of intersection accidents. These urban intersections experienced many more accidents than the other types of intersection evaluated, and only a small number of them experienced no accidents in the 3-year period. Geometric design features whose effects on safety were found to be statistically significant included number of lanes on major

road and crossroad, width of outside shoulder on major road, average lane width on major road, presence of major-road and crossroad right-turn channelization, presence of major-road left-turn prohibition, design speed of major road, and access control on major road. Traffic control features that were significant at signalized intersections included signal timing and signal phasing. The functional class of the major road and the presence of lighting at the intersection were also found to be statistically significant in some models. In some cases, however (as mentioned above), the observed effects on accidents of the above-identified variables were in the direction opposite that expected.

8. While the models presented in this report are the best that can be developed from the available data, they do not appear to be of direct use to practitioners. The models do not include effects for all geometric variables of potential interest to highway designers, and some of the effects they do include are in a direction opposite to that expected. Furthermore, the goodness of fit of the models is not as high as would be desired. Therefore, the models presented here are appropriate as a guide to future research, but do not appear appropriate for direct application in the field.

APPENDIX
ACCIDENT TYPE AND SEVERITY DISTRIBUTIONS

Tables 22 through 26 present distributions of accident type and severity for the 5 intersection types whose accident experience has been modeled in this report. For multiple-vehicle accidents, the accident type distribution is based primarily on standard Caltrans categories for collision type. However, all collisions that involve one or more left-turning or U-turning vehicles have been classified as left-turn accidents, and all collisions that involve one or more right-turning vehicles (but no left- or U-turning vehicles) have been classified as right-turn collisions. For single-vehicle accidents, the collision type has been based on standard Caltrans data for type of collision, movement preceding collision, and object struck. Accident severity is classified into three categories based on the most serious injury sustained in the accident: fatal, injury, and property-damage-only. Data were not available to subdivide injury accidents further by severity level.

Tables 22 through 26 include all accidents that occurred during a 3-year period at the intersections whose accident experience has been modeled in this report. The multiple-vehicle accidents shown in the tables are the accidents that were used in development of the models presented in Report No. FHWA-RD-96-125.[1]

Table 22. Distribution of Accident Type and Accident Severity for Rural, Four-Leg, STOP-Controlled Intersections

Accident Type	Number (percentage) of accidents by severity level							
	Fatal		Injury		Property damage only		Total	
Single-Vehicle Accidents								
Ran off road	4	(2.6)	28	(1.1)	27	(0.9)	59	(1.0)
Overturned in road	1	(0.7)	27	(1.0)	21	(0.7)	49	(0.9)
Other noncollision	0	(0.0)	18	(0.7)	73	(2.5)	91	(1.6)
Collision with parked car	1	(0.7)	14	(0.5)	81	(2.8)	96	(1.7)
Collision with train	0	(0.0)	0	(0.0)	0	(0.0)	0	(0.0)
Collision with pedestrian	0	(0.0)	0	(0.0)	0	(0.0)	0	(0.0)
Collision with bicycle	0	(0.0)	1	(0.0)	0	(0.0)	1	(0.0)
Collision with animal	0	(0.0)	4	(0.2)	2	(0.1)	6	(0.1)
Collision with fixed object	3	(2.0)	81	(3.1)	189	(6.6)	273	(4.8)
Other single-vehicle collision	4	(2.6)	81	(3.1)	113	(3.9)	198	(3.5)
Total single-vehicle accidents	13	(8.5)	254	(9.7)	506	(17.6)	773	(13.7)
Multiple-Vehicle Accidents								
Head-on collision	6	(3.9)	27	(1.0)	19	(0.7)	52	(0.9)
Sideswipe collision	1	(0.7)	39	(1.5)	130	(4.5)	170	(3.0)
Rear-end collision	3	(2.0)	251	(9.6)	369	(12.8)	623	(11.1)
Right-angle collision	85	(55.6)	1,000	(38.4)	671	(23.4)	1,756	(31.2)
Right-turn collision	1	(0.7)	84	(3.2)	226	(7.9)	311	(5.5)
Left-turn collision	34	(22.2)	816	(31.3)	840	(29.2)	1,690	(30.0)
Other multiple-vehicle collision	10	(6.5)	135	(5.2)	111	(3.9)	256	(4.5)
Total multiple-vehicle accidents	140	(91.5)	2,352	(90.3)	2,366	(82.4)	4,858	(86.3)
Total Accidents	153	(100.0)	2,606	(100.0)	2,872	(100.0)	5,631	(100.0)

Note: Numbers in parentheses are column percentages.

Table 23. Distribution of Accident Type and Accident Severity for Rural, Three-Leg, STOP-Controlled Intersections

Accident Type	Number (percentage) of accidents by severity level							
	Fatal		Injury		Property damage only		Total	
Single-Vehicle Accidents								
Ran off road	5	(5.1)	84	(3.0)	42	(1.2)	131	(2.0)
Overturned in road	1	(1.0)	73	(2.6)	35	(1.0)	109	(1.7)
Other noncollision	1	(1.0)	46	(1.6)	85	(2.4)	132	(2.1)
Collision with parked car	0	(0.0)	34	(1.2)	81	(2.3)	115	(1.8)
Collision with train	0	(0.0)	0	(0.0)	0	(0.0)	0	(0.0)
Collision with pedestrian	0	(0.0)	0	(0.0)	0	(0.0)	0	(0.0)
Collision with bicycle	0	(0.0)	0	(0.0)	1	(0.0)	1	(0.0)
Collision with animal	0	(0.0)	3	(0.1)	9	(0.3)	12	(0.2)
Collision with fixed object	5	(5.1)	196	(7.0)	294	(8.4)	495	(7.7)
Other single-vehicle collision	11	(11.1)	256	(9.1)	268	(7.7)	535	(8.4)
Total single-vehicle accidents	23	(23.2)	692	(24.7)	815	(23.3)	1,530	(23.9)
Multiple-Vehicle Accidents								
Head-on collision	12	(12.1)	74	(2.6)	36	(1.0)	122	(1.9)
Sideswipe collision	2	(2.0)	52	(1.9)	231	(6.6)	285	(4.5)
Rear-end collision	4	(4.0)	391	(13.9)	601	(17.2)	996	(15.6)
Right-angle collision	8	(8.1)	212	(7.6)	233	(6.7)	453	(7.1)
Right-turn collision	1	(1.0)	107	(3.8)	278	(8.0)	386	(6.0)
Left-turn collision	36	(36.4)	1,129	(40.2)	1,172	(33.5)	2,337	(36.5)
Other multiple-vehicle collision	13	(13.1)	149	(5.3)	128	(3.7)	290	(4.5)
Total multiple-vehicle accidents	76	(76.8)	2,114	(75.3)	2,679	(76.7)	4,869	(76.1)
Total Accidents	99	(100.0)	2,806	(100.0)	3,494	(100.0)	6,399	(100.0)

Note: Numbers in parentheses are column percentages.

Table 24. Distribution of Accident Type and Accident Severity for Urban, Four-Leg, STOP-Controlled Intersections

Accident Type	Number (percentage) of accidents by severity level							
	Fatal		Injury		Property damage only		Total	
Single-Vehicle Accidents								
Ran off road	0	(0.0)	8	(0.2)	10	(0.2)	18	(0.2)
Overturned in road	0	(0.0)	38	(0.9)	9	(0.2)	47	(0.5)
Other noncollision	0	(0.0)	40	(0.9)	93	(1.7)	133	(1.3)
Collision with parked car	3	(3.9)	81	(1.9)	351	(6.3)	435	(4.4)
Collision with train	0	(0.0)	0	(0.0)	0	(0.0)	0	(0.0)
Collision with pedestrian	0	(0.0)	1	(0.0)	0	(0.0)	1	(0.0)
Collision with bicycle	0	(0.0)	1	(0.0)	1	(0.0)	2	(0.0)
Collision with animal	0	(0.0)	0	(0.0)	0	(0.0)	0	(0.0)
Collision with fixed object	1	(1.3)	56	(1.3)	162	(2.9)	219	(2.2)
Other single-vehicle collision	0	(0.0)	115	(2.7)	225	(4.0)	340	(3.4)
Total single-vehicle accidents	4	(5.2)	340	(7.9)	851	(15.2)	1,195	(12.0)
Multiple-Vehicle Accidents								
Head-on collision	0	(0.0)	31	(0.7)	26	(0.5)	57	(0.6)
Sideswipe collision	0	(0.0)	82	(1.9)	361	(6.5)	443	(4.4)
Rear-end collision	1	(1.3)	989	(22.9)	1,249	(22.3)	2,239	(22.4)
Right-angle collision	21	(27.3)	870	(20.1)	927	(16.6)	1,818	(18.2)
Right-turn collision	1	(1.3)	255	(5.9)	417	(7.5)	673	(6.7)
Left-turn collision	14	(18.2)	1,201	(27.8)	1,499	(26.8)	2,714	(27.2)
Other multiple-vehicle collision	36	(46.8)	550	(12.7)	264	(4.7)	850	(8.5)
Total multiple-vehicle accidents	73	(94.8)	3,978	(92.1)	4,743	(84.8)	8,794	(88.0)
Total Accidents	77	(100.0)	4,318	(100.0)	5,594	(100.0)	9,989	(100.0)

Note: Numbers in parentheses are column percentages.

Table 25. Distribution of Accident Type and Accident Severity for Urban, Three-Leg, STOP-Controlled Intersections

Accident Type	Number (percentage) of accidents by severity level							
	Fatal		Injury		Property damage only		Total	
Single-Vehicle Accidents								
Ran off road	2	(1.8)	15	(0.3)	12	(0.1)	29	(0.2)
Overturned in road	0	(0.0)	52	(0.9)	14	(0.2)	66	(0.5)
Other noncollision	1	(0.9)	56	(1.0)	111	(1.4)	168	(1.2)
Collision with parked car	2	(1.8)	105	(1.8)	501	(6.2)	608	(4.4)
Collision with train	0	(0.0)	0	(0.0)	0	(0.0)	0	(0.0)
Collision with pedestrian	0	(0.0)	1	(0.0)	0	(0.0)	1	(0.0)
Collision with bicycle	0	(0.0)	2	(0.0)	1	(0.0)	3	(0.0)
Collision with animal	0	(0.0)	4	(0.1)	6	(0.1)	10	(0.1)
Collision with fixed object	8	(7.1)	114	(2.0)	245	(3.0)	367	(2.6)
Other single-vehicle collision	10	(8.9)	216	(3.8)	439	(5.4)	665	(4.8)
Total single-vehicle accidents	23	(20.5)	565	(9.9)	1,329	(16.5)	1,917	(13.8)
Multiple-Vehicle Accidents								
Head-on collision	7	(6.3)	69	(1.2)	43	(0.5)	119	(0.9)
Sideswipe collision	2	(1.8)	141	(2.5)	684	(8.5)	827	(6.0)
Rear-end collision	5	(4.5)	1,710	(30.1)	2,370	(29.4)	4,085	(29.5)
Right-angle collision	7	(6.3)	414	(7.3)	584	(7.2)	1,005	(7.2)
Right-turn collision	1	(0.9)	381	(6.7)	599	(7.4)	981	(7.1)
Left-turn collision	26	(23.2)	1,813	(31.9)	2,190	(27.1)	4,029	(29.1)
Other multiple-vehicle collision	41	(36.6)	597	(10.5)	268	(3.3)	906	(6.5)
Total multiple-vehicle accidents	89	(79.5)	5,125	(90.1)	6,738	(83.5)	11,952	(86.2)
Total Accidents	112	(100.0)	5,690	(100.0)	8,067	(100.0)	13,869	(100.0)

Note: Numbers in parentheses are column percentages.

Table 26. Distribution of Accident Type and Accident Severity for Urban, Four-Leg, Signalized Intersections

Accident Type	Number (percentage) of accidents by severity level							
	Fatal		Injury		Property damage only		Total	
Single-Vehicle Accidents								
Ran off road	0	(0.0)	9	(0.1)	7	(0.0)	16	(0.1)
Overturned in road	0	(0.0)	95	(0.8)	30	(0.2)	125	(0.4)
Other noncollision	0	(0.0)	65	(0.5)	92	(0.5)	157	(0.5)
Collision with parked car	1	(1.0)	56	(0.5)	310	(1.7)	367	(1.2)
Collision with train	0	(0.0)	0	(0.0)	0	(0.0)	0	(0.0)
Collision with pedestrian	0	(0.0)	3	(0.0)	0	(0.0)	3	(0.0)
Collision with bicycle	0	(0.0)	2	(0.0)	3	(0.0)	5	(0.0)
Collision with animal	0	(0.0)	0	(0.0)	4	(0.0)	4	(0.0)
Collision with fixed object	2	(2.1)	105	(0.8)	399	(2.2)	506	(1.7)
Other single-vehicle collision	9	(9.3)	237	(1.9)	768	(4.2)	1,014	(3.3)
Total single-vehicle accidents	12	(12.4)	572	(4.6)	1,613	(8.9)	2,197	(7.2)
Multiple-Vehicle Accidents								
Head-on collision	1	(1.0)	57	(0.5)	61	(0.3)	119	(0.4)
Sideswipe collision	0	(0.0)	253	(2.0)	1,446	(8.0)	1,699	(5.6)
Rear-end collision	7	(7.2)	4,405	(35.5)	6,163	(34.6)	10,575	(34.6)
Right-angle collision	21	(21.6)	1,803	(14.5)	1,991	(11.0)	3,815	(12.5)
Right-turn collision	5	(5.2)	993	(8.0)	1,730	(9.6)	2,728	(8.9)
Left-turn collision	19	(19.6)	3,298	(26.6)	4,305	(23.8)	7,622	(24.9)
Other multiple-vehicle collision	32	(33.0)	1,021	(8.2)	795	(4.4)	1,848	(6.0)
Total multiple-vehicle accidents	85	(87.6)	11,830	(95.4)	16,491	(91.1)	28,406	(92.8)
Total Accidents	97	(100.0)	12,402	(100.0)	18,104	(100.0)	30,603	(100.0)

Note: Numbers in parentheses are column percentages.

REFERENCES

1. Bauer, K.M., and D.W. Harwood, "Statistical Models of At-Grade Intersection Accidents," Report No. FHWA-RD-96-125, Federal Highway Administration, McLean, VA, 1996.

2. McCullagh, P., and J.A. Nelder, *Generalized Linear Models*, Second Edition (1989), New York: Chapman & Hall.

3. SAS Institute Inc., SAS® Technical Report P-243, SAS/STAT® Software: The GENMOD Procedure, Release 6.09, Cary, NC: SAS Institute, Inc., 1993, 88 pp.

4. Fridstrøm, L., et al., "Measuring the Contribution of Randomness, Exposure, Weather, and Daylight to the Variation in Road Accident Counts," *Accident Analysis and Prevention* (1995), Vol. 27, pp. 1–20.