Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA/TX-08/0-4703-5		

4. Title and Subtitle	5. Report Date
CALIBRATION FACTORS HANDBOOK: SAFETY PREDICTION MODELS CALIBRATED WITH TEXAS HIGHWAY SYSTEM DATA	June 2008 Published: October 2008
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
J. Bonneson and M. Pratt	Report 0-4703-5

9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
Texas Transportation Institute The Texas A&M University System College Station, Texas 77843-3135	11. Contract or Grant No. Project 0-4703

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Texas Department of Transportation Research and Technology Implementation Office P.O. Box 5080 Austin, Texas 78763-5080	Technical Report: September 2003-June 2008
	14. Sponsoring Agency Code

15. Supplementary Notes

Project performed in cooperation with the Texas Department of Transportation and the Federal Highway Administration.
Project Title: Incorporating Safety into the Highway Design Process
URL: http://tti.tamu.edu/documents/0-4703-5.pdf

16. Abstract

Highway safety is an ongoing concern to the Texas Department of Transportation (TxDOT). As part of its proactive commitment to improving highway safety, TxDOT is moving toward including quantitative safety analyses earlier in the project development process. The objectives of this research project are: (1) the development of safety design guidelines and evaluation tools to be used by TxDOT designers, and (2) the production of a plan for the incorporation of these guidelines and tools in the planning and design stages of the project development process.

This document summarizes the research conducted and the conclusions reached during the development of safety prediction models for intersections and highway segments in Texas. Models were developed for urban and suburban arterial intersections, urban and suburban arterial street segments, rural multilane highway segments, and urban and rural freeway segments. They were subsequently calibrated using Texas highway system data. Selected accident modification factors were also developed and calibrated. These factors address several geometric design elements, including turn bay presence, median width, barrier presence, and weaving section length.

17. Key Words	18. Distribution Statement
Highway Safety, Highway Design, Safety Management, Geometric Design	No restrictions. This document is available to the public through NTIS: National Technical Information Service Springfield, Virginia 22161 http://www.ntis.gov

19. Security Classif.(of this report)	20. Security Classif.(of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	160	

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

CALIBRATION FACTORS HANDBOOK: SAFETY PREDICTION MODELS CALIBRATED WITH TEXAS HIGHWAY SYSTEM DATA

by

J. Bonneson, P.E.
Senior Research Engineer
Texas Transportation Institute

and

M. Pratt
Associate Transportation Researcher
Texas Transportation Institute

Report 0-4703-5
Project 0-4703
Project Title: Incorporating Safety into the Highway Design Process

Performed in cooperation with the
Texas Department of Transportation
and the
Federal Highway Administration

June 2008
Published: October 2008

TEXAS TRANSPORTATION INSTITUTE
The Texas A&M University System
College Station, Texas 77843-3135

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data published herein. The contents do not necessarily reflect the official view or policies of the Federal Highway Administration (FHWA) and/or the Texas Department of Transportation (TxDOT). This report does not constitute a standard, specification, or regulation. It is not intended for construction, bidding, or permit purposes. The engineer in charge of the project was James Bonneson, P.E. #67178.

NOTICE

The United States Government and the State of Texas do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of this report.

ACKNOWLEDGMENTS

This research project was sponsored by the Texas Department of Transportation and the Federal Highway Administration. The research was conducted by Dr. James Bonneson and Mr. Michael Pratt. Both researchers are with the Texas Transportation Institute.

The researchers would like to acknowledge the support and guidance provided by the Project Monitoring Committee:

- Ms. Aurora (Rory) Meza, Project Coordinator (TxDOT);
- Ms. Elizabeth Hilton, Project Director (TxDOT);
- Mr. David Bartz (FHWA);
- Mr. Stan Hall (TxDOT);
- Mr. Richard Harper (TxDOT);
- Ms. Meg Moore (TxDOT);
- Ms. Wendy Simmons (TxDOT); and
- Ms. Joanne Wright (TxDOT).

In addition, the researchers would like to acknowledge the valuable assistance provided by Dr. Dominique Lord during the development of several report chapters.

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

Table **Page**

LIST OF TABLES (Continued)

Table **Page**

CHAPTER 1. INTRODUCTION

OVERVIEW

There is a growing public demand for safer streets and highways. In response to this demand, state and national transportation agencies have developed safety programs that emphasize public education, accelerated highway renewal, community-sensitive street systems, and innovative technology to facilitate safe highway design.

Historically, information about the safety effect of a design component has been based on anecdotal evidence, laws of physics, before-after studies, or comparisons of site safety (i.e., sites with and without the design component). However, the accuracy of this information is suspect because of the inherent random nature of crash data and the many factors (some of which pertain more to the driver and the vehicle than the roadway) that can lead to a crash at a specific location. As a result of this uncertainty, engineers have traditionally come to rely on design standards and policies to guide them in the design process, with the underlying premise that compliance with warrants and controls will yield a "safe" roadway.

In general, the safety experience with roadways built in compliance with warrants and controls has been good and the aforementioned premise largely validated. However, the weaknesses of this traditional design approach have become more apparent as traffic demands have increased over time, the performance of vehicles improved, and drivers became less patient. Points along the roadway having multiple, complex geometric components that tend to concentrate traffic and increase their interaction have started to show disproportionately high crash frequencies. Fortunately, in the past decade, new statistical analysis methods have been developed and the quality of crash data has improved. These advances have significantly increased both the accuracy and coverage of design-related safety information. Emerging technology is now making it possible to efficiently incorporate quantitative safety evaluations in the design process.

A significant amount of new safety information has been developed in recent years. The implementation of this information is now as pressing a problem as was the need for new research a decade ago. The forthcoming *Highway Safety Manual* (*HSM*) is expected to formalize the safety evaluation process; however, the *HSM* procedures will require local calibration to ensure that they accurately reflect the conditions in the jurisdiction for which they are used (*1*). Moreover, it is likely that each agency will want to tailor the calibrated procedures to ensure their consistency with local design policies.

OBJECTIVES

Highway safety concerns are also evident in Texas. Crashes in Texas continue to increase and currently exceed 300,000 per year. Nearly 3800 motorists die annually on Texas highways. As part of its proactive commitment to improving highway safety, the Texas Department of Transportation (TxDOT) is moving toward including quantitative safety analyses throughout the project development process. This research project has as its objectives: (1) the development of

safety design guidelines and evaluation tools to be used by TxDOT designers, and (2) the production of a plan for the incorporation of these guidelines and tools in the planning and design stages of the project development process.

The primary objective of this report is to document the activities undertaken to calibrate safety prediction models for the following facility types:

- Urban and Suburban Arterial Street Intersections,
- Urban and Suburban Arterial Street Segments,
- Rural Multilane Highway Segments, and
- Rural and Urban Freeway Segments.

These activities included an initial consideration of the feasibility of recalibrating the existing models in the *Interim Roadway Safety Design Workbook* (*2*). If it was determined that an alternative model form was more appropriate, then a new model form was derived and calibrated. Of the four facility types listed, only the intersection models were recalibrated. A new model form was developed and calibrated for each of the three segment models.

A secondary objective of this report is to document the development of several new accident modification factors. These factors were developed and calibrated in conjunction with the aforementioned safety prediction models. This development was typically dictated by the need to explain observed systematic variation in the crash data that is logically and statistically correlated with variation in a particular geometric design element, land use, or area type.

RESEARCH APPROACH

A six-year program of research was developed to satisfy the project's objectives. The research approach consists of 10 tasks that represent a logical sequence of needs assessment, research, evaluation, and workshop development. These tasks are identified in the following list.

1. Review the TxDOT design and safety evaluation processes.
2. Identify safety information sources and needs.
3. Determine the data needed for selected safety evaluation tools.
4. Evaluate use of accident modification factors for design evaluation.
5. Determine calibration factors for Texas application of safety prediction models.
6. Develop an *Interim Roadway Safety Design Workbook* and a *Roadway Safety Design Synthesis* (*2, 3*), and conduct research needed to develop a final version of the *Workbook*.
7. Develop workshop education materials.
8. Develop an Internet-based safety information resource.
9. Evaluate use of safety evaluation tools by TxDOT staff and refine as appropriate.

The first eight tasks have been undertaken in the first five years of the research. Tasks 1 through 5 have been completed. The findings from Task 5 are collectively documented in this report and a previous report by Bonneson et al. (*4*).

DEFINITIONS

This part of the chapter defines several terms related to highway safety. The definitions offered are consistent with their use in the safety-related literature; however, they may be refined for consistency with TxDOT design practice and the objectives of this research project.

Accident modification factor (AMF) is a constant or equation that represents the change in safety following a change in the character of a segment (or intersection). An AMF can be computed as the ratio $N_w/N_{w/o}$, where N_w represents the expected number of crashes experienced by a highway segment *with* one or more specified design components, and $N_{w/o}$ represents the expected number of crashes experienced by the same segment *without* the specified components. AMFs are often used as multiplicative factors to adjust the estimate obtained from a safety prediction model to a value that reflects the safety of a specific segment.

AMFs typically range in value from 0.5 to 2.0, with a value of 1.0 representing no effect on safety. AMFs less than 1.0 indicate that the specified component is associated with fewer crashes.

To illustrate the concept of AMF, consider a road segment that has an expected crash frequency of 3.0 crashes/yr. A change is made to the road cross section and, after a period of time, a follow-up evaluation indicates that the change resulted in an expected crash frequency of 4.0 crashes/yr. The AMF for this change is 1.3 (= 4.0/3.0).

As a second illustration, consider that a safety prediction model is used to estimate the expected crash frequency of a typical two-lane highway with a specified average daily traffic volume (ADT) and length. The model was developed to reflect the following as "typical": 12-ft lanes, 6-ft shoulders, no grade, no horizontal curves, 10-ft horizontal clearance, 1V:4H side slope, and no vertical grades. This model estimates an expected crash frequency of 5.0 crashes/yr for the "typical" road segment. It is desired to estimate the crash frequency of a specific road segment for which all geometric elements are "typical" except that the clear zone is 20 ft wide. The AMF for horizontal clearance has a value of 0.93 when the clearance distance is 20 ft. Thus, the expected crash frequency for the specific road segment is estimated as 4.6 crashes/yr (= 5.0 × 0.93).

Crash reduction factor (CRF) is a constant that represents the proportion of crashes reduced as a result of a safety improvement at a specific location or along a specific road segment. CRFs typically range in value from 0.10 to 0.90. Larger CRFs in this range indicate a more significant reduction in crashes due to the improvement. To illustrate, consider a road segment that has a crash frequency of 3.0 crashes/yr. An improvement is made to the road's cross section and, after a period of time passes, a follow-up evaluation indicates that the change resulted in a crash frequency of 2.0 crashes/yr. The CRF for this improvement is 0.33 (= [3.0 – 2.0]/3.0), representing a 33 percent reduction in crashes.

Injury crash is a crash wherein one or more of the persons involved is injured. The injury severity is reported as "possible," "non-incapacitating," or "incapacitating."

Safety (or "substantive safety") is the expected crash frequency associated with a segment (or intersection) for a given set of design components, traffic control devices, and exposure conditions (e.g., traffic volume, segment length). Given that crashes are random events and that conditions can change over time, the safety of a specific segment is best conceptualized as the long-run average of the crash frequencies reported for a large group of segments with similar features and traffic conditions.

Safety evaluation tool is, at its simplest level, a set of equations that can be used to predict: (1) the safety of a given segment (or intersection), and (2) the safety effect associated with a change in its design features. At this "simple" level, a tool is equivalent to a model. However, complex tools can incorporate additional analysis techniques. For example, complex tools can include techniques for incorporating the reported crash history of a specific segment to improve the accuracy of the safety prediction. Complex tools can also include techniques for evaluating alternative designs using safety and other data (e.g., benefit-cost analysis). Tools are sometimes represented in software to facilitate their application.

Safety prediction model is an equation, or set of equations, that can be used to estimate the safety of a typical segment (or intersection). The model includes factors related to crash risk and exposure. A figure or table is sometimes used to portray the relationship (instead of an equation). A model can be derived to include one or more AMFs. Models intended for practical application have one or more empirically based factors that require calibration to local conditions to ensure accurate predictions.

Safety surrogate is any statistic that is directly related to crash frequency or severity (e.g., conflicts) and that quantifies the relative risk of collision or injury.

REFERENCES

1. Hughes, W., K. Eccles, D. Harwood, I. Potts, and E. Hauer. *Research Results Digest 286: Development of a Highway Safety Manual*. National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C., March 2004.

2. Bonneson, J.A., K. Zimmerman, and K. Fitzpatrick. *Interim Roadway Safety Design Workbook*. FHWA/TX-06/0-4703-P4, Texas Department of Transportation, Austin, Texas, April 2006.

3. Bonneson, J.A., K. Zimmerman, and K. Fitzpatrick. *Roadway Safety Design Synthesis*. FHWA/TX-05/0-4703-P1. Texas Department of Transportation, Austin, Texas, November 2005.

4. Bonneson, J., D. Lord, K. Zimmerman, K. Fitzpatrick, and M. Pratt. *Development of Tools for Evaluating the Safety Implications of Highway Design Decisions*. FHWA/TX-07/0-4703-4. Texas Department of Transportation, Austin, Texas, September 2006.

CHAPTER 2. URBAN AND SUBURBAN ARTERIAL STREET INTERSECTIONS

OVERVIEW

This chapter describes the activities undertaken to recalibrate existing safety prediction models for urban and suburban arterial street intersections in Texas. A series of accident modification factors were also developed as part of the recalibration activity. The models considered for recalibration are described in the *Interim Roadway Safety Design Workbook* (*1*). This *Workbook* describes separate models for the following intersection configuration and control-type combinations:

- three-leg intersections with signal control,
- three-leg intersections with one-way stop control,
- four-leg intersections with signal control, and
- four-leg intersections with two-way stop control.

Each model provides an estimate of the expected crash frequency for the intersection, given specified traffic volume and geometric design conditions. Each model was initially calibrated using data from locations outside of Texas. To ensure that the estimate is not biased for Texas applications, each model was subsequently recalibrated using data specific to Texas.

The objective of this research was to recalibrate models for each of the aforementioned intersection combinations using crash data for Texas highways. However, as part of the recalibration process, several additional AMFs were developed to control for geometric differences among the intersections in the database. These AMFs can be used with the recalibrated base models to estimate the change in safety associated with a change in geometry.

This chapter is divided into four parts. The first part provides a review of the literature on the topic of safety prediction model calibration. The second part describes the development of a generalized intersection AMF equation and its application to turn lane installations. The third part summarizes the method used to recalibrate the existing models. The last part describes the analysis and findings from the recalibration of the intersection models.

LITERATURE REVIEW

This part of the chapter describes a framework for safety prediction model calibration. It consists of three sections. The first section provides a brief overview of safety prediction and the role of "base" prediction models. The second section summarizes the base models described in the *Workbook*. The last section reviews the data available in TxDOT's road inventory database.

Safety Prediction Models

The expected crash frequency for an intersection with specified attributes is computed using a safety prediction model. This model represents the combination of a "base" model and one or more AMFs. The base model is used to estimate the expected crash frequency for a typical intersection. The AMFs are used to adjust the base estimate when the attributes of the specific segment are not considered typical. For intersections, the safety prediction model is shown below as Equation 1, and its base model component is shown as Equation 2.

with,

$$E[N] = E[N]_b \times AMF_{lw} \times AMF_{dd} \ldots \quad \textbf{(1)}$$

$$E[N]_b = a\ ADT_{major}^{b_1}\ ADT_{minor}^{b_2} \quad \textbf{(2)}$$

where,

$E[N]$ = expected crash frequency, crashes/yr;
$E[N]_b$ = expected base crash frequency, crashes/yr;
AMF_{lw} = lane width accident modification factor;
AMF_{dd} = driveway density accident modification factor;
a, b_1, b_2 = calibration coefficients;
ADT_{major} = average daily traffic volume on the major road, veh/d; and
ADT_{minor} = average daily traffic volume on the minor road, veh/d.

Equation 2 predicts 0.0 crashes when either *ADT* variable is equal to 0.0. This boundary condition is illogical because some types of intersection-related crashes (e.g., rear-end) are likely to occur even when the cross street ADT is negligible. An alternative form for the intersection base model that does not share this limitation is:

$$E[N]_b = a\ (ADT_{major} + ADT_{minor})^b \quad \textbf{(3)}$$

When the coefficient *b* in this equation is equal to 1.0, then the coefficient *a* is effectively equal to the crash rate, with units of "crashes per million-entering-vehicles."

In practice, the selected base model form is typically used for all intersection configurations (i.e., three-leg, four-leg, etc.) and control types (i.e., signalized, unsignalized). However, it is separately calibrated for each configuration and control-type combination. This approach yields unique calibration coefficients for each combination.

Base Models

This section describes base models derived by Bonneson et al. (*1*) for urban and suburban arterial street intersections. These models are based on a review and synthesis of various models described in the research literature (*2*). The base model described in the *Workbook* is:

$$E[N]_b = 0.000365\ Base\ (ADT_{major} + ADT_{minor}) \quad \textbf{(4)}$$

where,

Base = base injury (plus fatal) crash rate, crashes/million-entering-vehicles.

The base crash rate used in Equation 4 is obtained from either Table 2-1 or Table 2-2.

Table 2-1. Crash Rates for Three-Leg Urban and Suburban Intersections.

Control Type	Crash Rate, injury + fatal crashes per million-entering-vehicles				
	Ratio of Minor-Street to Major-Street Volume				
	0.05	**0.10**	**0.15**	**0.20**	**0.25**
Unsignalized [1]	0.18	0.21	0.22	0.22	0.23
Signalized	0.12	0.15	0.17	0.18	0.19

Note:
1 - Unsignalized intersections have an uncontrolled major street and a stop-controlled minor street.

Table 2-2. Crash Rates for Four-Leg Urban and Suburban Intersections.

Control Type	Major-Street Volume, veh/d	Crash Rate, injury + fatal crashes per million-entering-vehicles				
		Ratio of Minor-Street to Major-Street Volume				
		0.10	**0.30**	**0.50**	**0.70**	**0.90**
Unsignalized [1]	5000	0.25	0.29	0.28	0.27	0.26
	10,000	0.23	0.26	0.26	0.25	0.24
	15,000	0.22	0.24	0.24	0.23	0.22
	20,000	0.21	Intersection very likely to meet signal warrants			
	25,000	0.20				
Signalized	5000	0.19	0.24	0.26	0.26	0.26
	10,000	0.17	0.22	0.23	0.23	0.23
	15,000	0.16	0.21	0.22	0.22	0.22
	20,000	0.15	0.20	0.21	0.21	0.21
	25,000	0.15	0.19	0.20	0.21	0.20
	30,000	0.14	0.19	0.20	0.20	0.20
	40,000	0.14	0.18	0.19	0.19	0.19
	≥50,000	0.13	0.17	0.18	0.19	0.18

Note:
1 - Unsignalized intersections have an uncontrolled major street and a stop-controlled minor street.

The base crash rates in Tables 2-1 and 2-2 were derived using Equation 4 and the expected crash frequency computed using the following equation:

$$E[N]_b = a \left(\frac{ADT_{major}}{1000} \right)^{b_1} \left(\frac{ADT_{minor}}{1000} \right)^{b_2} \quad \textbf{(5)}$$

The constant and coefficients listed in Table 2-3 can be used in Equation 5 to yield the same expected crash frequency as obtained from Equation 4 using the rates from Table 2-1 or 2-2.

Table 2-3. Intersection Base Model Coefficients.

Control Type	Three Intersection Legs			Four Intersection Legs		
	a	b_1	b_2	a	b_1	b_2
Unsignalized [1]	0.143	0.766	0.248	0.234	0.596	0.260
Signalized	0.143	0.629	0.385	0.234	0.459	0.397

Note:
1 - Unsignalized intersections have an uncontrolled major street and a stop-controlled minor street.

Either of the two aforementioned base models is applicable to intersections that have typical geometric design conditions. These typical conditions are identified in the *Workbook* and are restated in columns two and four of Table 2-4. The base conditions listed are representative of intersections on urban and suburban arterial streets.

Table 2-4. Base Conditions for Urban and Suburban Intersections.

Characteristic	Unsignalized		Signalized	
	Base Condition	Sensitivity of Continuous Variables	Base Condition	Sensitivity of Continuous Variables
Left-turn lanes on major street	present	--	present	--
Right-turn lanes on major street	not present	--	not present	--
Number of lanes on major street	4	--	4	--
Number of lanes on minor street	2	--	2	--
Lane width on major street	12 ft	11 to 13 ft	12 ft	11 to 13 ft
Shoulder width on major street	1.5 ft	0 to 3 ft	not applicable	--
Median presence on major street	not present	16 to 20 ft	not applicable	--

AMFs are available in the *Workbook* to assess each of the design characteristics listed in Table 2-4. For routine applications, the base model is combined with the AMFs in the safety prediction model (i.e., Equation 1) to estimate the crash frequency for a particular intersection that has design elements that are not consistent with the base conditions. The base condition for two characteristics is designated as "not applicable" in this table. This designation is a reflection of the fact that AMFs have not been developed for the corresponding characteristics and, for this reason, a base condition has not been defined.

Several of the characteristics in Table 2-4 can be represented as a continuous variable. These characteristics were evaluated for a range of values to examine the sensitivity of the corresponding AMF to a change in the continuous variable. The results of this sensitivity analysis are listed in columns three and five of Table 2-4. The ranges listed in these columns equate to AMF values between 0.95 and 1.05. For example, a lane width between 11 and 13 ft yields AMF values of about 1.05 and 0.95, respectively. Lane widths less than 11 ft yield an AMF value in excess of 1.05. The ranges listed provide some indication of the correlation between a change in the characteristic and the change in crash risk.

Texas Highway Data

The geometric and traffic attributes for the Texas state highway system were obtained from the Texas Reference Marker (TRM) system highway database. This database is maintained by TxDOT and contains data for several hundred highway intersections, each of which is described in terms of its geometry, traffic volume, and location attributes. In fact, about 140 attributes are used to describe each intersecting highway in TRM. However, only about 20 attributes are used to describe highway geometry or traffic characteristics, and about 10 more attributes are used to describe the street name and physical location. The remaining attributes describe administrative designations and road-management-related information. The relevant intersection characteristics that are available in the TRM database are identified in Table 2-5.

Table 2-5. Urban and Suburban Intersection Calibration Data.

Characteristic	Availability in TRM Database
Number of lanes on major street	Yes
Number of lanes on minor street	Yes [1]
Lane width on major street	Yes (using surface width)
Shoulder width on major street	Yes
Median presence on major street	Yes
Alignment skew angle	Yes
Truck presence	Yes [1]

Note:
1 - Data are available for the minor street only if it is a state-maintained highway.

An intersection is defined in TRM as a point entity. It has a pair of reference numbers that uniquely identify it in the database and locate it on the highway system. One reference number is assigned to each of the intersecting roadways. Data in TRM only apply to roadways on the state highway system. Hence, TRM only has information about both intersecting roadways when these two roadways are part of the state highway system.

AMF DEVELOPMENT

This part of the chapter describes the development of a general AMF model for quantifying the relationship between safety and some geometric characteristic of an intersection leg or perhaps one of the intersecting streets. This model was developed because most of the AMFs in the *Workbook* apply only to the major street. This limitation is a reflection of the manner in which the corresponding AMFs were derived, as documented in Reference 2. Nevertheless, it leaves a void in the information about the effect of similar treatments on the minor street. The general AMF model is intended to provide a means of extrapolating an AMF derived for one street to the intersecting street.

This part of the chapter consists of three sections. The first section describes a general model for extending AMFs developed for one street to the intersecting street. The second section extends

the general model to allow the extrapolation of AMFs developed for one intersection leg to other intersection legs. The last section demonstrates an application of the general model to turn lane installation.

General AMF Model for Street Treatments

The general AMF model is developed for the situation where an AMF is developed to estimate the relationship between "intersection crash frequency" and a change that is made to one of the two intersecting streets. In this context, "intersection crash frequency" is meant to include all crashes that occur at the intersection plus those that occur in the immediate vicinity of the intersection and which are identified as "intersection related."

Intuitively, a treatment applied to one street (e.g., increase lane width) should not have a measurable correlation with crashes associated with the intersecting street. However, if the AMF is derived to estimate the effect of a change in one street (e.g., increase lane width on major street) on intersection crash frequency, then this AMF cannot be used to estimate the effect of a similar change on the cross street without the risk of double-counting the effect of the treatment. A rational solution to this problem is described in this section.

General Model

A treatment applied to one intersecting street will have its most significant effect on crashes associated with traffic on that street. The magnitude of this effect can be quantified using an AMF that is based on just the crashes associated with the treated street (i.e., AMF_{str}). Alternatively, this effect can also be quantified using an AMF based on intersection crashes (i.e., AMF_{int}). The relationship between these two AMFs is:

$$AMF_{int,i} = AMF_{str} P_{str,i} + 1.0 (1 - P_{str,i}) \quad \textbf{(6)}$$

where,

$AMF_{int,\,i}$ = AMF for a specified treatment to street i, quantified in terms of intersection crashes (i = 1 for major street or 2 for minor street);

AMF_{str} = AMF for a specified treatment to any street, quantified in terms of the crashes that occur on the subject street; and

$P_{str,\,i}$ = proportion of intersection crashes that occur on treated street i.

Using Equation 4, the expected crash frequency associated with street i can be computed as:

$$E[N]_{str,i} = 0.000365\ Base\ ADT_i \quad \textbf{(7)}$$

where,

ADT_i = average daily traffic for street i (i = 1 for major street or 2 for minor street), veh/d.

The proportion of intersection crashes that occur on street i can be estimated by dividing Equation 7 by Equation 4 to yield:

$$P_{str,i} = \frac{ADT_i}{ADT_1 + ADT_2} \quad \textbf{(8)}$$

Equation 6 (with Equation 8) represents the "General AMF Model for Street Treatments." It recognizes that the effect of a treatment is more accurately quantified in terms of AMF_{str}. The effect of the treatment on intersection crashes is a function of the ADT on the treated street.

The general model is most appropriate for treatments that occur along the street, on both sides of the intersection. Changes in lane width, shoulder width, or median width are examples of such treatments. A model is described in the next section that is applicable to treatments that occur on a specific intersection leg.

It can be argued that the value of the street-specific AMF (i.e., AMF_{str}) is not sensitive to which street is treated. In fact, if both streets are treated, then the intersection AMF (i.e., $AMF_{int,1,2}$) can be computed as:

$$AMF_{int,1,2} = \left[AMF_{str} P_{str,1} + 1.0 (1 - P_{str,1})\right] \times \left[AMF_{str} P_{str,2} + 1.0 (1 - P_{str,2})\right] \quad \textbf{(9)}$$

Extension of General Model

As noted in the previous part of this chapter, AMFs are available in the *Workbook* that are derived for the major street (but not the minor street). These AMFs were quantified in terms of their impact to intersection crash frequency and, thereby, represent the dependent variable (i.e., $AMF_{int,1}$) in Equation 6. In this instance, the following form of Equation 6 holds.

$$AMF_{int,1} = AMF_{str} P_{str,1} + 1.0 (1 - P_{str,1}) \quad \textbf{(10)}$$

Unfortunately, the proportion of intersection crashes that occur on the subject street $P_{str,1}$ is not described with these AMFs, so AMF_{str} cannot be estimated. The remainder of this subsection describes how Equation 10 can be used to convert an AMF derived for the major street to an AMF for the minor street (where the AMF for the major street is based on intersection crash frequency).

Equation 6 can also be written in the following form when the AMF is derived for the minor street (but not the major street):

$$AMF_{int,2} = AMF_{str} P_{str,2} + 1.0 (1 - P_{str,2}) \quad \textbf{(11)}$$

Equations 10 and 11 can be combined to obtain an equation for estimating an AMF for the minor street by extrapolating an AMF that was derived for the major street.

$$AMF_{int,2} = \left(AMF_{int,1} - 1\right) \frac{P_{str,2}}{P_{str,1}} + 1.0 \quad \textbf{(12)}$$

Equation 12 is useful for estimating an AMF for the minor street when the only available AMF is for the major street. Inspection of the equation indicates that it adheres to logical boundary conditions (e.g., $AMF_{int,2} = 1.0$ when $P_{str,2} = 0$; and $AMF_{int,2} = AMF_{int,1}$ when $P_{str,2} = P_{str,1}$).

General AMF Model for Leg Treatments

The framework developed in the previous section is extended in this section to the development of a general AMF model when the treatment is specific to an intersection leg (as opposed to one of the intersecting streets). This type of treatment could be represented by the installation of a turn lane, a channelized right-turn, or the addition of a driveway. The form of this model is:

$$AMF_{int,j} = AMF_{leg} P_{leg,j} + 1.0 (1 - P_{leg,j}) \quad \textbf{(13)}$$

$$P_{leg,j} = \frac{ADT_j}{ADT_1 + ADT_2 + ADT_3 + ADT_4} \quad \textbf{(14)}$$

where,

$AMF_{int,j}$ = AMF for a specified treatment to leg j, quantified in terms of intersection crashes (j = 1 for the highest volume major-street leg, 2 for the other major-street leg, 3 for the highest volume minor-street leg, and 4 for the other minor-street leg);
AMF_{leg} = AMF for a specified treatment to any leg, quantified in terms of the crashes that occur on the subject leg;
$P_{leg,j}$ = proportion of intersection crashes that occur on treated leg j; and
ADT_j = average daily traffic for leg j, veh/d.

Equations 13 and 14 are equally applicable to three- or four-leg intersections. However, if applied to a three-leg intersection, ADT_4 is set equal to zero (0).

A rational argument can be made that the value of the leg-specific AMF (i.e., AMF_{leg}) is not sensitive to which leg is treated or the number of legs at the intersection. In fact, if two legs are treated, then the intersection AMF (i.e., $AMF_{int,j,k}$) can be computed as:

$$AMF_{int,j,k} = \left[AMF_{leg} P_{leg,j} + 1.0 (1 - P_{leg,j})\right] \times \left[AMF_{leg} P_{leg,k} + 1.0 (1 - P_{leg,k})\right] \quad \textbf{(15)}$$

Equation 13 can be used in this multiplicative manner to address treatments made to any number of approach legs. One term is included in Equation 15 for each leg treated.

The applicability of Equation 15 to a given treatment must reflect consideration of leg traffic control and the manner in which it may influence the AMF_{leg} value. For example, the value of AMF_{leg} for left-turn lane installation on an uncontrolled leg of an intersection with two-way stop control is not likely to be the same as AMF_{leg} for left-turn lane installation on a stop-controlled leg. Thus, the number of terms included in Equation 15 should be dependent on the number of legs that are treated and for which the leg-specific AMF is applicable.

Development of Generalized Turn Lane AMF Model

This section describes an application of the general AMF model for leg treatments to the AMFs developed by Harwood et al. (*3*) for turn lanes. They developed AMFs for left-turn lanes and right-turn lanes installed individually and in pairs at signalized and unsignalized intersections. The

intersections were located in urban and rural areas. Some intersections had three legs and others had four legs. Finally, some intersections had signal control installed at the time of the lane installation (they were referred to as "newly signalized" intersections).

The database assembled by Harwood et al. (*3*) represents 280 treated intersections (143 rural and 137 urban intersections) collectively located in eight states. At these intersections, 392 left-turn lanes were added and 182 right-turn lanes were added.

Harwood et al. (*3*) developed the AMFs using before-after data. The AMFs were quantified using three methods. They determined that the AMFs derived using the empirical Bayes (EB) method were the most accurate. They also determined that those AMFs derived from the comparison group (CG) method were acceptable when statistically significant results were not obtained from the EB method. They also considered using the AMFs derived using a yoked comparison method when the findings from the EB and CG methods were not acceptable.

The objective of the analysis described in the remainder of this section is to quantify the leg-specific AMFs (i.e., AMF_{leg}) for the left-turn lane and right-turn lane treatments at signalized, newly signalized, and unsignalized intersections in urban and rural settings using the data reported by Harwood et al. (*3*). These combinations represent a maximum of 12 AMFs (= 2×3×2).

Model Development

A series of equations were constructed from the generalized model structure described in the previous subsection. These equations included the desired leg-specific AMFs as independent variables. The form of these models is shown below.

I3L - Left-Turn Lane Added to Three-Leg Intersection - Intersection Crash Basis

$$\hat{AMF}_{int,unsig,left} = AMF_{leg,unsig,left} P_{leg,1} + 1.0\,(1 - P_{leg,1}) \quad \textbf{(16)}$$

I4L - Left-Turn Lane Added to Four-Leg Intersection - Intersection Crash Basis

$$\hat{AMF}_{int,unsig,left} = \left[AMF_{leg,unsig,left} P_{leg,1} + 1.0\,(1 - P_{leg,1})\right] \times \left[AMF_{leg,unsig,left} P_{leg,2} + 1.0\,(1 - P_{leg,2})\right]^{(N_{left\ lanes} - 1)} \quad \textbf{(17)}$$

I4LR - Left- and Right-Turn Lane Added to Four-Leg Intersection - Intersection Crash Basis

$$\hat{AMF}_{int,unsig\,left+right} = \left[AMF_{leg,unsig,left} P_{leg,1} + 1.0\,(1 - P_{leg,1})\right] \times \left[AMF_{leg,unsig,left} P_{leg,2} + 1.0\,(1 - P_{leg,2})\right]^{(N_{left\ lanes} - 1)} \times \left[AMF_{leg,unsig,right} P_{leg,1} + 1.0\,(1 - P_{leg,1})\right] \times \left[AMF_{leg,unsig,right} P_{leg,2} + 1.0\,(1 - P_{leg,2})\right]^{(N_{right\ lanes} - 1)} \quad \textbf{(18)}$$

I3R - Right-Turn Lane Added to Three-Leg Intersection - Intersection Crash Basis

$$\hat{AMF}_{int,unsig,right} = AMF_{leg,unsig,right} P_{leg,1} + 1.0\,(1 - P_{leg,1}) \quad \textbf{(19)}$$

I4R - Right-Turn Lane Added to Four-Leg Intersection - Intersection Crash Basis

$$\hat{AMF}_{int,unsig,right} = \left[AMF_{leg,unsig,right} P_{leg,1} + 1.0\,(1 - P_{leg,1})\right] \times \left[AMF_{leg,unsig,right} P_{leg,2} + 1.0\,(1 - P_{leg,2})\right]^{(N_{right\ lanes} - 1)} \quad \textbf{(20)}$$

A3L - Left-Turn Lane Added to Three-Leg Intersection - Approach Leg Crash Basis

$$\hat{AMF}_{leg,unsig,left} = AMF_{leg,unsig,left} \quad \textbf{(21)}$$

A4L - Left-Turn Lane Added to Four-Leg Intersection - Approach Leg Crash Basis

$$\hat{AMF}_{leg,unsig,left} = AMF_{leg,unsig,left} \quad \textbf{(22)}$$

A4LR - Left- and Right-Turn Lane Added to Four-Leg Intersection - Approach Leg Crash Basis

$$\hat{AMF}_{leg,unsig,left+right} = AMF_{leg,unsig,left} \times AMF_{leg,unsig,right} \quad \textbf{(23)}$$

A3R - Right-Turn Lane Added to Three-Leg Intersection - Approach Leg Crash Basis

$$\hat{AMF}_{leg,unsig,right} = AMF_{leg,unsig,right} \quad \textbf{(24)}$$

A4R - Right-Turn Lane Added to Four-Leg Intersection - Approach Leg Crash Basis

$$\hat{AMF}_{leg,unsig,right} = AMF_{leg,unsig,right} \quad \textbf{(25)}$$

where,

$\hat{AMF}_{l,m,n}$ = predicted AMF for crash location l, with control type m, and project type n (l = "int" for intersection crashes, and "leg" for crashes on the leg; m = "unsig" for an unsignalized intersection, "sig" for a signalized intersection, and "new" for newly signalized; n = "left" for left-turn lane installation, "right" for right-turn lane installation, and "left+right" when both lanes were added to a leg);

$AMF_{leg,m,n}$ = estimated leg-specific AMF for project type n applied to the leg of an intersection with control type m (used as a calibration coefficient in the regression model);

$P_{leg,j}$ = proportion of intersection crashes that occur on treated leg j (j = 1 for the highest volume major-street leg, 2 for the other major-street leg, 3 for the highest volume minor-street leg, and 4 for the other minor-street leg);

$N_{left\ lanes}$ = number of legs on the major street with a left-turn lane (= 1, 2); and

$N_{right\ lanes}$ = number of legs on the major street with a right-turn lane (= 1, 2).

Preliminary versions of Equations 16 through 25 included a fourth subscript with the leg-specific AMF variable (i.e., $AMF_{x,leg,m,n}$). This fourth subscript was used to designate the number of legs x at the intersection at which the turn lane was installed. However, preliminary statistical analysis indicated that there was no significant difference in the value of the leg-specific AMFs when the lane was added to a four-leg intersection or a three-leg intersection. For this reason, the subscript was not added, and the value of $AMF_{leg,m,n}$ for a given project type and control type was defined to be the same, regardless of the number of legs at the intersection. This treatment is reflected in all of the Equations 16 through 25 but it is most notable in Equations 21 and 22, as well as Equations 24 and 25, where the two equations in each pair are the same, regardless of the specified number of legs.

The proportion of intersection crashes $P_{leg,j}$ was estimated using Equation 14 with the average ADTs reported by Harwood et al. (*3*, Table 17) for each combination of area type (i.e., urban or rural) and intersection control type.

Equations 16 through 25 are derived to represent turn lane installations at unsignalized intersections. This sensitivity is most notable in the formulation of Equations 16 through 20 where the terms shown only reflect lanes being added to the uncontrolled major-street legs. This approach reflects the fact that Harwood et al. (*3*) only collected data for unsignalized intersections at which turn lane modifications were made to one or both of the uncontrolled major-street legs. In contrast, they collected data for both the major-street and the minor-street approach legs at signalized intersections. To accommodate signalized intersections, Equations 16 through 20 were expanded to include additional terms (much like those terms shown) to reflect the installation of turn lanes on up to four legs.

Although not shown in the subscripts of Equations 16 through 25, these equations were used to separately evaluate intersections in urban areas and those in rural areas.

Model Calibration

The equations developed in the previous section were calibrated using the crash reduction factors for fatal and injury crashes provided in Appendix C of the report by Harwood et al. (*3*). Specifically, the "mean" crash reduction factors CRF_m applicable to "Total Intersection" crashes (referred to herein as "intersection" crashes) and to "Intersection Approach" crashes (referred to herein as "leg-specific" crashes) were used. The "mean" factors represent the average crash reduction per treated site. The "mean per added lane" factors CRF_{mpl} provided in the report were used with the corresponding "standard error per added lane" S_{mpl} values to compute the standard error of the "mean" factor S_m (i.e., $S_m = S_{mpl} \times CRF_m / CRF_{mpl}$).

As noted previously, Harwood et al. (*3*) developed factors using the empirical Bayes (EB), comparison group (CG), and yoked comparison analysis methods. Of these, they gave the EB and CG methods greater consideration than the yoked comparison methods. Hence, only the factors derived using the EB and CG methods were used for model calibration. The standard error for those factors derived from the CG method were multiplied by a "method correction factor" of 2.0, based on the guidance provided by Bahar et al. (*4*).

The mean crash reduction factors reported by Harwood et al. (*3*) were converted into AMFs using the following equation:

$$AMF = 1.0 + \frac{CRF_m}{100} \quad \textbf{(26)}$$

The AMFs obtained from Equation 26 were used as the dependent variable in a regression analysis. A search algorithm was used to simultaneously evaluate all of the equations (described in the previous subsection), compare the AMFs from the Harwood data with $\hat{AMF}_{l,m,n}$ from the prediction equations, and find the value of the leg-specific AMFs (i.e., $AMF_{leg,m,n}$) that minimized the sum of the squared error. This algorithm was automated using the Nonlinear Regression (NLIN)

procedure in the SAS software (*5*). The SAS procedure was coded to minimize the weighted squared error, where the weight for each observation was equal to the reciprocal of the squared standard error (i.e., weight = $1/S_m^2$).

The regression analysis of the AMFs indicated that those applicable to rural intersections were consistently smaller than those applicable to urban intersections. Further examination indicated that the rural AMFs were approximately equal to the square of the urban AMFs, for common combinations of control type and project type. Based on this finding, the same set of equations were used for both the urban and rural AMFs with the exception that the following relationship was substituted in each equation for $AMF_{leg,m,n}$ when applied to an AMF derived from a rural intersection:

$$AMF_{leg,m,n} = (AMF^*_{leg,m,n})^{1+(b_0-1)I_{rural}} \quad (27)$$

where,

$AMF^*_{leg,m,n}$ = estimated AMF for project type n applied to the leg of an *urban* intersection with control type m;
b_o = calibration coefficient; and
I_{rural} = indicator variable (= 1.0 if intersection is rural, 0.0 if it is urban).

Equation 27 reduces the number of equations evaluated by the regression algorithm by one-half (i.e., both urban and rural equations are not needed) but still allows for a relationship to exist between the calibrated urban and rural AMFs.

The results of the model calibration are presented in Table 2-6. Calibration of this model focused on AMFs for injury (plus fatal) crash frequency. The R^2 for the model is 0.53. This statistic indicates that about 53 percent of the variability in the AMF data is explained by the model.

Figure 2-1 provides a graphical indication of the fit of the model to the reported AMFs. The figure compares the residual error of the predicted AMF over the range of predicted AMFs. Each residual error was "standardized" by dividing it by its standard error S_m. Each data point represents one of the 73 AMF analysis method, crash location, area-type, traffic control, and project-type combinations reported by Harwood et al. (*3*). The data are centered on "0" which indicates that there is no bias in the predicted AMFs. The scatter in the data indicates that almost all observations are within three standard deviations. The fit would likely improve if the ADT for each observation was available, instead of the average values that were used (as reported by Harwood et al. [*3*, Table 17]). The two residuals that exceed four standard deviations were associated with AMF estimates that had very small standard errors (and thus exceptionally large weight). Access to the data reported by Harwood et al. (*3*) would be needed to determine why these two values are so deviant.

The AMFs reported in Table 2-6 for urban intersections are repeated in Table 2-7. The rural AMF exponent was used with Equation 27 to compute the AMFs for rural intersections. The AMFs in this table are applicable to leg-specific, injury (plus fatal) crashes. They can be used with Equations 13 or 15 to estimate the relationship between of turn lane presence and intersection crash frequency.

Table 2-6. Turn Lane AMF Model Statistical Description.

Model Statistics		Value		
R^2:		0.53		
Observations n_o:		73 AMFs (from 248 intersections and 440 legs)		
Standard Deviation s_e:		±0.10		
Calibrated Coefficient Values				
Variable	**Definition**	**Value**	**Std. Dev.**	**t-statistic [1]**
b_0	Rural AMF exponent	1.916	0.252	-3.6
$AMF_{leg,sig,left}$	Leg-specific AMF for left-turn lane installation at signalized urban intersection	0.654	0.011	31.5
$AMF_{leg,sig,right}$	Leg-specific AMF for right-turn lane installation at signalized urban intersection	0.758	0.024	10.3
$AMF_{leg,new,left}$	Leg-specific AMF for left-turn lane installation at newly signalized urban intersection	0.573	0.031	13.9
$AMF_{leg,new,right}$	Leg-specific AMF for right-turn lane installation at newly signalized urban intersection	0.779	0.066	3.4
$AMF_{leg,unsig,left}$	Leg-specific AMF for left-turn lane installation at unsignalized urban intersection (uncontrolled)	0.589	0.041	10.0
$AMF_{leg,unsig,right}$	Leg-specific AMF for right-turn lane installation at unsignalized urban intersection (uncontrolled)	0.866	0.039	3.4

Note:
1 - Test of null hypothesis that coefficient value is equal to 1.0.

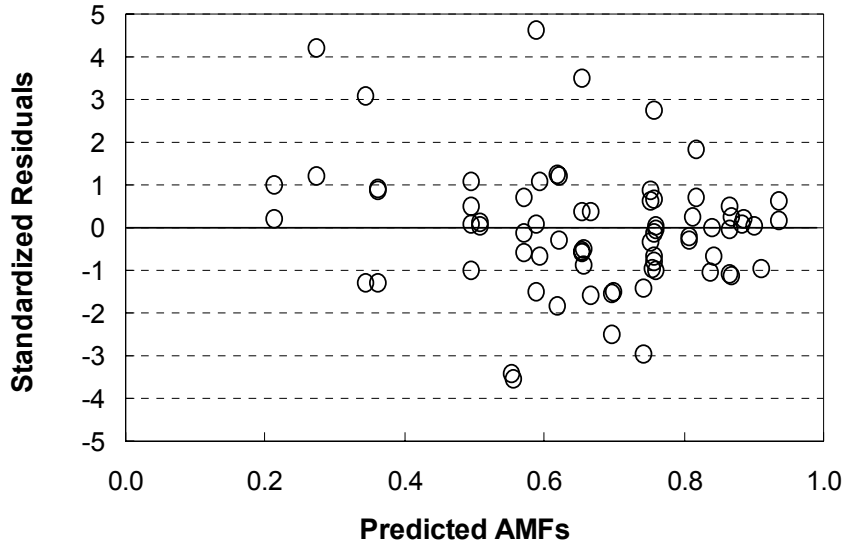


Figure 2-1. Comparison of Predicted AMFs with Standardized Residuals.

Table 2-7. Recommended Leg-Specific Turn Lane AMFs.

Area Type	Control Type	Left-Turn Lane	Right-Turn Lane
Rural [1]	Unsignalized [2]	0.36	0.76
	Signalized	0.44	0.59
	Newly Signalized	0.34	0.62
Urban (and Suburban)	Unsignalized [2]	0.59	0.87
	Signalized	0.65	0.76
	Newly Signalized	0.57	0.78

Notes:
1 - Rural AMFs computed from urban AMF for common control type and turn lane configuration using the following equation $AMF_{rural} = (AMF_{urban})^{1.916}$.
2 - AMFs listed are for the uncontrolled approach at an unsignalized intersection.

To illustrate the use of the AMFs in Table 2-7, consider the installation of a left-turn lane on one approach of a four-leg urban intersection with the same ADT on all legs. The intersection is signal controlled. Table 2-7 indicates the AMF for leg-specific crashes is 0.65. Thus, if there are 10 crashes/yr on the leg before the lane is installed, the expected crash frequency after installation is 6.5 crashes/yr (= 10 ×0.65), which corresponds to a reduction of 3.5 crashes/yr.

Continuing with the same example, the leg-specific AMF of 0.65 can be converted into an equivalent AMF for intersection crashes using Equation 13. Based on Equation 14, the proportion of intersection crashes for the treated leg $P_{leg,j}$ is equal to 0.25 because each leg has the same ADT in this example. This calculation yields an AMF of 0.91 (= 0.65 ×0.25 + 1.0 [1 - 0.25]). Thus, if the intersection experiences 40 crashes/yr, the expected crash frequency after installation is 36.5 crashes/yr (= 40 ×0.65), which corresponds to a reduction of 3.5 crashes/yr.

Consider the installation of a left-turn lane and a right-turn lane on one approach of a four-leg urban intersection with the same ADT on all legs. The intersection is signal controlled. Table 2-7 indicates the AMF for leg-specific crashes is 0.65 and 0.76 for left-turn and right-turn lanes, respectively. The combined AMF for both lanes is equal to the product of these two AMFs, or 0.49. Thus, if there are 10 crashes/yr on the leg before the lanes are installed, the expected crash frequency after installation is 4.9 crashes/yr (= 10 ×0.49), which corresponds to a reduction of 5.1 crashes/yr.

Continuing with the same example, the leg-specific AMFs of 0.65 and 0.76 can be converted into an equivalent AMF for intersection crashes using Equation 15. Based on Equation 14, the proportion of intersection crashes for the treated leg $P_{leg,j}$ is equal to 0.25 because each leg has the same ADT in this example. This calculation yields an AMF of 0.86 (={0.65 ×0.25 + 1.0 [1 - 0.25]} ×{0.76 ×0.25 + 1.0 [1 - 0.25]}). Thus, if the intersection experiences 40 crashes/yr, the expected crash frequency after installation is 34.3 crashes/yr (= 40 × 0.86), which corresponds to a reduction of 5.7 crashes/yr. This same result would be realized if the lanes were implemented on opposing approaches, provided that the leg ADTs were the same. If they are not the same, then Equation 14 would be used to compute $P_{leg,j}$ for each leg, and these two proportions would be used in Equation 15 to obtain the combined AMF.

METHODOLOGY

This part of the chapter describes the methodology used to recalibrate the urban intersection safety prediction models described in the *Workbook* (*1*). The objective of this activity was to adjust one or more model coefficients such that the resulting model provides an accurate estimate of the frequency of crashes at intersections in Texas.

This part is divided into three sections. The first section describes the data used to recalibrate the intersection models. The second section describes the technique used to quantify the calibration coefficients. The last section describes the technique used to adjust the over-dispersion parameter for bias due to a small sample size. The goodness-of-fit statistics used to assess the accuracy of the recalibrated models are described in Reference 6.

Database

The database assembled from the TRM data for intersection model recalibration can be described as a "full" database. In this type of database, all of the relevant intersection characteristics are represented as attributes in the database. A "base" database could not be formed from the full database because several of the desired intersection characteristics are not included in the TRM data. A more complete discussion of the differences between full and base databases, as related to model recalibration, is described by Bonneson et al. (*6*).

The full database was assembled using data from the TRM database combined with data obtained from aerial photography. Aerial photos were acquired for the time period coincident with the crash data (i.e., circa 2000) and used to determine turn lane presence and intersection traffic control. They were also used to quantify the number of driveways near the intersection and the presence of right-turn channelization. Given the time-consuming nature of data collection from aerial photos, it was not feasible to selectively acquire only data for "typical" intersections. Thus, a base database could not be formed using manual methods because of time limitations.

Recalibration of the base model component of a safety prediction model using a full database requires the use of AMFs. Specifically, the AMFs in the *Workbook* (*1*) that are associated with the design elements in the database were used to compute appropriate AMF values for each intersection. These values are then used in the full model during recalibration such that only the values for a, b_1, and b_2 in Equation 2 are adjusted.

Recalibration Technique

The recalibration activity used statistical analysis software that employs maximum likelihood methods based on a negative binomial distribution of crash frequency. In the initial analysis, the constant a, ADT exponents b_1 and b_2, and over-dispersion parameter k were quantified. If the ADT exponents were found to be significantly different from their counterparts in the existing model (shown in Table 2-3), then the new values were used in replacement of the existing coefficients. If these exponents were not significantly different, then the coefficients in the existing model were

retained. Regardless of the decision about the ADT exponents, a new value for *a* was computed and used in the recommended model.

Over-Dispersion Parameter Adjustment

It was assumed that crash occurrence at an intersection is Poisson distributed, and that the distribution of the mean crash frequency for a group of similar intersections is gamma distributed. In this manner, the distribution of crashes for a group of similar intersections can be described by the negative binomial distribution. The variance of this distribution is:

$$V[X] = y\,E[N] + \frac{(y\,E[N])^2}{k} \quad \textbf{(28)}$$

where,

X = reported crash count for y years, crashes;
y = time interval during which X crashes were reported, yr; and
k = over-dispersion parameter.

Research by Lord (7) has indicated that databases with low sample mean values and small sample size may not exhibit as much variability as suggested by Equation 28. Rather, the regression model may "over-explain" some of the random variability in the small database, or the low sample mean may introduce an instability in the model coefficients (7). These issues are especially present in model recalibration because of the inherent use of small databases and the desire to recalibrate several model coefficients. To minimize the adverse impact of these issues, the number of recalibrated coefficients is kept to a reasonable minimum.

Also in recognition of the concerns expressed by Lord (7), the over-dispersion parameter k obtained from the regression analysis is adjusted downward, such that it yields a more reliable estimate of the variance of the crash distribution. The derivation of the adjustment technique used is described in Reference *6*. The following equation was used to estimate the true over-dispersion parameter, given the estimated over-dispersion parameter obtained from the regression analysis.

$$k_t = \frac{\sqrt{(n-p)^2 m^2 + 69\,k_r (n-p)\,m} - (n-p)\,m}{34.5} \quad \textbf{(29)}$$

where,

k_t = true over-dispersion parameter;
k_r = estimated over-dispersion parameter obtained from database analysis;
n = number of observations (i.e., intersections in database);
p = number of model variables; and
m = average number of crashes per observation (= total crashes in database / n).

Equation 29 was used to estimate the true over-dispersion parameter for each of the models described in this chapter. All subsequent references to the over-dispersion parameter k in this chapter denote the estimated true parameter obtained from Equation 29 (i.e., hereafter, $k = k_t$).

CALIBRATION DATA AND ANALYSIS

This part of the chapter describes the recalibration of four intersection models. Initially, the process used to select the intersection sites is identified. Then, the characteristics of the sites are summarized. Next, the techniques used to assemble the crash data are described. Finally, the findings from the model calibration are discussed.

Site Selection and Data Collection

The intersections used to calibrate the safety prediction model were required to satisfy the following criteria:

- Area type: urban or suburban;
- Curve presence: radius of 5730 ft or larger on the major street at the intersection; and
- Intersection spacing: 650 ft or more between adjacent intersections.

All intersections were required to have both of the intersecting highways on the state highway system. This requirement ensured that data for both highways could be found in the TRM data. Also, the intersections selected were required to be within the regions of Texas for which high-resolution aerial photography was available on the Internet.

Guidance provided by Harwood et al. (*8*) was used to establish the desired sample size. Specifically, they indicate that a desirable minimum number of 25 four-leg signalized intersections are needed for model recalibration. They also suggest that the desirable minimum number of unsignalized three-leg intersections is 100, and the desirable minimum number of four-leg unsignalized intersections is 100. Based on typical crash frequencies for each of these intersection types, the minimum number of sites equates to a minimum of 150 reported crashes for each intersection type. Harwood et al. (*8*) did not define a desired minimum number of three-leg signalized intersections.

Site Characteristics

Traffic and geometry data were identified for several thousand highway intersections using the TRM data for year 2003, the first year for which a complete database was available. All total, 991 intersections were identified that satisfied the selection criteria (except the need for high-resolution aerial photography). A review of several sources of aerial photography on the Internet revealed that only about 50 percent of the identified intersections were in areas that the photography was of sufficient resolution that lane markings could be identified. Of those intersections that were in areas of high-resolution photography, about 40 percent were found to be undesirable for a variety of reasons (e.g., lane markings suggest a type of control that is contradictory to that found in the crash data, intersection was in a rural area, intersection had unusual geometry, etc.). As a result of the screening process, 246 intersections were identified that satisfied all of the site selection criteria. Selected characteristics are provided for these intersections in Table 2-8.

Table 2-8. Summary Characteristics for Urban and Suburban Intersections.

Legs	Control Type	Total Intersections	Major-Street Volume, veh/d		Minor-Street Volume, veh/d	
			Minimum	Maximum	Minimum	Maximum
Three	Signal	19	5,960	45,200	1,400	22,500
	Unsignalized [1]	38	1,720	23,000	560	10,400
Four	Signal	133	4,040	55,200	330	38,000
	Unsignalized [1]	56	400	27,600	270	6,200

Note:
1 - Unsignalized intersections have an uncontrolled major street and a stop-controlled minor street.

The number of unsignalized intersections listed in Table 2-8 does not satisfy the desired minimum number, as discussed in the previous section. The number of three-leg signalized intersections is also somewhat small. Nevertheless, as the discussion in a subsequent section will show, this number was believed sufficient to recalibrate the corresponding base models.

Data Collection

Crash data were identified for each intersection using the TxDOT highway system crash database. Three years of crash data, corresponding to years 1999, 2000, and 2001 were identified for each intersection. The ADT for each of these three years was obtained from the TRM database and averaged to obtain one ADT for each intersecting roadway. Only those crashes that occurred within 325 ft of the intersection, and that were identified as "intersection-related," were included in the database. A distance of 325 ft was used in recognition of the 0.1 mi crash location precision used by TxDOT. Use of a shorter distance significantly increased the probability that some intersection-related crashes would not be identified for any intersection.

Data Analysis

This section is divided into two subsections. The first subsection summarizes the crash data at the selected study sites. The second subsection describes the formulation of the calibration model and summarizes the statistical analysis methods used to calibrate it.

Database Summary

The crash data for each of the 246 intersections are summarized in Table 2-9. These intersections were associated with 3090 crashes, of which 2175 were injury or fatal crashes. The trends in crash rate are provided in the last two columns of the table. Although the number of unsignalized intersections is significantly below the desired minimum, the number of fatal-plus-injury crashes for each intersection type is very near to the desired minimum of 150 crashes.

Table 2-9. Crash Data Summary for Urban and Suburban Intersections.

Legs	Control Type	Exposure,[1] mev	Crashes / 3 years			Crash Rate, cr/mev	
			PDO [2]	I+F [3]	Total	I+F [3]	Total
Three	Signal	166.7	81	126	207	0.25	0.41
	Unsignalized	146.3	47	103	150	0.23	0.34
Four	Signal	1417.6	720	1802	2522	0.42	0.59
	Unsignalized	156.3	67	144	211	0.31	0.45
		Total:	915	2175	3090		

Notes:
1 - mev: million-entering-vehicles per year.
2 - PDO: property-damage-only crashes.
3 - I+F: injury plus fatal crashes.

Model Development and Statistical Analysis Methods

This section describes the intersection calibration model and the methods used to recalibrate it. The general form for this model is shown below:

$$E[N] = a\left(\frac{ADT_{major}}{1000}\right)^{b_1}\left(\frac{ADT_{minor}}{1000}\right)^{b_2} \times AMF_{LT} \times AMF_{RT} \times AMF_{SW} \times AMF_{LW} \times AMF_{MP} \qquad (30)$$

where,

$E[N]$ = expected number of crashes, crashes/yr;
a, b_1, b_2 = calibration coefficients;
AMF_{LT} = left-turn lane accident modification factor;
AMF_{RT} = right-turn lane accident modification factor;
AMF_{SW} = shoulder width accident modification factor;
AMF_{LW} = lane width accident modification factor; and
AMF_{MP} = median width accident modification factor.

As discussed in the section titled Methodology, AMFs must be included in the model during the recalibration process because a full database is used. In this manner, only the coefficients associated with the base model component of Equation 30 (i.e., a, b_1, b_2) are recalibrated. The AMFs listed in Equation 30 are specified in the following equations:

$$\begin{aligned} AMF_{LT} = {} & \left[AMF_{leg,lt} P_{leg,1} + 1.0\,(1 - P_{leg,1})\right]^{I_{lt,1}} \times \\ & \left[AMF_{leg,lt} P_{leg,2} + 1.0\,(1 - P_{leg,2})\right]^{I_{lt,2}} \times \\ & \left[AMF_{leg,lt} P_{leg,3} + 1.0\,(1 - P_{leg,3})\right]^{I_{lt,3}} \times \\ & \left[AMF_{leg,lt} P_{leg,4} + 1.0\,(1 - P_{leg,4})\right]^{I_{lt,4}} \end{aligned} \qquad (31)$$

where,

$AMF_{leg,\,lt}$ = leg-specific left-turn lane accident modification factor (from Table 2-7);

$P_{leg,j}$ = proportion of intersection crashes that occur on treated leg j (j = 1 for the highest volume major-street leg, 2 for the other major-street leg, 3 for the highest volume minor-street leg, and 4 for the other minor-street leg), (computed with Equation 14); and
$I_{lt,j}$ = indicator variable for leg j (= 1.0 if left-turn lane is present, 0.0 otherwise).

$$AMF_{RT} = \left[AMF_{leg,rt} P_{leg,1} + 1.0\,(1 - P_{leg,1})\right]^{I_{rt,1}} \times \left[AMF_{leg,rt} P_{leg,2} + 1.0\,(1 - P_{leg,2})\right]^{I_{rt,2}} \times \left[AMF_{leg,rt} P_{leg,3} + 1.0\,(1 - P_{leg,3})\right]^{I_{rt,3}} \times \left[AMF_{leg,rt} P_{leg,4} + 1.0\,(1 - P_{leg,4})\right]^{I_{rt,4}} \quad (32)$$

where,
$AMF_{leg,rt}$ = leg-specific right-turn lane accident modification factor (from Table 2-7); and
$I_{rt,j}$ = indicator variable for leg j (= 1.0 if right-turn lane is present, 0.0 otherwise).

$$AMF_{SW} = e^{b_{s,1}(W_{s,1} - 1.5)} \times \left[\left(e^{b_{s,1}(W_{s,2} - 1.5)} - 1\right)\frac{P_{str,2}}{P_{str,1}} + 1.0\right] \quad (33)$$

where,
$W_{s,1}$ = major-street shoulder width, ft;
$W_{s,2}$ = minor-street shoulder width, ft;
$b_{s,1}$ = major-street-specific shoulder width coefficient (= -0.020 for unsignalized intersections, unknown for signalized intersections); and
$P_{str,i}$ = proportion of intersection crashes that occur on treated street i (i = 1 for major street or 2 for minor street), (computed with Equation 8).

$$AMF_{LW} = e^{b_{l,1}(W_{l,1} - 12)} \times \left[\left(e^{b_{l,1}(W_{l,2} - 12)} - 1\right)\frac{P_{str,2}}{P_{str,1}} + 1.0\right] \quad (34)$$

where,
$W_{l,1}$ = major-street lane width, ft;
$W_{l,2}$ = minor-street lane width, ft; and
$b_{l,1}$ = major-street-specific lane width coefficient (= -0.057 for unsignalized intersections, -0.053 for signalized intersections).

$$AMF_{MP} = AMF_{base,1} \times e^{b_{m,1}(W_{m,1} - 16)} \times \left[\left(AMF_{base,2} \times e^{b_{m,1}(W_{m,2} - 16)} - 1\right)\frac{P_{str,2}}{P_{str,1}} + 1.0\right] \quad (35)$$

where,
$W_{m,1}$ = major-street median width, ft;
$W_{m,2}$ = minor-street median width, ft;
$AMF_{base,1}$ = base median presence AMF for major street (= 0.83 if raised curb without left-turn bay, 1.0 otherwise);
$AMF_{base,2}$ = base median presence AMF for minor street (= 0.83 if raised curb without left-turn bay and four legs, 0.0 if three legs, 1.0 otherwise); and
$b_{m,1}$ = major-street-specific median width coefficient (= 0.0076 if unsignalized with median width of 16 ft or more and three legs, 0.0160 if unsignalized with median width of 16 ft or more and four legs, 0.0 otherwise).

Equations 31 and 32 are based on the leg-specific AMFs in Table 2-7. They are shown with four terms and are directly applicable to four-leg signalized intersections. Only three of the four terms were used for the three-leg signalized intersections, and only two of the terms were used for the unsignalized intersections. Only two terms were used for unsignalized intersections because the AMFs in Table 2-7 only apply to turn lanes on uncontrolled approaches. No AMFs are currently available for turn lanes on stop-controlled approaches.

Equations 33, 34, and 35 were taken from the *Workbook*. In each case, the *Workbook* AMF is specific to the major street at an intersection. Equations 33, 34, and 35 show the extrapolation of these AMFs to both the major and minor streets by incorporating Equation 12. Equations 33 and 35 are applicable only to unsignalized intersections. The effect of shoulder width and median width at signalized intersections is unknown at this time.

During the preliminary recalibration analysis, two notable correlations were found with crash frequency. One correlation was with free right-turn channelization, and the other was with the number of through lanes on the intersecting streets. Two additional AMFs were added to the model to account for these influences. The coefficients in these additional AMFs were also calibrated during the recalibration process. The two AMFs are described in the following equations:

$$\begin{aligned} AMF_{CH} = {} & [AMF_{leg,ch,1} P_{leg,1} + 1.0 (1 - P_{leg,1})] \times \\ & [AMF_{leg,ch,2} P_{leg,2} + 1.0 (1 - P_{leg,2})] \times \\ & [AMF_{leg,ch,3} P_{leg,3} + 1.0 (1 - P_{leg,3})] \times \\ & [AMF_{leg,ch,4} P_{leg,4} + 1.0 (1 - P_{leg,4})] \end{aligned} \quad \textbf{(36)}$$

with,

$$AMF_{leg,ch,j} = e^{b_{ch,m} I_{ch,j}} \quad \textbf{(37)}$$

where,

AMF_{CH} = free right-turn channelization accident modification factor;
$AMF_{leg,ch}$ = leg-specific channelization accident modification factor for leg j;
$I_{ch,j}$ = indicator variable for leg j (= 1.0 if channelization is present, 0.0 otherwise); and
$b_{ch,m}$ = calibration coefficient for intersection control type m (= "unsig" for an unsignalized intersection and "sig" for signalized intersection).

$$AMF_{NL} = \left[AMF_{str,nl,1} P_{str,1} + 1.0 (1 - P_{str,1})\right] \times \left[AMF_{str,nl,2} P_{str,2} + 1.0 (1 - P_{str,2})\right] \quad \textbf{(38)}$$

with,

$$AMF_{str,nl,i} = e^{b_{nl,m} (N_i - N_{base,i})} \quad \textbf{(39)}$$

where,

AMF_{NL} = number-of-lanes accident modification factor;
$AMF_{str,nl,i}$ = street-specific number-of-lanes accident modification factor for street i (i = 1 for major street or 2 for minor street);
N_i = number of through lanes on street i;

$N_{base,i}$ = base number of through lanes on street i (= 4 lanes for a major street; 2 lanes for a minor street); and

$b_{nl,m}$ = calibration coefficient for intersection control type m (= "unsig" for an unsignalized intersection and "sig" for signalized intersection).

Equations 36 and 38 were separately calibrated to unsignalized and signalized intersections, as suggested by the calibration coefficients in companion Equations 37 and 39. When Equation 36 is applied to an unsignalized intersection, no distinction is made as to whether the AMF is applied to an uncontrolled leg or a stop-controlled leg. This approach is based on a preliminary analysis of the data that indicated that the measured effect of channelization on uncontrolled approaches is almost identical to that for stop-controlled approaches. Nevertheless, it is noted that this approach is in contrast to that used for the turn lane AMFs, where the turn lane AMFs for uncontrolled approaches are reasoned to not be applicable to stop-controlled approaches.

Equation 36, as shown, is applicable to four-leg intersections. Only three terms should be used when it is applied to intersections with three legs.

The Nonlinear Regression procedure (NLIN) in the SAS software was used to estimate the model coefficients (5). This procedure was used because the safety prediction model is both nonlinear and discontinuous. The loss function associated with NLIN was specified to equal the log likelihood function for the negative binomial distribution. Equation 28 was used to define the variance function. The procedure was set up to estimate model coefficients based on maximum-likelihood methods.

One disadvantage of the NLIN procedure is that it is not able to compute the best-fit value of k for the enhanced model. This disadvantage is overcome by using the Generalized Modeling (GENMOD) procedure in SAS. GENMOD automates the over-dispersion factor estimation process using maximum-likelihood methods. Initially, NLIN is used to calibrate the prediction model. Then, GENMOD is used to regress the relationship between the reported and predicted crash frequencies (where the natural log of the predicted values is specified as an offset variable, and the log link function is used) to obtain a best-fit over-dispersion estimate a using its internal variance function ($V[X] = E[N] + a\,E[N]^2$). The estimate of a from GENMOD is then inverted to estimate a new value of k (i.e., $k_{new} = 1/a$). This new value is then used in a second application of NLIN and the process is repeated until convergence is achieved between the k value used in NLIN and that obtained from GENMOD. Convergence is typically achieved in two iterations. As a last step, the value of k obtained from this iterative method was used in Equation 29 to estimate the true over-dispersion parameter.

Model Calibration

This section describes the results of the model recalibration analysis. For each of the four intersection configuration and control-type combinations considered, the ADT coefficients for the recalibrated model (i.e., b_1, b_2) were not significantly different from those in the existing model (as listed previously in Table 2-3). For this reason, the ADT coefficients in the existing model were retained, and only coefficient a was recalibrated for each model.

A preliminary recalibration of each model combination indicated that the measured effect of channelization and number-of-lanes was very similar at the unsignalized intersections for both the three-leg and four-leg configurations. A similar finding was obtained in the initial recalibration of the signalized intersections. Based on this finding, the three-leg and four-leg intersection data were combined and used to calibrate a combined model (and selected AMFs) for a given control type. The following model variation of Equation 30 was used for this purpose:

$$E[N] = \left(E[N_3]\ I_3\ +\ E[N_4]\ (1.0 - I_3)\right) \times AMF_{CH} \times AMF_{LN} \quad \textbf{(40)}$$

with,

$$E[N_3] = a_3 \left(\frac{ADT_{major}}{1000}\right)^{b_{1,3}} \left(\frac{ADT_{minor}}{1000}\right)^{b_{2,3}} \times AMF_{LT} \times AMF_{RT} \times AMF_{SW} \times AMF_{LW} \times AMF_{MP} \quad \textbf{(41)}$$

$$E[N_4] = a_4 \left(\frac{ADT_{major}}{1000}\right)^{b_{1,4}} \left(\frac{ADT_{minor}}{1000}\right)^{b_{2,4}} \times AMF_{LT} \times AMF_{RT} \times AMF_{SW} \times AMF_{LW} \times AMF_{MP} \quad \textbf{(42)}$$

where,

$E[N_m]$ = expected number of crashes at intersection with m legs, crashes/yr;
a_m = calibration coefficient for intersection with m legs;
$b_{1,m}$, $b_{2,m}$ = ADT exponents for intersection with m legs (from Table 2-3); and
I_3 = indicator variable for intersection legs (= 1.0 if three legs, 0.0 if four legs).

The results of the recalibration analysis are presented in four subsections. The first subsection describes the results for three-leg signalized intersections. The second subsection describes the results for three-leg unsignalized intersections. The third subsection describes the results for four-leg signalized intersections. The last subsection describes the results for four-leg unsignalized intersections.

Three-Leg Signalized Intersection

The results of the three-leg signalized intersection model calibration are presented in Table 2-10. Calibration of this model focused on injury (plus fatal) crash frequency. The Pearson χ^2 statistic for the model is 17.2, and the degrees of freedom are 16 (= $n - p$ = 19 – 3). As this statistic is less than $\chi^2_{0.05,\,16}$ (= 26.3), the hypothesis that the model fits the data cannot be rejected. The R^2 for the model is 0.54. An alternative measure of model fit that is better suited to the negative binomial distribution is R_k^2 (9). The R_k^2 for the calibrated model is 0.65. This statistic indicates that about 65 percent of the variability due to systematic sources is explained by the model.

Table 2-10. Model Statistical Description–Urban Three-Leg Signalized Intersection.

Model Statistics		Value		
R^2 (R_k^2):		0.54 (0.65)		
Scale Parameter φ:		0.96		
Pearson χ^2:		17.2 ($\chi^2_{0.05,\,16}$ = 26.3)		
Over-Dispersion Parameter k:		2.89		
Observations n_o:		19 intersections (126 injury + fatal crashes in 3 years)		
Standard Deviation s_e:		±1.24 crashes/yr		
Range of Model Variables				
Variable	**Variable Name**	**Units**	**Minimum**	**Maximum**
ADT_{major}	Major-street ADT	veh/d	5960	45,200
ADT_{minor}	Minor-street ADT	veh/d	1400	22,500
I_{ch}	Major-street channelization	legs with presence	0	1
I_{ch}	Minor-street channelization	legs with presence	0	1
N	Major-street number of through lanes	lanes	2	6
N	Minor-street number of through lanes	lanes	2	4
Calibrated Coefficient Values				
Variable	**Definition**	**Value**	**Std. Dev.**	**t-statistic**
a	Constant	0.159	0.025	6.4
b_{ch}	Channelization AMF	0.248	0.107	2.3
b_{nl}	Number-of-lanes AMF coefficient	0.197	0.053	3.7

The recommended model is:

$$E[N]_{3LSG} = 0.159\left(\frac{ADT_{major}}{1000}\right)^{0.629}\left(\frac{ADT_{minor}}{1000}\right)^{0.385} \quad \textbf{(43)}$$

where,

$E[N]_{3LSG}$ = expected crash frequency for a three-leg signalized intersection, crashes/yr.

The ADT coefficients in Equation 43 were obtained from Table 2-3. The AMF coefficients are explained in a subsequent section.

Three-Leg Unsignalized Intersection

The results of the three-leg unsignalized intersection model calibration are presented in Table 2-11. Calibration of this model focused on injury (plus fatal) crash frequency. The Pearson χ^2 statistic for the model is 33.2, and the degrees of freedom are 35 (= $n - p$ = 38 – 3). As this statistic is less than $\chi^2_{0.05,\,35}$ (= 49.8), the hypothesis that the model fits the data cannot be rejected. The R^2 for the model is 0.38. The R_k^2 for the calibrated model is 0.57. This statistic indicates that about 57 percent of the variability due to systematic sources is explained by the model.

Table 2-11. Model Statistical Description–Urban Three-Leg Unsignalized Intersection.

Model Statistics		Value		
R^2 (R_k^2):		0.38 (0.57)		
Scale Parameter φ:		0.90		
Pearson χ^2:		33.2 ($\chi^2_{0.05,\,35}$ = 49.8)		
Over-Dispersion Parameter k:		1.20		
Observations n_o:		38 intersections (103 injury + fatal crashes in 3 years)		
Standard Deviation s_e:		±0.96 crashes/yr		
Range of Model Variables				
Variable	**Variable Name**	**Units**	**Minimum**	**Maximum**
ADT_{major}	Major-street ADT	veh/d	1720	23,000
ADT_{minor}	Minor-street ADT	veh/d	400	27,600
I_{ch}	Major-street channelization	legs with presence	0	1
I_{ch}	Minor-street channelization	legs with presence	0	1
N	Major-street number of through lanes	lanes	2	4
N	Minor-street number of through lanes	lanes	2	3
Calibrated Coefficient Values				
Variable	**Definition**	**Value**	**Std. Dev.**	**t-statistic**
a	Constant	0.088	0.018	4.9
b_{ch}	Channelization AMF	0.937	0.225	4.2
b_{nl}	Number-of-lanes AMF coefficient	-0.135	0.127	-1.1

The recommended model is:

$$E[N]_{3LST} = 0.088\left(\frac{ADT_{major}}{1000}\right)^{0.766}\left(\frac{ADT_{minor}}{1000}\right)^{0.248} \quad \textbf{(44)}$$

where,

$E[N]_{3LST}$ = expected crash frequency for a three-leg unsignalized intersection, crashes/yr.

The ADT coefficients in Equation 44 were obtained from Table 2-3. The AMF coefficients are explained in a subsequent section.

Four-Leg Signalized Intersection

The results of the four-leg signalized intersection model calibration are presented in Table 2-12. Calibration of this model focused on injury (plus fatal) crash frequency. The Pearson χ^2 statistic for the model is 122, and the degrees of freedom are 130 (= $n - p$ = 133 – 3). As this statistic is less than $\chi^2_{0.05,\,130}$ (= 158), the hypothesis that the model fits the data cannot be rejected. The R^2 for the model is 0.46. The R_k^2 for the calibrated model is 0.58. This statistic indicates that about 58 percent of the variability due to systematic sources is explained by the model.

Table 2-12. Model Statistical Description–Urban Four-Leg Signalized Intersection.

Model Statistics		Value		
R^2 (R_k^2):		0.46 (0.58)		
Scale Parameter φ:		0.92		
Pearson χ^2:		122 ($\chi^2_{0.05,\,130}$ = 158)		
Over-Dispersion Parameter k:		3.53		
Observations n_o:		133 intersections (1802 injury + fatal crashes in 3 years)		
Standard Deviation s_e:		±2.9 crashes/yr		
Range of Model Variables				
Variable	**Variable Name**	**Units**	**Minimum**	**Maximum**
ADT_{major}	Major-street ADT	veh/d	4040	55,200
ADT_{minor}	Minor-street ADT	veh/d	330	38,000
I_{ch}	Major-street channelization	legs with presence	0	2
I_{ch}	Minor-street channelization	legs with presence	0	2
N	Major-street number of through lanes	lanes	2	6
N	Minor-street number of through lanes	lanes	2	6
Calibrated Coefficient Values				
Variable	**Definition**	**Value**	**Std. Dev.**	**t-statistic**
a	Constant	0.353	0.028	12.6
b_{ch}	Channelization AMF	0.248	0.107	2.3
b_{nl}	Number-of-lanes AMF coefficient	0.197	0.053	3.7

The recommended model is:

$$E[N]_{4LSG} = 0.353\left(\frac{ADT_{major}}{1000}\right)^{0.459}\left(\frac{ADT_{minor}}{1000}\right)^{0.397} \quad \textbf{(45)}$$

where,
$E[N]_{4LSG}$ = expected crash frequency for a four-leg signalized intersection, crashes/yr.

The ADT coefficients in Equation 45 were obtained from Table 2-3. The AMF coefficients are explained in a subsequent section.

Four-Leg Unsignalized Intersection

The results of the four-leg unsignalized intersection model calibration are presented in Table 2-13. Calibration of this model focused on injury (plus fatal) crash frequency. The Pearson χ^2 statistic for the model is 62.9, and the degrees of freedom are 53 (= $n - p$ = 56 – 3). As this statistic is less than $\chi^2_{0.05,\,53}$ (= 71.0), the hypothesis that the model fits the data cannot be rejected. The R^2 for the model is 0.12. The R_k^2 for the calibrated model is 0.27. This statistic indicates that about 27 percent of the variability due to systematic sources is explained by the model.

Table 2-13. Model Statistical Description–Urban Four-Leg Unsignalized Intersection.

Model Statistics		Value		
R^2 (R_k^2):		0.12 (0.27)		
Scale Parameter φ:		1.14		
Pearson χ^2:		62.9 ($\chi^2_{0.05,\,53}$ = 71.0)		
Over-Dispersion Parameter k:		2.07		
Observations n_o:		56 intersections (144 injury + fatal crashes in 3 years)		
Standard Deviation s_e:		±0.78 crashes/yr		
Range of Model Variables				
Variable	**Variable Name**	**Units**	**Minimum**	**Maximum**
ADT_{major}	Major-street ADT	veh/d	400	27,600
ADT_{minor}	Minor-street ADT	veh/d	270	6200
I_{ch}	Major-street channelization	legs with presence	0	2
I_{ch}	Minor-street channelization	legs with presence	0	2
N	Major-street number of through lanes	lanes	2	4
N	Minor-street number of through lanes	lanes	2	4
Calibrated Coefficient Values				
Variable	**Definition**	**Value**	**Std. Dev.**	**t-statistic**
a	Constant	0.172	0.030	5.7
b_{ch}	Channelization AMF	0.937	0.225	4.2
b_{nl}	Number-of-lanes AMF coefficient	-0.135	0.127	-1.1

The recommended model is:

$$E[N]_{4LST} = 0.172\left(\frac{ADT_{major}}{1000}\right)^{0.596}\left(\frac{ADT_{minor}}{1000}\right)^{0.260} \quad \textbf{(46)}$$

where,

$E[N]_{4LST}$ = expected crash frequency for a four-leg unsignalized intersection, crashes/yr.

The ADT coefficients in Equation 46 were obtained from Table 2-3. The AMF coefficients are explained in a subsequent section.

Calibrated AMFs

Two AMFs were calibrated in conjunction with the model recalibration, both are applicable to injury (plus fatal) crashes. The channelization AMF is a leg-specific AMF. It can be used directly to estimate the correlation between the addition of free right-turn channelization and the frequency of crashes on a given intersection leg. Alternatively, it can be used with Equations 13 or 15 to compute an AMF that is applicable to intersection crashes. The leg-specific free right-turn channelization AMF for signalized intersections is:

$$AMF_{leg,ch,sig} = e^{0.248\,I_{ch,j}} \quad \textbf{(47)}$$

where,
$AMF_{leg,ch,sig}$ = leg-specific channelization accident modification factor for signalized intersections.

The AMF in Equation 47 is obtained from Table 2-10 or Table 2-12. It indicates that leg-specific crash frequency is 28 percent larger (= 100 × [$e^{0.248}$ - 1]) when right-turn channelization is present. A similar effect was noted by Bauer and Harwood (*10*) at urban four-leg signalized intersections. It is likely attributable to the yield or stop control associated with right-turn channelization and the fact that the driver stopped in the channelized (free right-turn) lane is typically in a poor position to evaluate safe gaps in the crossing traffic stream. This poor position stems from the large-radius travel path associated with the free right-turn lane. The stop (or yield) line on this path is typically placed near the point of entry to the major road and requires drivers to look back, over their left shoulder, to search for gaps in the conflicting traffic stream. This maneuver is difficult for most drivers and may compromise the diligence with which they search for safe gaps.

The leg-specific free right-turn channelization AMF for unsignalized intersections is:

$$AMF_{leg,ch,unsig} = e^{0.937 I_{ch,j}} \quad \textbf{(48)}$$

where,
$AMF_{leg,ch,unsig}$ = leg-specific channelization accident modification factor for unsignalized intersections.

The AMF in Equation 48 is obtained from Table 2-11 or Table 2-13. It indicates that the leg-specific crash frequency is about 155 percent larger when right-turn channelization is present. The value obtained from Equation 48 is consistent with the effect found by Bauer and Harwood (*10*) for channelization on the stop-controlled approaches at unsignalized intersections. They did not report finding a similar effect when channelization was added to the uncontrolled major-street approaches.

The number-of-lanes AMF coefficients in Tables 2-10 through 2-13 can be used with Equation 39 to compute the street-specific AMF based on the number of through lanes present. This computation was undertaken for the development of the AMF values listed in Table 2-14. These AMF values can be directly used to estimate the correlation between the addition of a through lane and the frequency of crashes on the subject street. Alternatively, they can be used with Equations 6 or 9 to compute an AMF that is applicable to intersection crashes.

Table 2-14. Recommended Street-Specific Number-of-Lanes AMFs.

Number of Through Lanes	Signalized Intersection		Unsignalized Intersection	
	Major Street N_{base} = 4 lanes	Minor Street N_{base} = 2 lanes	Major Street N_{base} = 4 lanes	Minor Street N_{base} = 2 lanes
2	0.67	1.00	1.31	1.00
3	0.82	1.22	1.14	0.87
4	1.00	1.48	1.00	0.76
5	1.22	1.81	0.87	0.67
6	1.48	2.20	0.76	0.58

The trends in Table 2-14 suggest that signalized intersections become less safe as they increase in size. This trend may be due to the increasing exposure to conflict associated with an increasing number of intersecting traffic lanes. In contrast, unsignalized intersections become more safe as they increase in size. This trend suggests minor-movement drivers are more cautious in their gap selection at large unsignalized intersections or, possibly that the more risky maneuvers (e.g., crossing movement) are less frequently attempted. Similar trends are reported in the *Workbook* for rural intersections.

Several AMFs have been described in this chapter. Some major-street-specific AMFs were taken from the *Workbook* and extended to the minor street. Turn lane AMFs were recalculated from data reported in the literature and extended to full range of intersection legs and control conditions. Two new AMFs were developed in association with the model recalibration analysis. The base conditions for this set of AMFs are described in Table 2-15.

Table 2-15. Base Conditions for Urban and Suburban Intersection Models.

Characteristic	Signalized		Unsignalized	
	Major Street	Minor Street	Major Street	Minor Street
Left-turn lanes	present	present	present	--
Right-turn lanes	not present	not present	not present	--
Number of lanes	4	2	4	2
Channelization	not present	not present	not present	not present
Lane width	12 ft	12 ft	12 ft	12 ft
Shoulder width	--	--	1.5 ft	1.5 ft
Median presence	--	--	not present	not present

Note:
"--" - Relationship between the stated characteristic and safety is unknown at this time.

Sensitivity Analysis

This section examines the injury (plus fatal) crash rates that are estimated from the recalibrated models and compares them with those obtained from the existing models. The new crash rates were computed by dividing the expected crash frequency from the base model by the sum of the major-street and minor-street ADTs (with conversion to million-entering-vehicles). The rates computed from the existing models were shown previously in Tables 2-1 and 2-2.

The injury (plus fatal) crash rates computed with the recalibrated models are listed in Tables 2-16 and 2-17 for three-leg and four-leg intersections, respectively. A comparison of the three-leg intersection rates in Tables 2-1 and 2-16 indicate that those for the unsignalized intersections decreased by about 40 percent. In contrast, those for the signalized intersections increased by 10 percent. A comparison of the four-leg intersection rates in Tables 2-2 and 2-17 indicates that those for the unsignalized intersections decreased by about 25 percent. In contrast, those for the signalized intersections increased by 50 percent.

Table 2-16. Crash Rates for Three-Leg Urban and Suburban Intersections.

Control Type	Crash Rate, injury + fatal crashes per million-entering-vehicles				
	Ratio of Minor-Street to Major-Street Volume				
	0.05	**0.10**	**0.15**	**0.20**	**0.25**
Unsignalized [1]	0.11	0.13	0.13	0.14	0.14
Signalized	0.14	0.17	0.19	0.20	0.21

Note:
1 - Unsignalized intersections have an uncontrolled major street and a stop-controlled minor street.

Table 2-17. Crash Rates for Four-Leg Urban and Suburban Intersections.

Control Type	Major-Street Volume, veh/d	Crash Rate, injury + fatal crashes per million-entering-vehicles				
		Ratio of Minor-Street to Major-Street Volume				
		0.10	**0.30**	**0.50**	**0.70**	**0.90**
Unsignalized [1]	5000	0.19	0.21	0.21	0.20	0.19
	10,000	0.17	0.19	0.19	0.18	0.17
	15,000	0.16	0.18	0.18	0.17	0.16
	20,000	0.15	Intersection very likely to meet signal warrants			
	25,000	0.15				
Signalized	5000	0.28	0.37	0.39	0.39	0.39
	10,000	0.25	0.33	0.35	0.35	0.35
	15,000	0.24	0.31	0.33	0.33	0.33
	20,000	0.23	0.30	0.32	0.32	0.32
	25,000	0.22	0.29	0.31	0.31	0.31
	30,000	0.22	0.28	0.30	0.30	0.30
	40,000	0.21	0.27	0.29	0.29	0.29
	≥50,000	0.20	0.26	0.28	0.28	0.28

Note:
1 - Unsignalized intersections have an uncontrolled major street and a stop-controlled minor street.

The rates listed in Tables 2-16 and 2-17 apply to urban intersections with four lanes on the major street and two lanes on the minor street. Crash rates can be produced for other lane combinations using the number-of-lanes AMF in Table 2-14 and Equation 9. In fact, crash rates produced in this manner for intersections with two lanes on the major and minor streets are very similar to those shown in Tables 2-1 and 2-2.

The trends in Table 2-14 indicate that urban intersection crash rate is significantly influenced by the number of lanes present on each street. The trends in Tables 2-16 and 2-17 suggest that crash rate is also influenced by control type. When these two influences are combined, crash rates tend to increase when signalization is added to intersections with many lanes. On the other hand, crash rates tend to decrease (or increase by only a small amount) when signalization is added to intersections with few lanes.

The crash rates for urban signalized intersections in Tables 2-16 and 2-17 are consistent with those for rural signalized intersections. In contrast, the crash rates for unsignalized intersections in Tables 2-16 and 2-17 are lower than those for rural unsignalized intersections. This trend is likely due to (1) lower speed at urban intersections, relative to rural intersections; and (2) a larger number of unsignalized intersections per mile in urban areas than rural areas. With regard to the first point, minor movement drivers at low-speed unsignalized intersections are better able to identify safe gaps in which they can cross or turn. With regard to the second point, major-street drivers more frequently encounter stop-controlled cross streets and are likely to be alert to potential cross street vehicles accessing the major street.

A comparison of the three-leg signalized intersection rates in Table 2-16 with those for three-leg unsignalized intersections suggests that signalization increases crash frequency by 20 to 50 percent, with larger percentages in this range associated with higher minor-street volume. The rates in Table 2-17 indicate that signalization increases crash frequency from 50 to 100 percent at four-leg intersections. Recognizing that the rates in Tables 2-16 and 2-17 correspond to intersections with four lanes on the major street and two lanes on the minor street, these changes suggest that signalizing larger intersections may increase crash frequency.

Although not shown in a table, crash rates were also estimated for urban intersections with two lanes on the major and minor streets. The trend in these rates suggests that signalization reduces crashes from 11 to 27 percent for three-leg intersections, which is consistent with the range found by McGee et al. (*11*) for the conversion of three-leg intersections from unsignalized to signalized control. However, for four-leg intersections, the trend in rates suggests that signalization may still increase crashes by a small amount at smaller intersections in urban areas.

REFERENCES

1. Bonneson, J.A., K. Zimmerman, and K. Fitzpatrick. *Interim Roadway Safety Design Workbook*. FHWA/TX-06/0-4703-P4, Texas Department of Transportation, Austin, Texas, April 2006.

2. Bonneson, J.A., K. Zimmerman, and K. Fitzpatrick. *Roadway Safety Design Synthesis*. FHWA/TX-05/0-4703-P1. Texas Department of Transportation, Austin, Texas, November 2005.

3. Harwood, D., K. Bauer, I. Potts, D. Torbic, K. Richard, E. Kohlman Rabbani, E. Hauer, and L. Elefteriadou. *Safety Effectiveness of Intersection Left- and Right-Turn Lanes*. Report No. FHWA-RD-02-089. Federal Highway Administration, Washington, D.C., July 2002.

4. Behar, G., M. Parkhill, E. Hauer, F. Council, B. Persaud, C. Zegeer, R. Elvik, A. Smiley, and B. Scott. *NCHRP Project 17-27: Prepare Parts I and II of the Highway Safety Manual - Final Report.* Transportation Research Board, National Research Council, Washington, D.C., June 2007.

5. *SAS/STAT User's Guide, Version 6*, 4th ed., SAS Institute, Inc., Cary, North Carolina, 1990.

6. Bonneson, J., D. Lord, K. Zimmerman, K. Fitzpatrick, and M. Pratt. *Development of Tools for Evaluating the Safety Implications of Highway Design Decisions.* FHWA/TX-07/0-4703-4. Texas Department of Transportation, Austin, Texas, September 2006.

7. Lord, D. "Modeling Motor Vehicle Crashes Using Poisson-Gamma Models: Examining the Effects of Low Sample Mean Values and Small Sample Size on the Estimation of the Fixed Dispersion Parameter." *Accident Analysis & Prevention*, Vol. 38, No. 4, Elsevier Ltd., Oxford, Great Britain, July 2006, pp. 751-766.

8. Harwood, D.W., F.M. Council, E. Hauer, W.E. Hughes, and A. Vogt. *Prediction of the Expected Safety Performance of Rural Two-Lane Highways.* FHWA-RD-99-207. Federal Highway Administration, Washington, D.C., 2000.

9. Miaou, S.P. *Measuring the Goodness-of-Fit of Accident Prediction Models.* FHWA-RD-96-040. Federal Highway Administration, Washington, D.C., 1996.

10. Bauer, K., and D. Harwood. *Statistical Models of At-Grade Intersections–Addendum*. Report No. FHWA-RD-99-094. Federal Highway Administration, Washington, D.C., March 2000.

11. McGee, H., S. Taori, and B. Persaud. *NCHRP Report 491: Crash Experience Warrant for Traffic Signals*. Transportation Research Board, National Research Council, Washington, D.C., 2003.

CHAPTER 3. URBAN AND SUBURBAN ARTERIAL STREET SEGMENTS

OVERVIEW

This chapter describes the activities undertaken to calibrate safety prediction models for urban and suburban arterial streets in Texas. Each model provides an estimate of the expected crash frequency for a street segment with a specified length, traffic volume, and geometric design. A series of accident modification factors were also developed as part of the calibration activity. Six models were calibrated and are recommended for use in Texas. Collectively, these models address the following cross section combinations:

- undivided cross section with two through lanes,
- undivided cross section with four through lanes,
- divided cross section with a nonrestrictive median and four through lanes,
- divided cross section with a nonrestrictive median and six through lanes,
- divided cross section with a restrictive median and four through lanes, and
- divided cross section with a restrictive median and six or eight through lanes.

Equivalent models are described in the *Interim Roadway Safety Design Workbook* (*1*). These models were developed using data from locations outside of Texas. They are replaced by the models described in this chapter, all of which have been calibrated using data specific to Texas.

The objective of this research was to calibrate models for each of the aforementioned cross section combinations using crash data for Texas streets. Streets with two-way traffic flow were the focus of the calibration activity. As part of the calibration process, several AMFs were also developed. These AMFs can be used with the calibrated models to estimate the change in safety associated with a specified change in geometry.

This chapter is divided into four parts. The first part provides a review of the literature on the topic of safety prediction models for urban and suburban street segments. The second part describes the development of a generalized curve radius AMF that is applicable to streets and highways. The third part describes the method used to calibrate the proposed models. The last part describes the analysis and findings from the calibration process.

LITERATURE REVIEW

This part of the chapter describes a framework for safety prediction model calibration. It consists of five sections. The first section identifies the various median types found on urban streets and establishes a vocabulary in this regard. The second section provides a brief overview of safety prediction and the role of "base" prediction models. The third section summarizes the base models described in the *Workbook*. The fourth section summarizes the models developed by Harwood et al. for NCHRP Project 17-26 (*2*). The last section reviews the data available in TxDOT's road inventory database.

Median Types in Urban and Suburban Areas

A detailed review of about 700 miles of two-way urban and suburban streets in Texas indicates that about 80 percent of these miles have a divided cross section with one of the following median types:

- two-way left-turn lane (TWLTL) median,
- flush-paved median,
- raised-curb median, or
- depressed median.

The first two median types listed are described herein as "nonrestrictive" median types. They describe the type of median used on about 55 percent of the street mileage reviewed. Of those streets with a nonrestrictive median, about 98 percent have a TWLTL median.

The third and fourth median types listed are described herein as "restrictive" median types. They describe the type of median used on about 25 percent of the street mileage reviewed. Of those streets with a restrictive median, about 90 percent have a raised-curb median.

Undivided cross sections have no median. This cross section was found on about 20 percent of the street-miles reviewed.

Safety Prediction Models

The expected crash frequency for a street segment with specified attributes is computed using a safety prediction model. This model represents the combination of a "base" model and one or more AMFs. The base model is used to estimate the expected crash frequency for a typical segment. The AMFs are used to adjust the base estimate when the attributes of the specific segment are not considered typical. The general form of the typical safety prediction model is shown below as Equation 1, and its base model component is shown as Equation 2.

with,

$$E[N] = E[N]_b \times AMF_{lw} \times AMF_{mw} \ldots \quad \textbf{(1)}$$

$$E[N]_b = a\ ADT^{\,b}\ L \quad \textbf{(2)}$$

where,

$E[N]$ = expected crash frequency, crashes/yr;
$E[N]_b$ = expected base crash frequency, crashes/yr;
AMF_{lw} = lane width accident modification factor;
AMF_{mw} = (restrictive) median width accident modification factor;
a, b = calibration coefficients;
ADT = average daily traffic volume, veh/d; and
L = street segment length, mi.

When the coefficient b for the ADT variable is equal to 1.0, then the coefficient a in Equation 2 is effectively equal to the crash rate, with units of "crashes per million vehicle-miles."

Interim Workbook Base Models

This section describes base models derived by Bonneson et al. (*3*) for urban and suburban arterial street segments. These models were calibrated using data and models described in the research literature. They were subsequently documented in the *Workbook* in the form of a base model and associated base crash rates. The base model in the *Workbook* is:

$$E[N]_b = 0.000365\ Base\ \ ADT\ \ L \quad \textbf{(3)}$$

where,

Base = injury (plus fatal) crash rate (from Table 3-1), crashes/million vehicle-miles.

A comparison of Equation 3 with Equation 2 indicates that the constant *a* is equal to the product of the base crash rate *Base* and "0.000365." The constant represents a conversion factor to convert the ADT into "millions of vehicles per year." Also, the ADT exponent *b* is equal to 1.0 in Equation 3.

The crash rate used in Equation 3 is provided in Table 3-1. They are based on a synthesis of the trends reported by four researchers in their examination of urban street crashes (*3*). The rates for "Undivided or TWLTL Median" are directly applicable to undivided cross sections (i.e., no median). For their application to segments with a TWLTL median, they must be adjusted downward using an AMF provided in the *Workbook*.

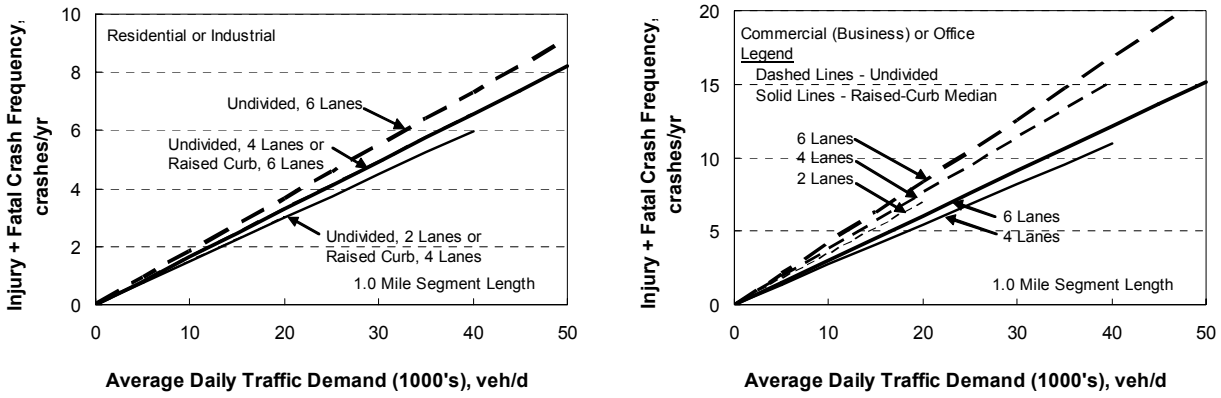
Table 3-1. Base Crash Rates for Urban and Suburban Arterial Street Segments.

Adjacent Land Use	Attributes	Base Crash Rate, injury + fatal crashes/mvm [1]				
	Median Type:	Undivided or TWLTL Median [2]			Raised-Curb Median	
	Through Lanes:	2	4	6	4	6
Commercial (business) or office		0.95	1.04	1.15	0.75	0.83
Residential or industrial		0.41	0.45	0.50	0.41	0.45

Notes:
1 - mvm: million vehicle-miles.
2 - Rates for the TWLTL median must be adjusted using the TWLTL median type AMF.

As indicated in Table 3-1, the crash rates listed are sensitive to the land use adjacent to the subject street. Land use by itself is not a direct cause of crashes. Hence, its inclusion in this table serves as a surrogate for other factors that are related to land use and associated with crash causation or exposure. In terms of exposure, land use affects the level of driveway traffic activity. Commercial business land uses are likely to generate more daily driveway traffic demand than residential land uses. In terms of crash causation, land use is related to the visual intensity of the driving environment. Commercial business land uses tend to have many signs, billboards, and other visual distractions for drivers, relative to other land uses. This land use may also have more vertical obstructions (e.g., trash receptacles, public benches, sign supports, etc.) adjacent to the roadway.

The relationship between crash frequency and traffic demand (represented by Equation 3 and the rates in Table 3-1) is shown in Figure 3-1 for a 1-mile street segment. The trends for residential and industrial land uses are shown in Figure 3-1a, and those for commercial business or office land uses are shown in Figure 3-1b. The trends indicate that undivided cross sections have about 10 percent more crashes than divided cross sections in residential or industrial areas and about 40 percent more crashes in business or office areas. Crash frequency in residential or industrial areas is about one-half that in business or office areas for the same traffic demand and cross section.



a. Residential or Industrial. **b. Business or Office.**

Figure 3-1. Crash Prediction Model Reported in the *Workbook*.

For an ADT of 20,000 veh/d, the trends in Figure 3-1 indicate that a four-lane undivided street in a residential area has an expected crash frequency of 3.3 crashes/yr (7.6 crashes/yr if located in a business area). In contrast, a four-lane divided street in a residential area has an expected crash frequency of 3.0 crashes/yr (5.5 crashes/yr if located in a business area).

The base model in Equation 3 is applicable to street segments that have typical geometric design conditions. These conditions are identified in the *Workbook* and are restated in column two of Table 3-2. The base conditions listed are representative of roads in Texas that are classified as urban or suburban arterial streets. The base model is combined with the AMFs described in the *Workbook* to yield the safety prediction model for street segments.

The characteristics in column 1 of Table 3-2 are each associated with an AMF to account for segment conditions that deviate from the base conditions. Each AMF was evaluated for a range of values to examine its sensitivity to a change in the associated variable. The results of this sensitivity analysis are listed in column 3 of Table 3-2. The ranges listed in these columns equate to AMF values between 0.95 and 1.05. For example, a shoulder width between 0 and 4 ft yields AMF values of about 1.05 and 0.95, respectively. The ranges listed provide some indication of the correlation between a change in the characteristic and the change in crash risk.

Table 3-2. Base Conditions for Urban and Suburban Arterial Street Segments.

Characteristic	Base Condition	Sensitivity of Continuous Variables
Horizontal curve radius	tangent (no curve)	900 ft radius or flatter
Lane width	12 ft	10.5 to 12 ft
Paved shoulder width [1]	1.5 ft (curb-and-gutter)	0 to 4 ft
Restrictive median width	16 ft	8 to 28 ft
Curb parking	none	varies
Utility pole density and offset	50 poles/mi 2.0 ft average offset	Any pole density 1 to 30 ft
Driveway density	Undivided - 40 drives/mi TWLTL - 40 drives/mi Raised curb - 25 drives/mi	Undivided - 35 to 45 drives/mi TWLTL - 35 to 45 drives/mi Raised curb - 20 to 30 drives/mi
Truck presence	6% trucks	4.5 to 7.5%

Note:
1 - Curb-and-gutter is assumed as typical. Width shown is an "effective" shoulder width for curb-and-gutter.

NCHRP Project 17-26 Base Models

This section describes base model derived by Harwood et al. (*2*) for urban and suburban arterial street segments. This model was calibrated using data obtained from state and county transportation agencies in Minnesota and Michigan. The base model they developed is described using the following equations:

with,

$$E[N]_b = E[N]_{mv} + E[N]_{sv} + E[N]_{dw} \quad \textbf{(4)}$$

$$E[N]_{mv} = a_{mv}\, ADT^{\,b_{mv}}\, L \quad \textbf{(5)}$$

$$E[N]_{sv} = a_{sv}\, ADT^{\,b_{sv}}\, L \quad \textbf{(6)}$$

$$E[N]_{dw} = \sum_i a_{dw,i}\, (ADT/15000)^{b_{dw}}\, n_i \quad \textbf{(7)}$$

where,

$E[N]_{mv}$ = expected multiple-vehicle nondriveway crash frequency for base conditions, crashes/yr;
$E[N]_{sv}$ = expected single-vehicle crash frequency for base conditions, crashes/yr;
$E[N]_{dw}$ = expected driveway-related crash frequency for base conditions, crashes/yr;
a_{mv}, b_{mv} = calibration coefficients for multiple-vehicle nondriveway crashes;
a_{sv}, b_{sv} = calibration coefficients for single-vehicle crashes;
$a_{dw,i}$, b_{dw} = calibration coefficients for driveway-related crashes for driveway type i; and
n_i = number of driveways of type i.

The calibration coefficients were separately calibrated for each of the following five combinations of lanes and median type:

- undivided cross section with two through lanes (2U),
- undivided cross section with four through lanes (4U),
- divided cross section with a TWLTL median and two through lanes (2N),
- divided cross section with a TWLTL median and four through lanes (4N), and
- divided cross section with a restrictive median and four through lanes (4R).

It is notable that the calibration coefficients for multiple-vehicle nondriveway and for single-vehicle models (i.e., Equations 5 and 6) are not sensitive to adjacent land use.

Equation 7 estimates the number of driveway crashes as the summation of the crashes that occur at each driveway on the segment, where each driveway is categorized as being one of the following seven types:

- major commercial (retail business),
- minor commercial (retail business),
- major industrial/institutional (factories, warehouses, schools, hospitals, churches, etc.),
- minor industrial/institutional (factories, warehouses, schools, hospitals, churches, etc.),
- major residential,
- minor residential (single-family), and
- other.

In this manner, the driveway crash prediction model incorporates a sensitivity to land use. Major driveways were defined as those that serve sites with 50 or more parking spaces.

Equation 7 was calibrated using data from Minnesota only. The coefficient $a_{dw,i}$ is sensitive to lanes, median type, and driveway type. Hence, it is represented by 35 separate values (= 5 x 7). An examination of these coefficients indicates that, for a given traffic demand, minor residential driveways have the lowest crash frequency per driveway. Major residential driveways have about 5.3 times as many crashes as minor residential driveways. Similarly, minor industrial, minor commercial business, major commercial business, and major industrial driveways have about 1.4, 3.2, 10, and 11 times, respectively, as many crashes as minor residential driveways.

The relationship between crash frequency and traffic demand, as obtained from Equations 4 through 7, is illustrated in Figure 3-2 for a 1-mile street segment. The individual component models are illustrated in Figures 3-2a, 3-2b, and 3-2c. Figure 3-2c illustrates the expected crash frequency when 30 minor residential driveways are present. The sum of the individual component crash frequencies is illustrated in Figure 3-2d. The trends in Figure 3-2d are comparable to those in Figure 3-1a since both are based on residential land use.

Figure 3-1 indicates that, for the same median type, cross sections with more lanes tend to have more crashes for the same traffic demand. In contrast, Figure 3-2 indicates that cross sections with more lanes tend to have *less* crashes. It is likely that the true correlation between lanes and crash frequency over the full range of ADT is more accurately depicted in Figure 3-2. This observation is based on the fact that the models in Figure 3-2 more accurately reflect the nonlinear relationship between traffic demand and crash frequency, while the model depicted in Figure 3-1 is

based on a linear relationship (i.e., rate). The trend with number-of-lanes shown in Figure 3-1 may result from using a linear relationship to extrapolate beyond the range of the data used to compute the rate.

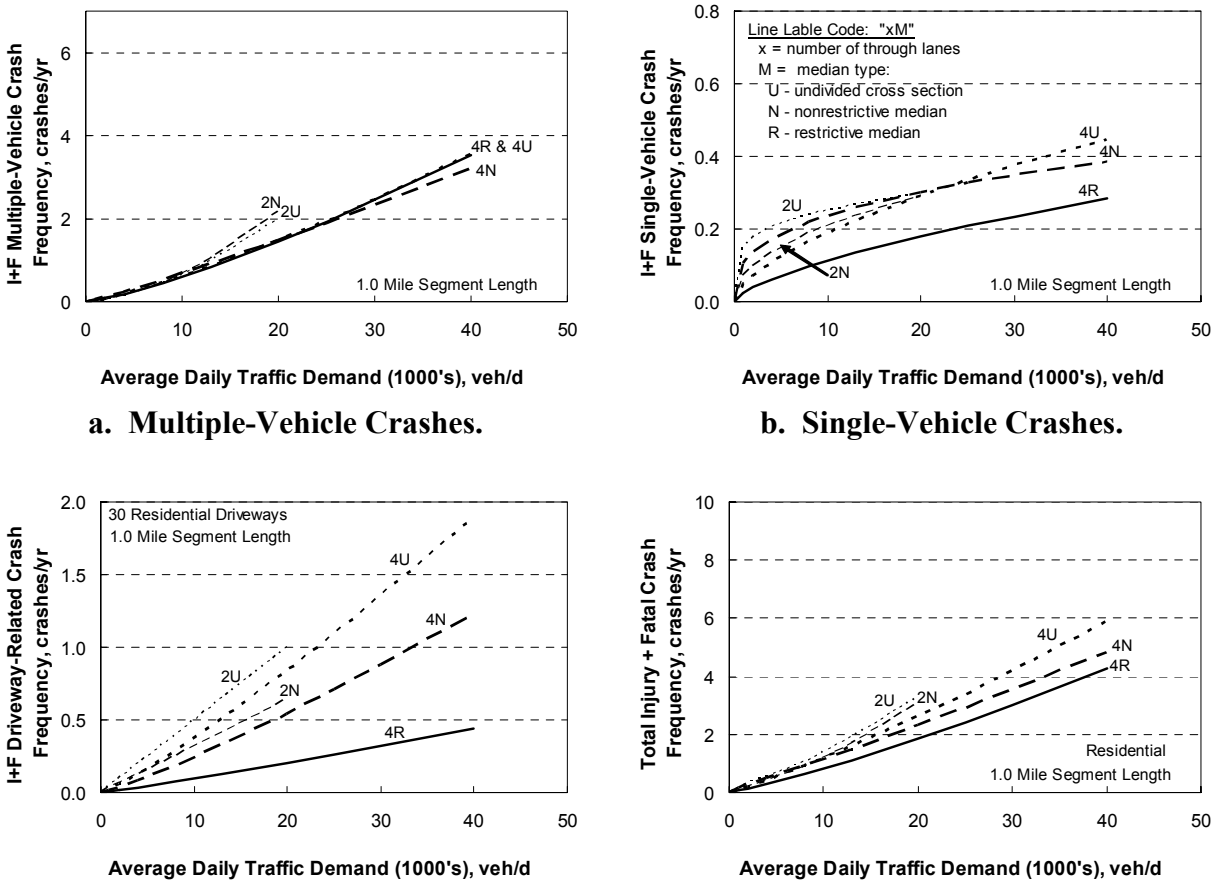


Figure 3-2. Crash Prediction Model Reported by Harwood et al.

For an ADT of 20,000 veh/d, the trends in Figure 3-2d indicate that a four-lane undivided street in a residential area has an expected crash frequency of 2.6 crashes/yr. In contrast, a four-lane divided street in a residential area has an expected crash frequency of 1.8 crashes/yr. When compared with the trends in Figure 3-1a, the expected crash frequency in Figure 3-2d is between 45 and 80 percent of that from Figure 3-1a for any given volume. That is, the expected crash frequency obtained from Equation 4 is notably lower than the crash frequency obtained from the four models used to derive Figure 3-1 (*3*). This trend is consistent when comparing the corresponding models (i.e., Equations 3 and 4) for other land uses. It is evidence of the need to calibrate crash prediction models to "local" conditions.

Texas Highway Data

The geometric and traffic attributes for the Texas state highway system were obtained from the Texas Reference Marker (TRM) system highway database. This database is maintained by TxDOT and contains data for several thousand highway segments, each of which is described in terms of its geometry, traffic, and location attributes. In fact, about 140 attributes are used to describe each highway segment in TRM. However, only about 20 attributes are used to describe segment geometry or traffic characteristics, and about 10 more attributes are used to describe road name and physical location. The remaining attributes describe administrative designations and road-management-related data that are not particularly relevant to safety analysis.

A highway segment in TRM is defined as a length of road along which no one attribute changes. A change in one attribute dictates the end of one segment and the start of a new segment. The average tangent (i.e., not curved) segment length in TRM is about 0.4 mi.

Several segment characteristics are needed in the calibration database. They are identified in Table 3-3. Column 3 of this table indicates whether the corresponding characteristic is available in the TRM database.

Table 3-3. Urban and Suburban Street Segment Calibration Data.

Attribute Category	Attribute	Availability in TRM Database
General	Area type (e.g., urban, rural)	Yes
Traffic control	Two-way vs. one-way operation	Yes
Traffic characteristics	ADT (by specified year)	Yes
	Truck percentage	Yes
Geometry	Horizontal curve radius	Yes
	Spiral transition curve	Yes
	Number of lanes	Yes
	Lane width	Yes for restrictive median types.
	Inside & outside shoulder width	Yes
	Median type	Partially, data does not distinguish between no median, TWLTL median, and flush-paved median.
	Median width	Yes for restrictive median types.
	Curb parking	No
	Utility pole density & offset	No
	Curb presence	Yes
	Driveway location and use	No

The information in Table 3-3 indicates that data for several segment characteristics are not available in the TRM database. However, it was possible to obtain lane width, median type, median width, curb parking, and driveway location data using high-resolution aerial photography available

from the Internet. Utility pole density could not be reliably obtained from any source. It was judged that the base condition value would be sufficiently representative of the collective set of segments in the database. This assumption allowed the corresponding AMF value to be set to 1.0.

For restrictive medians, median width was measured as the distance between the near edges of the traveled way for the two opposing travel directions. Thus, median width includes the width of both inside shoulders.

AMF DEVELOPMENT

This part of the chapter consists of two sections. The first section describes a general model for incorporating cross section type into those AMFs that are found to influence only specific crash types (e.g., run-off-the-road). The second section describes the development of a generalized horizontal curve radius AMF. This AMF incorporates a sensitivity to speed in the AMF estimate.

General AMF Model for Street Treatments

This section describes the development of a general AMF model for incorporating the influence of median type and number-of-lanes into those AMFs that are associated with only specific crash types (e.g., run-off-the-road). The general AMF model was first introduced in the *Synthesis* (*3*). It was subsequently extended to intersection applications in Chapter 2.

The general model is appropriate when a treatment applied to a street segment (e.g., change shoulder width) has its most significant effect on specific crash types. The magnitude of this effect can be quantified using an AMF that is based on just the crashes of the specified types (i.e., AMF_{spc}). Alternatively, it can also be quantified using an AMF based on all crash types (i.e., AMF_{seg}). The relationship between these two AMFs is:

$$AMF_{seg,i} = AMF_{spc,i} P_{spc,i} + 1.0 (1 - P_{spc,i}) \quad \textbf{(8)}$$

where,

$AMF_{seg,i}$ = AMF for a specified treatment to a segment with cross section type i, quantified in terms of all crash types;

$AMF_{spc,i}$ = AMF for a specified treatment to a segment with cross section type i, quantified in terms of one or more specific crash types; and

$P_{spc,i}$ = proportion of specific crash types on a typical segment with cross section i.

Typically, the proportion $P_{spc,i}$ used in Equation 8 varies by cross section type (i.e., number of lanes and median type). Hence, AMF_{seg} also varies by cross section type. Moreover, AMF_{seg} is a more useful form for the AMF than AMF_{spc} when used in the type of model represented by Equation 1 (i.e., a model that predicts crashes of all types combined).

In some instances, AMF_{seg} is desired for a different cross section type than that for which an AMF is available. The remainder of this subsection describes how Equation 8 can be used to convert an $AMF_{seg,i}$ for cross section type i into an $AMF_{seg,j}$ for cross section type j.

Equation 8 can be re-written in the following form to describe the AMF for cross section type *j*:

$$AMF_{seg,j} = AMF_{spc,j} P_{spc,j} + 1.0 (1 - P_{spc,j}) \quad (9)$$

Equations 8 and 9 can be combined to obtain an equation for estimating AMF_{seg} for cross section *j* by extrapolating the AMF_{seg} that was derived for cross section *i*. The following equation was derived from their combination:

$$AMF_{seg,j} = \left(AMF_{seg,i} - 1\right) \frac{P_{spc,j}}{P_{spc,i}} + 1.0 \quad (10)$$

Inspection of Equation 10 indicates that it adheres to logical boundary conditions (e.g., $AMF_{seg,j} = 1.0$ when $P_{spc,j} = 0$; and $AMF_{seg,j} = AMF_{seg,i}$ when $P_{spc,j} = P_{spc,i}$).

The proportion of specific crash types used in Equation 10 is provided in the *Workbook* (*1*) for several AMFs. These proportions were based on Texas crash data for the years 1999 and 2000. Since their derivation, crash data were made available for year 2001. These data were used to update the proportions to reflect a three-year average (i.e., 1999, 2000, and 2001). The results of this analysis are shown in Table 3-4. The proportions shown in each row are based on the total listed in the last column.

Table 3-4. Urban and Suburban Arterial Street Crash Distribution.

Median Type [1]	Through Lanes	Proportion of Injury + Fatal Crashes					Total Crash Frequency [2], crashes per 3 years
		Single-Vehicle		Multiple-Vehicle		Single-Vehicle with Pole	
		Either Side	Right Side	Opp. Dir.	Sideswipe		
Undivided or Non-restrictive	2	0.1885	not applicable	0.0603	0.0243	0.0318	9,639
	4	0.0935		0.0280	0.0436	0.0294	19,607
	6	0.0496		0.0244	0.0600	0.0150	5,198
Restrictive	4	0.1822	0.1072	not applicable	0.0584	0.0348	6,409
	6	0.1476	0.0798		0.0955	0.0327	6,322
	8	0.0406	0.0237		0.0881	0.0135	1,181

Notes:
1 - Nonrestrictive: TWLTL or flush-paved median. Restrictive: raised-curb or depressed median.
2 - Injury (plus fatal) crashes reported during the years 1999, 2000, and 2001.

The proportions listed in Table 3-4 were combined as needed for three AMFs in the *Workbook*. The combinations reflect the types of crashes that were found to be influenced by a change in the corresponding geometric element, as a function of median type and number of lanes. The resulting proportions to be used in Equation 10 are listed in Table 3-5. The values listed in this table are very similar to those reported in the *Workbook*.

Table 3-5. Urban and Suburban Arterial Street Crash Distribution for Selected AMFs.

Median Type	Through Lanes	Proportion of Injury + Fatal Crashes		
		Lane Width AMF	Shoulder Width AMF	Utility Pole Offset AMF
Undivided or Non-restrictive	2	0.27 [a]	0.19 [b]	0.032 [d]
	4	0.17 [a]	0.094 [b]	0.029 [d]
	6	0.13 [a]	0.050 [b]	0.015 [d]
Restrictive	4	0.26 [a]	0.11 [c]	0.035 [d]
	6	0.26 [a]	0.080 [c]	0.033 [d]
	8	0.14 [a]	0.024 [c]	0.014 [d]

Notes:
a - Single-vehicle run-off-road, same-direction sideswipe, and multiple-vehicle opposite direction crashes.
b - Single-vehicle run-off-road crashes, either side.
c - Single-vehicle run-off-road crashes, right side.
d - Single-vehicle-with-pole crashes.

Horizontal Curve Radius AMF

The relationship between curve radius and crash frequency has been the subject of several research projects (*3, 4, 5, 6*). The trends reported by three researchers are shown in Figure 3-3.

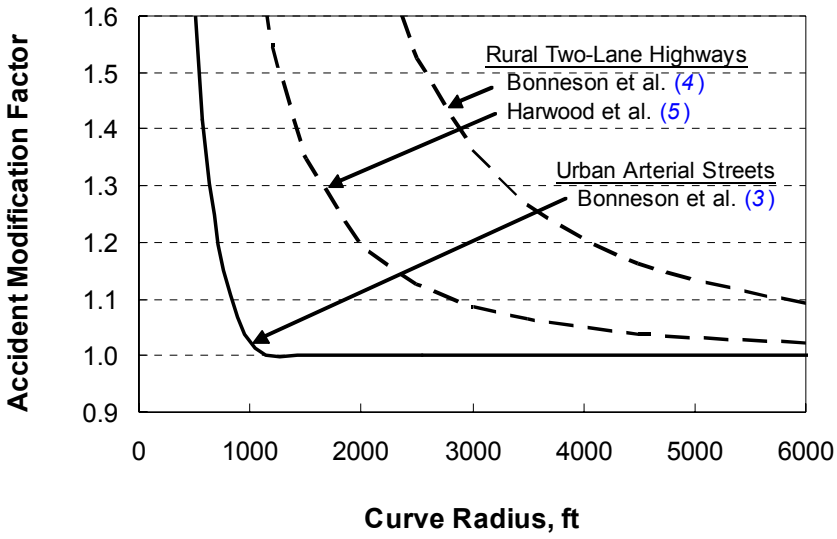


Figure 3-3. Relationship between Curve Radius and AMF for Various Road Types.

The trend in Figure 3-3 that is identified as "urban arterial streets" was derived from a model originally developed by Hauer et al. (*6*). It represents the trend found in injury crashes. The average speed limit of the segments in their database was 40 mph. The two-lane highway model developed by Bonneson et al. (*4*) is based on injury (plus fatal) crashes. The average speed limit of the segments in their database was about 70 mph. The two-lane highway model developed by Harwood

et al. (5) is based on all crash severities. The average speed was not reported but the data used corresponded to the mid-1980s when the national speed limit was 55 mph. The trend line shown for Harwood et al. is based on a typical deflection angle of 20 degrees. The relationship between the three models implies a possible influence of speed on crash frequency.

The curve radius AMF is fundamentally defined as follows:

$$AMF_{cr} = \frac{E[N]_{tangent} + E[N]_{curve}}{E[N]_{tangent}} \quad \textbf{(11)}$$

where,

AMF_{cr} = horizontal curve radius accident modification factor;
$E[N]_{tangent}$ = expected crash frequency on tangent segment, crashes/yr; and
$E[N]_{curve}$ = expected crash frequency on segment due to curvature, not including those crashes that would have occurred had the segment not been curved; crashes/yr.

The following equations are rationalized to estimate the components of Equation 11:

$$E[N]_{tangent} = a\ ADT^{b}\ L_c\ AMF_i \quad \textbf{(12)}$$

and

$$E[N]_{curve} = a\ ADT^{b}\ I_c\ f_d\ AMF_i \quad \textbf{(13)}$$

$$f_d = \frac{(1.47\ V)^2}{32.2\ R} - \frac{e}{100} \quad \textbf{(14)}$$

where,

AMF_i = accident modification factor for element i (i = lane width, shoulder width, etc.);
L_c = length of horizontal curve (= $I_c \times R$ / 5280 / 57.3), mi;
I_c = curve deflection angle, degrees;
f_d = side friction demand factor;
V = speed limit, mph;
e = superelevation rate, percent; and
R = curve radius, ft.

Equation 13 does not include curve length L_c as a measure of exposure. Rather, curve deflection angle I_c is rationalized as the appropriate measure of exposure for crashes that occurred because the segment is curved (i.e., $E[N]_{curve}$). Just as segment length L represents the product of forward velocity and travel time, deflection angle represents the product of angular velocity and travel time.

Side friction demand is included in Equation 13 as being logically correlated with crash frequency. As side friction demand reaches the roll threshold for tall vehicles, the propensity for a roll-over crash increases. Also, as side friction demand reaches the limit of available friction, the propensity for sliding off the road increases. The relationship between crash rate and side friction demand derived by Bonneson et al. (7) is shown in Figure 3-4.

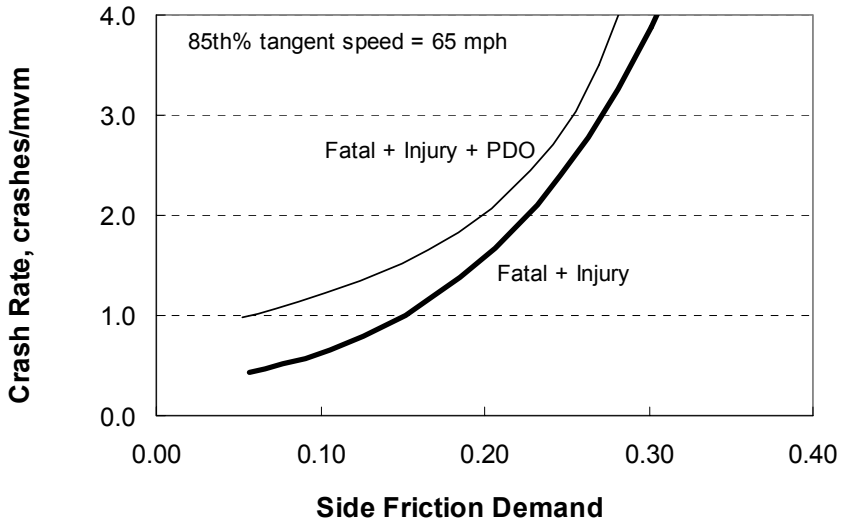


Figure 3-4. Relationship between Crash Rate and Side Friction Demand.

Equations 11 through 14 can be combined to yield the following curve radius AMF form:

$$AMF_{cr} = 1.0 + b_0 \frac{(1.47\ V)^2}{32.2\ R^2} \quad \textbf{(15)}$$

where,
b_0 = calibration coefficient.

The contribution of the superelevation term (i.e., $e/[100\ R]$) in Equation 14 is small and is removed for mathematical convenience.

As shown by Bonneson et al. (*3*), the curve radius AMF developed by Harwood et al. (*5*) reduces to a form that is similar to that found in Equation 15. The Harwood AMF does not include the speed variable V^2 but it does include a deflection angle variable I_c in the divisor of the second term of Equation 15. By including deflection angle in the divisor, the AMF developed by Harwood et al. indicates that curves are safer (i.e., have a lower AMF value) when they have a *larger* deflection angle. A possible explanation for this trend is that curve-related crashes are primarily associated with the curve entrance (or exit), rather than its length or deflection angle (*8*). However, another explanation for this trend is that deflection angle is negatively correlated with speed. This explanation is supported by the relationship between speed limit and deflection angle shown in Figure 3-5. The data shown are based on 66,000 rural highway curves in Texas.

The data points in Figure 3-5 indicate a similar relationship for major and minor collector highways, which are predominantly two-lane highways. The trend line shown represents a line of best fit to the collector highway data. The trend is less clear for arterial highways. Similar trends were found for urban street curves.

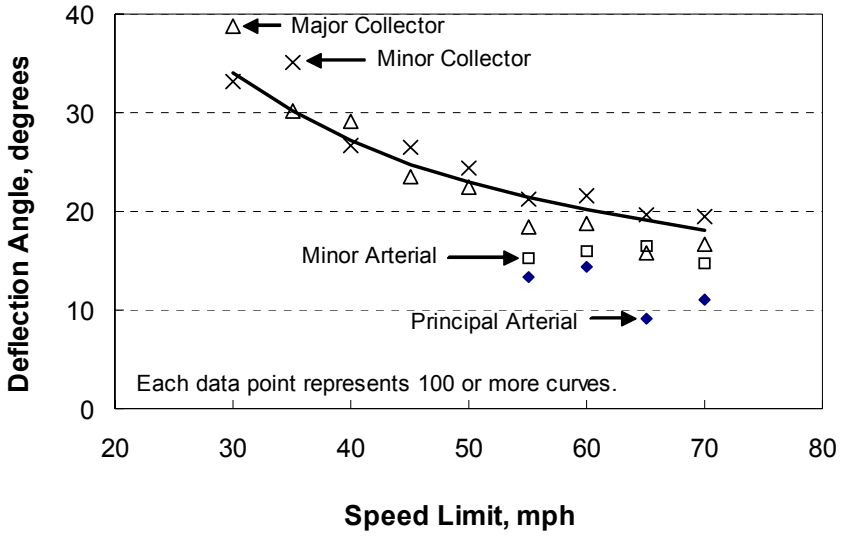


Figure 3-5. Relationship between Speed Limit and Deflection Angle.

The trend in Figure 3-5 may explain why deflection angle was included in the Harwood et al. AMF. That is, deflection angle was included because it statistically "explained" some of the variation in the crash data. Figure 3-5 suggests that what was found (i.e., deflection angle is correlated with crashes) was a surrogate for speed (perhaps more appropriately design speed) and that speed is more likely a causal factor in curve crashes than deflection angle. Based on this analysis, it is rationalized that the deflection angle term in the Harwood et al. model is actually explaining the effect of speed, which provides additional justification for the use of speed in the numerator of the second term of Equation 15.

Equation 15 was fit to the "urban arterial streets" trend line in Figure 3-3. The calibration coefficient b_0 yielding the best fit is 1154. A similar exercise for the Harwood et al. model in Figure 3-3 yielded a best-fit coefficient of 3873. The best-fit coefficient for the Bonneson et al. model is 10,030. The variation of this coefficient among the three models suggests that there is still some unexplained influence that is not represented in the AMF. Further examination of the coefficients indicated an additional sensitivity to speed that is not explained by Equation 15. This relationship is shown in Figure 3-6. The best-fit trend line in this figure suggests an additional sensitivity to speed (i.e., $V^{3.8608}$) should be included in Equation 15.

Based on the preceding analysis, the following generalized curve radius AMF was developed:

$$AMF_{cr} = 1 + 0.97\ (0.147\ V)^4\ \frac{(1.47\ V)^2}{32.2\ R^2} \qquad \textbf{(16)}$$

This AMF is shown in Figure 3-7 for selected speed limits. The AMF models shown previously in Figure 3-3 are shown in Figure 3-7 using dashed lines. Equation 16 is shown to explain the

differences between the three models reported in the literature. Other, unknown factors may also explain these differences. Additional research is needed to confirm the effect of speed on curve safety. This research should also examine the correlation between superelevation rate and crash frequency.

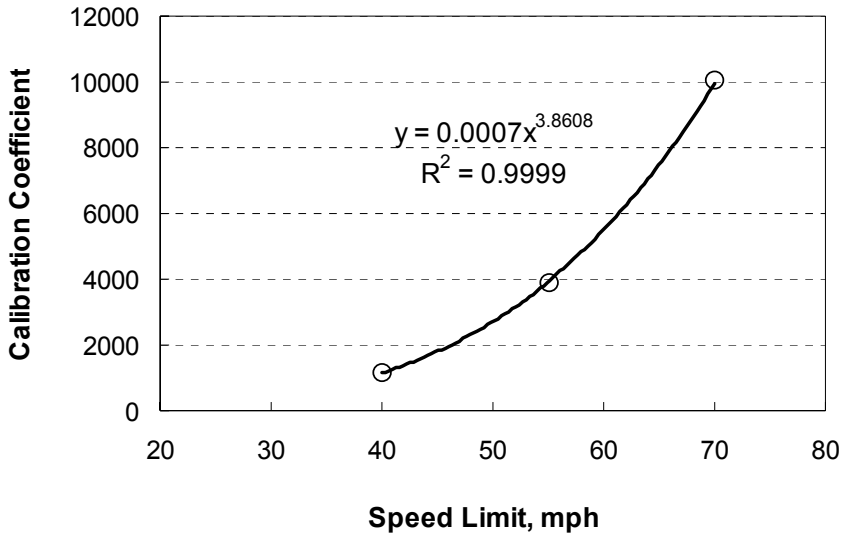


Figure 3-6. Relationship between Speed Limit and Model Calibration Coefficient.

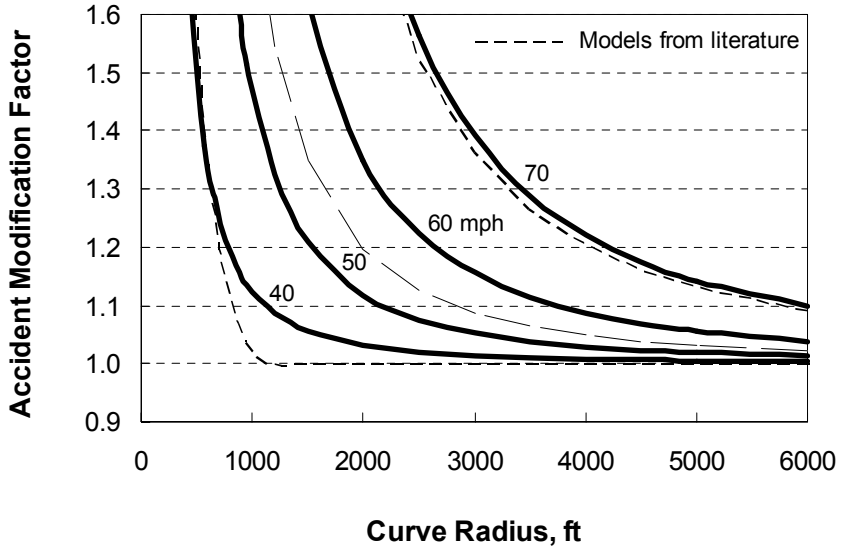


Figure 3-7. Relationship between Curve Radius, Speed, and AMF.

METHODOLOGY

This part of the chapter describes the methodology used to calibrate the urban and suburban arterial street segment safety prediction model. It is divided into three sections. The first section describes the database used to calibrate the model. The second section describes the technique used to quantify the calibration coefficients. The last section describes the technique used to adjust the over-dispersion parameter for bias due to a small sample size.

Database

A database was assembled from the TRM database for segment calibration. Variables associated with the AMFs described in the *Workbook* (*1*) were included in the database to ensure that their relationship with crash occurrence was included in the analysis. However, as previously noted in the discussion associated with Table 3-3, utility pole density and offset are not in the TRM database, and measurement of these characteristics in the field was not a feasible option. A review of the aerial photos for each segment indicated (to the extent permitted by photo resolution) that the base condition values listed in Table 3-2 for utility poles were representative of the collective set of segments in the database.

Adjacent Land Use

The database was assembled to allow for the exploration of correlation between adjacent land use and crash frequency. Land use was characterized as residential or undeveloped, industrial, commercial business, and office. Table 3-6 identifies the characteristics of each land use. These characteristics were used to determine the land use adjacent to, and on each side of, the subject segment. The characterization was focused on the land use within a 100-ft wide zone adjacent to the segment, as measured from the outside edge of the near shoulder and away from the segment.

For each street segment, land use was measured as a length along the segment. It was separately measured on each side of the street. Driveways and public street approaches to the segment were included in the measured length and assigned a land use equivalent to that of the property to which they were adjacent. With this method of measurement, the total length of land use on a segment equaled twice the segment's length.

Driveway Count

Driveways along a segment were counted and categorized according to the land use they served. Public street approaches and driveways served by a signalized intersection were not included in the count of driveways. The following two types of driveways were identified:

- Full driveway - a driveway that allowed the following four turn movements for access: right-in, right-out, left-in, and left-out.
- Partial driveway - a driveway that allowed right-in and right-out movements but prevented left-turns with channelizing islands, a restrictive median, or one-way driveway operation.

Table 3-6. Adjacent Land Use Characteristics.

Land Use	Characteristics	Examples
Residential or Undeveloped	• Buildings are small • A small percentage of the land is paved • If driveways exist, they have very low volume • Ratio of land-use acreage to parking stalls is large	• Single-family home • Undeveloped property • Graveyard • Farmland • Park or green-space recreation area
Industrial	• Buildings are large and production oriented • Driveways and parking may be designed to accommodate large trucks • Driveway volume is moderate at shift change times and is low throughout the day • Ratio of land-use acreage to parking stalls is moderate	• Factory • Warehouse • Storage tanks • Farmyard with barns and machinery
Commercial Business	• Buildings are larger and separated by convenient parking between building and roadway • Driveway volume is moderate from mid-morning to early evening • Ratio of land-use acreage to parking stalls is small	• Strip commercial • Shopping mall • Apartment complex • Trailer park • Airport • Gas station • Restaurant
Office	• Buildings typically have two or more stories • Most parking is distant from the building or behind it • Driveway volume is high at morning and evening peak traffic hours; otherwise, it is low • Ratio of land-use acreage to parking stalls is small	• Office tower • Public building • Church • School • Clubhouse (buildings at a park) • Parking lot for "8 to 5" workers

In the count of driveways, a full driveway was counted as "1" driveway, and a partial driveway was counted as "0.5" driveways. This distinction was rationalized as appropriate given the difference in the number of conflicting paths present with each driveway type and the correlation between the number of these paths and driveway crash frequency. Partial driveways were most frequently found on segments with a restrictive median.

Driveways that were apparently unused were not counted. Similarly, driveways leading into fields, small utility installations (e.g., cellular phone tower), and abandoned buildings were not counted. A circular driveway at a residence was counted as one driveway even though both ends of the driveway intersected the subject segment. A similar approach was used when a small business (e.g., gas station) had two curb cuts separated by only 10 or 20 ft.

An alley that intersects the subject street was considered to be a driveway. In residential areas, each alley was counted as one driveway unless it provided the primary driveway access to the homes it served. In this latter case, the alley was considered to represent an "equivalent" number of residential driveways. The count of equivalent driveways equaled the number of residences with driveways served by the alley and for which the associated drivers were likely to access the subject street (as opposed to using the alley to access a different public street).

Curb Parking Presence and Type

Two types of curb parking were identified: angle and parallel. Curb parking presence and type was measured in terms of its length along the segment. The length of each parking type was measured along both sides of the street. The measured length of parking represents only that of the angle parking stalls, or the legal length of street for which parallel parking is permitted. With parallel parking, it was assumed that parking was legally required to be 30 ft back from any public street approach. With this method of measurement, if parking is provided along both sides of the segment, the total length of parking would add up to twice the segment's length minus all of the following: (1) the width of all driveways and associated curb returns, (2) the width of all public street approaches and associated curb returns, and (3) 60 ft (= 2 × 30 ft) for each public street approach.

Curve Radius and Length

The segments included in the database were selected from the TRM database such that they did not have curves with a radius of less than 5000 ft. This approach was undertaken to allow the calibration activity to focus on the calibration of models for tangent street segments. However, when reviewing the aerial photograph of each segment, a small percentage of segments were found to have moderate curvature. In this situation, the curve's length and radius were recorded and added to the database.

The curve radius AMF described in Equation 16 was used account for the presence of curvature. The following equation was used to compute the aggregated AMF for segments with combined curve and tangent portions:

$$AMF_{cr|agg} = \frac{(L - L_c)\,1.0 \; - \; L_c\,AMF_{cr}}{L} \qquad \textbf{(17)}$$

where,

$AMF_{cr|agg}$ = aggregated curve radius AMF for a segment with both tangent and curved portions.

In those few cases where a curve was found on the segment, it was often located entirely within the segment. However, sometimes only a portion of the curve would be located on the segment. In this situation, the length of curve recorded in the database represented only that portion of the curve that was within the segment boundaries.

Calibration Technique

The calibration activity used statistical analysis software that employs maximum likelihood methods based on a negative binomial distribution of crash frequency. The calibration model consisted of Equation 1 and subtle variations of Equations 4 through 7. The models represented by Equations 5, 6, and 7 are referred to as "component models." Equations 1 and 4 through 7 represent one "aggregate model" and were simultaneously analyzed to find the best-fit calibration coefficients. In this manner, the component models were calibrated to compute the expected frequency of the corresponding crash type for a common set of AMFs (as shown in Equation 1). The analysis approach is described in more detail in the next part of this chapter.

Over-Dispersion Parameter Adjustment

It was assumed that crash frequency on a segment is Poisson distributed, and that the distribution of the mean crash frequency for a group of similar intersections is gamma distributed. In this manner, the distribution of crashes for a group of similar intersections can be described by the negative binomial distribution. The variance of this distribution is:

$$V[X] = y\,E[N] + \frac{(y\,E[N])^2}{k\,L} \quad \textbf{(18)}$$

where,

$V[X]$ = crash frequency variance for a group of similar locations, crashes^2;
X = reported crash count for y years, crashes;
y = time interval during which X crashes were reported, yr; and
k = over-dispersion parameter, mi^{-1}.

Research by Lord (*9*) has indicated that databases with low sample mean values and small sample size may not exhibit as much variability as suggested by Equation 18. Rather, the regression model may "over-explain" some of the random variability in the small database, or the low sample mean may introduce an instability in the model coefficients (*9*).

Also in recognition of the concerns expressed by Lord (*9*), the over-dispersion parameter k obtained from the regression analysis is adjusted downward, such that it yields a more reliable estimate of the variance of the crash distribution. The derivation of the adjustment technique used is described in Reference *4*. The following equation was used to estimate the true over-dispersion parameter, given the estimated over-dispersion parameter obtained from the regression analysis.

$$k_t = \frac{\sqrt{(n-p)^2\,m^2 + 69\,k_r L.\,(n-p)\,m} - (n-p)\,m}{34.5\;L.} \quad \textbf{(19)}$$

where,

k_t = true over-dispersion parameter;
k_r = estimated over-dispersion parameter obtained from database analysis;
n = number of observations (i.e., segments in database);
p = number of model variables;
$L.$ = average segment length for all n observations (= 1.0 for intersection applications); and
m = average number of crashes per observation (= total crashes in database / n).

Equation 19 was used to estimate the true over-dispersion parameter for each of the models described in this chapter. All subsequent references to the over-dispersion parameter k in this chapter denote the estimated true parameter obtained from Equation 19 (i.e., hereafter, $k = k_t$).

CALIBRATION DATA AND ANALYSIS

This part of the chapter describes the calibration of the aggregate arterial street segment safety prediction model. Initially, the process used to select the segments is identified. Then, the characteristics of the segments are summarized. Next, the techniques used to assemble the crash data

are described. Finally, the findings from the model calibration are discussed. In this section, a "site" is used to refer to one street segment.

Site Selection and Data Collection

The street segments used to calibrate the safety prediction model were required to satisfy the following criteria:

- Area type: urban;
- Functional class: arterial;
- Segment length: 0.1 mi or more; and
- Curve presence: no curves with a radius less than 5000 ft.

All segments were required to have a minimum length of 0.1 mi. The value of 0.1 mi reflects a recognition of the precision of crash location in TxDOT's state-wide crash database.

The TRM database was screened to exclude segments that had one or more curves with a radius less than 5000 ft. This criterion was intended to allow the calibration to focus on tangent segments. However, a visual inspection of aerial photography for each segment revealed the presence of a curve on some segments. Those segments with flatter curves were kept in the database, and the effect of curvature was modeled using Equation 17. Segments with sharp curvature were removed from the database.

Research by Harwood et al. (*2*) and the synthesis by Bonneson et al. (*3*) indicate that segment crash frequency is significantly influenced by cross section (i.e., median type and number of lanes). This influence is so significant and varied that a separate model should be calibrated for each unique combination of median type and lanes that exist in the database. A review of the TRM database indicated that the following eight cross section combinations are most common in Texas:

- undivided cross section with two through lanes (2U),
- undivided cross section with four through lanes (4U),
- divided cross section with a nonrestrictive median and two through lanes (2N),
- divided cross section with a nonrestrictive median and four through lanes (4N),
- divided cross section with a nonrestrictive median and six through lanes (6N),
- divided cross section with a restrictive median and four through lanes (4R),
- divided cross section with a restrictive median and six through lanes (6R), and
- divided cross section with a restrictive median and eight through lanes (8R).

Thus, in addition to the aforementioned site selection criteria, effort was expended during the site selection process to target enough segments for each combination such that separate models could be accurately calibrated. To achieve this goal, a minimum sample size criterion based on total exposure was established for each combination.

Site Characteristics

Traffic and geometry data were identified for several thousand highway segments using the TRM database for year 2003, the first year for which a complete database was available. All total, 942 segments satisfied the site selection criteria. These segments represent streets located in 24 of the 25 TxDOT districts. Selected characteristics are provided for these segments in Table 3-7.

Table 3-7. Summary Characteristics for Urban and Suburban Arterial Street Segments.

Median Type	Through Lanes	Total Segments	Segment Length, mi	Seg. Length Range, mi		Volume Range, veh/d	
				Minimum	Maximum	Minimum	Maximum
Undivided (U)	2	72	32.2	0.10	2.1	850	15,600
	4	140	53.6	0.10	1.6	2,500	31,300
Non-restrictive (N)	2	12	4.4	0.11	1.0	5,700	13,200
	4	436	195.9	0.10	2.3	1,950	48,100
	6	89	53.6	0.10	2.5	7,400	54,000
Restrictive (R)	4	56	25.0	0.10	2.0	4,450	32,400
	6	128	84.9	0.10	3.7	3,450	56,700
	8	9	8.4	0.14	3.0	33,000	67,300
Overall:		942	458.0	0.10	3.7	850	67,300

Examination of the data in Table 3-7 indicates that there are relatively few segments for two of the cross section combinations. Specifically, the 2N and the 8R cross sections were few in number and may not have sufficient sample size to calibrate robust models. The 2N cross section is rationalized to have traffic flows that interact in a unique manner relative to other cross sections and thus, should not be combined. In contrast, the 8R cross section was rationalized to be operationally similar to the 6R cross section and could be combined to calibrate one model applicable to both cross section types.

The adjacent land use characteristics associated with the segments are described in Table 3-8. In general, the commercial business category is the most common land use found on the arterial street segments. The undeveloped or single-family residential land uses are the second most commonly found land use. Industrial and office land uses collectively comprise about 10 percent of all land use.

The number of driveways associated with each cross section combination is also shown in Table 3-8. Driveways serving business land uses dominate all other categories. Driveway counts to the nearest 0.5 driveway are shown for the restrictive median types because this median type was frequently associated with "partial" driveways (as defined in a previous section).

Table 3-8. Land Use Characteristics for Urban and Suburban Arterial Street Segments.

Median Type	Through Lanes	Proportion of Adjacent Land Use				Number of Driveways			
		U-R [1]	Industrial	Business	Office	U-R [1]	Industrial	Business	Office
Undivided	2	0.702	0.068	0.202	0.028	481	67	249	31
	4	0.405	0.048	0.485	0.062	635	74	1212	55
Non-restrictive	2	0.437	0.004	0.434	0.125	13	1	60	2
	4	0.397	0.052	0.500	0.051	948	301	4425	227
	6	0.323	0.072	0.558	0.047	135	117	1432	55
Restrictive	4	0.459	0.026	0.471	0.044	35.5	27.5	297.0	31.5
	6	0.380	0.030	0.496	0.094	132.0	51.5	1084.5	112.5
	8	0.229	0.025	0.655	0.092	2.0	0.5	148.0	9.5
Overall:		0.417	0.041	0.475	0.068	2382	640	8908	524

Note:
1 - U-R.: undeveloped or single-family residential property.

Data Collection

Crash data were identified for each segment using the TxDOT crash database. Three years of crash data, corresponding to years 1999, 2000, and 2001, were identified for each segment. The ADT for each of these three years was obtained from the TRM database and averaged to obtain one ADT for each segment. Crash data prior to 1999 were not used due to the increase in speed limit that occurred on many Texas highways in 1997 and 1998. Crashes that were associated with intersections were excluded from the database.

The crash records associated with each segment were categorized in terms of whether they were multiple-vehicle nondriveway crashes, single-vehicle crashes, or driveway-related crashes. Only crashes that were identified as associated with one or more injuries or fatalities (i.e., K, A, B, C severity rating) were included in the database.

Data Analysis

This section is divided into two subsections. The first subsection summarizes the crash data at the selected study sites. The second subsection describes the formulation of the calibration model and summarizes the statistical analysis methods used to calibrate it.

Database Summary

The crash data for the 942 segments are summarized in Table 3-9. These segments were collectively associated with 7860 injury or fatal crashes. The trends in crash rate are provided in the last column of the table. Table 3-1 indicates that the injury (plus fatal) crash rate for urban and suburban arterial street segments ranges from 0.41 to 1.15 crashes/mvm, depending on land use and median type. This range is consistent with the injury (plus fatal) crash rates in Table 3-9.

Table 3-9. Crash Data Summary for Arterial Street Segments.

Median Type	Through Lanes	Exposure,[1] mvm	Injury + Fatal Crashes / 3 years				Crash Rate, cr/mvm
			Multiple-Vehicle[2]	Single-Vehicle	Driveway-Related	Total	
Undivided	2	71.9	35	29	50	114	0.53
	4	228.5	315	94	345	754	1.10
Non-restrictive	2	14.9	4	7	2	13	0.29
	4	1192.8	1276	255	1221	2752	0.77
	6	594.4	867	141	816	1824	1.02
Restrictive	4	144.3	122	35	112	269	0.62
	6	818.6	705	191	539	1435	0.58
	8	165.8	421	31	247	699	1.41
Overall:		3231.2	3745	783	3332	7860	0.81

Notes:
1 - mvm: million vehicle-miles.
2 - Multiple-vehicle crashes do not include driveway-related crashes.

Model Development and Statistical Analysis Methods

This section describes the proposed arterial street segment safety prediction model and the methods used to calibrate it. The form of this model is:

$$E[N] = E[N]_b \times AMF_{lw} \times AMF_{mw} \times AMF_{pk} \times AMF_{cr|agg} \times AMF_{sw} \times AMF_{nmw} \times AMF_{tk} \quad \textbf{(20)}$$

with,

$$E[N]_b = E[N]_{mv} I_{mv} + E[N]_{sv} I_{sv} + E[N]_{dw} I_{dw} \quad \textbf{(21)}$$

$$E[N]_{mv} = E[N]_{mv2u} I_{2u} + E[N]_{mv2n} I_{2n} + E[N]_{mv4u} I_{4u} + E[N]_{mv4n} I_{4n} + E[N]_{mv4r} I_{4r} + E[N]_{mv6n} I_{6n} + E[N]_{mv6r} I_{6r} \quad \textbf{(22)}$$

$$E[N]_{mv2u} = e^{b_{mv2u} + b_{mvu} Ln(ADT/1000) + Ln(L) + b_{vi} P_{ind} + b_{vb} P_{bus} + b_{vo} P_{off}} \quad \textbf{(23)}$$

$$E[N]_{mv2n} = e^{b_{mv2n} + (b_{mvu} + b_{mvn}) Ln(ADT/1000) + Ln(L) + b_{vi} P_{ind} + b_{vb} P_{bus} + b_{vo} P_{off}} \quad \textbf{(24)}$$

$$E[N]_{mv4u} = e^{b_{mv4u} + b_{mvu} Ln(ADT/1000) + Ln(L) + b_{vi} P_{ind} + b_{vb} P_{bus} + b_{vo} P_{off}} \quad \textbf{(25)}$$

$$E[N]_{mv4n} = e^{b_{mv4n} + (b_{mvu} + b_{mvn}) Ln(ADT/1000) + Ln(L) + b_{vi} P_{ind} + b_{vb} P_{bus} + b_{vo} P_{off}} \quad \textbf{(26)}$$

$$E[N]_{mv4r} = e^{b_{mv4r} + (b_{mvu} + b_{mvr}) Ln(ADT/1000) + Ln(L) + b_{vi} P_{ind} + b_{vb} P_{bus} + b_{vo} P_{off}} \quad \textbf{(27)}$$

$$E[N]_{mv6n} = e^{b_{mv6n} + (b_{mvu} + b_{mvn}) Ln(ADT/1000) + Ln(L) + b_{vi} P_{ind} + b_{vb} P_{bus} + b_{vo} P_{off}} \quad \textbf{(28)}$$

$$E[N]_{mv6r} = e^{b_{mv6r} + (b_{mvu} + b_{mvr}) Ln(ADT/1000) + Ln(L) + b_{vi} P_{ind} + b_{vb} P_{bus} + b_{vo} P_{off}} \quad \textbf{(29)}$$

$$E[N]_{sv} = E[N]_{sv2u} I_{2u} + E[N]_{sv2n} I_{2n} + E[N]_{sv4u} I_{4u} + E[N]_{sv4n} I_{4n} + E[N]_{sv4r} I_{4r} + E[N]_{sv6n} I_{6n} + E[N]_{sv6r} I_{6r} \quad (30)$$

$$E[N]_{sv2u} = e^{b_{sv2u} + b_{svu} Ln(ADT/1000) + Ln(L) + b_{vi} P_{ind} + b_{vb} P_{bus} + b_{vo} P_{off}} \quad (31)$$

$$E[N]_{sv2n} = e^{b_{sv2n} + (b_{svu} + b_{svn}) Ln(ADT/1000) + Ln(L) + b_{vi} P_{ind} + b_{vb} P_{bus} + b_{vo} P_{off}} \quad (32)$$

$$E[N]_{sv4u} = e^{b_{sv4u} + b_{svu} Ln(ADT/1000) + Ln(L) + b_{vi} P_{ind} + b_{vb} P_{bus} + b_{vo} P_{off}} \quad (33)$$

$$E[N]_{sv4n} = e^{b_{sv4n} + (b_{svu} + b_{svn}) Ln(ADT/1000) + Ln(L) + b_{vi} P_{ind} + b_{vb} P_{bus} + b_{vo} P_{off}} \quad (34)$$

$$E[N]_{sv4r} = e^{b_{sv4r} + (b_{svu} + b_{svr}) Ln(ADT/1000) + Ln(L) + b_{vi} P_{ind} + b_{vb} P_{bus} + b_{vo} P_{off}} \quad (35)$$

$$E[N]_{sv6n} = e^{b_{sv6n} + (b_{svu} + b_{svn}) Ln(ADT/1000) + Ln(L) + b_{vi} P_{ind} + b_{vb} P_{bus} + b_{vo} P_{off}} \quad (36)$$

$$E[N]_{sv6r} = e^{b_{sv6r} + (b_{svu} + b_{svr}) Ln(ADT/1000) + Ln(L) + b_{vi} P_{ind} + b_{vb} P_{bus} + b_{vo} P_{off}} \quad (37)$$

$$E[N]_{dw} = E[N]_{dw2u} I_{2u} + E[N]_{dw2n} I_{2n} + E[N]_{dw4u} I_{4u} + E[N]_{dw4n} I_{4n} + E[N]_{dw4r} I_{4r} + E[N]_{dw6n} I_{6n} + E[N]_{dw6r} I_{6r} \quad (38)$$

$$E[N]_{dw2u} = (n_{res} + n_{ind} e^{b_{dwi}} + n_{bus} e^{b_{dwb}} + n_{off} e^{b_{dwo}}) e^{b_{dw2u} + b_{dwu} Ln(ADT/15000) + b_D Ln(S_d)} \quad (39)$$

$$E[N]_{dw2n} = (n_{res} + n_{ind} e^{b_{dwi}} + n_{bus} e^{b_{dwb}} + n_{off} e^{b_{dwo}}) e^{b_{dw2n} + (b_{dwu} + b_{dwn}) Ln(ADT/15000) + b_D Ln(S_d)} \quad (40)$$

$$E[N]_{dw4u} = (n_{res} + n_{ind} e^{b_{dwi}} + n_{bus} e^{b_{dwb}} + n_{off} e^{b_{dwo}}) e^{b_{dw4u} + b_{dwu} Ln(ADT/15000) + b_D Ln(S_d)} \quad (41)$$

$$E[N]_{dw4n} = (n_{res} + n_{ind} e^{b_{dwi}} + n_{bus} e^{b_{dwb}} + n_{off} e^{b_{dwo}}) e^{b_{dw4n} + (b_{dwu} + b_{dwn}) Ln(ADT/15000) + b_D Ln(S_d)} \quad (42)$$

$$E[N]_{dw4r} = (n_{res} + n_{ind} e^{b_{dwi}} + n_{bus} e^{b_{dwb}} + n_{off} e^{b_{dwo}}) e^{b_{dw4r} + (b_{dwu} + b_{dwr}) Ln(ADT/15000) + b_D Ln(S_d)} \quad (43)$$

$$E[N]_{dw6n} = (n_{res} + n_{ind} e^{b_{dwi}} + n_{bus} e^{b_{dwb}} + n_{off} e^{b_{dwo}}) e^{b_{dw6n} + (b_{dwu} + b_{dwn}) Ln(ADT/15000) + b_D Ln(S_d)} \quad (44)$$

$$E[N]_{dw6r} = (n_{res} + n_{ind} e^{b_{dwi}} + n_{bus} e^{b_{dwb}} + n_{off} e^{b_{dwo}}) e^{b_{dw6r} + (b_{dwu} + b_{dwr}) Ln(ADT/15000) + b_D Ln(S_d)} \quad (45)$$

$$AMF_{sw,i} = (e^{b_{sw}(W_s - 1.5)} - 1.0) \frac{P_i}{0.11} + 1.0 \quad (46)$$

$$AMF_{nmw} = e^{(I_{2n} + I_{4n} + I_{6n}) b_{nmw} (W_{nm} - 12)} \quad (47)$$

$$AMF_{tk} = e^{b_{tk}(P_t - 6.0)} \quad (48)$$

where,

AMF_{lw} = lane width accident modification factor;

AMF_{mw} = (restrictive) median width accident modification factor;

AMF_{pk} = curb parking accident modification factor;
$AMF_{cr|agg}$ = aggregated curve radius accident modification factor (Equation 17);
AMF_{sw} = (outside) shoulder width accident modification factor;
AMF_{nmw} = nonrestrictive median width accident modification factor;
AMF_{tk} = truck presence accident modification factor;
I_{mv} = crash indicator variable (= 1.0 if multiple-vehicle nondriveway crash data, 0.0 otherwise);
I_{sv} = crash indicator variable (= 1.0 if single-vehicle crash data, 0.0 otherwise);
I_{dw} = crash indicator variable (= 1.0 if driveway-related crash data, 0.0 otherwise);
$E[N]_{mvXY}$ = expected multiple-vehicle nondriveway crash frequency for number-of-lanes X and median type Y, (X = 2, 4, 6; Y = u: undivided, n: nonrestrictive, r: restrictive); crashes/yr;
I_{XY} = cross section indicator variable (= 1.0 if cross section has X lanes and median type Y, 0.0 otherwise);
$E[N]_{svXY}$ = expected single-vehicle crash frequency for number-of-lanes X and median type Y; crashes/yr;
$E[N]_{dwXY}$ = expected driveway-related crash frequency for number-of-lanes X and median type Y; crashes/yr;
P_{ind} = proportion of adjacent land use that is industrial (= 0.5 L_{ind} /L);
L_{ind} = curb miles associated with industrial land use, mi;
P_{bus} = proportion of adjacent land use that is commercial business (= 0.5 L_{bus} /L);
L_{bus} = curb miles associated with commercial business land use, mi;
P_{off} = proportion of adjacent land use that is office (= 0.5 L_{off} /L);
L_{off} = curb miles associated with office land use, mi;
n_{res} = number of driveways serving residential land uses;
n_{ind} = number of driveways serving industrial land uses;
n_{bus} = number of driveways serving business land uses;
n_{off} = number of driveways serving office land uses;
S_d = average driveway spacing (= 2 L / [n_{res} + n_{ind} + n_{bus} + n_{off} + 1]), mi/driveway;
W_s = shoulder width, ft;
P_i = proportion of influential crashes of type i (see Table 3-5);
W_{nm} = nonrestrictive median width, ft;
P_t = percent trucks represented in ADT, %; and
b_i = calibration coefficient for condition i (see Table 3-10).

The mathematical relationships for each of the first three AMFs identified in the variable list are provided in the *Workbook*. A preliminary examination of the data indicated that these AMFs were accurate estimators of the relationship between lane width, restrictive median width, or parking presence and multiple-vehicle, single-vehicle, or driveway-related crashes. For this calibration activity, the proportions in Table 3-5 were used with the lane width AMF from the *Workbook* (as opposed to the proportions provided in the *Workbook*).

An examination of the data indicated that the number of driveway crashes on a segment did not increase linearly with the count of driveways. Segments with many driveways tended to have more crashes than those with few driveways but the increase in crashes was smaller than the increase

in driveways. The effect was rationalized to represent a decrease in driveway volume as the spacing between driveways decreased. A property's driveway volume is positively correlated its length of frontage and negatively correlated with its number of driveway access points. Thus, average driveway spacing was included in the model as a surrogate for driveway volume.

The mathematical relationships for the shoulder width AMF and the truck presence AMF identified in the *Workbook* were not used in the proposed safety prediction model. A preliminary examination of the data indicated that these AMFs did not accurately replicate the trends in the crash data. For this reason, it was determined that these AMFs should be calibrated using the assembled database. An examination of their effect on multiple-vehicle, single-vehicle, and driveway-related crashes did not reveal a compelling reason to develop separate AMFs for each crash type.

An examination of the data indicated that the median width AMF in the *Workbook* did not accurately replicate the trends in the crash data for segments with a nonrestrictive median. For this reason, it was determined that a nonrestrictive median width AMF should be calibrated using the database. An examination of median effect on multiple-vehicle, single-vehicle, and driveway-related crashes did not reveal a compelling reason to develop separate AMFs for each crash type.

The utility pole density AMF identified in the *Workbook* is not included in the proposed model because pole count and location data were not readily available. It is assumed that the collective set of segments in the database is represented by the base condition of a 2.0 ft pole offset and 50 poles/mi. As such, the model should not be biased as a result of excluding this AMF.

The driveway density AMF and the TWLTL median type AMF identified in the *Workbook* are not included in the proposed model because their correlation with safety is represented by alternative model components. Specifically, driveway-related crashes and TWLTL median type are explicitly represented by the component models of the proposed safety prediction model.

The correlation between the presence of a curb on the outside of the roadway and crash frequency was examined. However, negligible correlation between curb presence and crash frequency was found. It was noted that 80 percent of the segments in the database had a curb present. The average shoulder width when a curb is present was 1.4 ft and when a curb is not present was 2.2 ft. The similarity of the shoulder widths for the "with" and "without" curb present categories suggested that the curb effect was not likely being inadvertently explained by the observed effect of shoulder width.

The Nonlinear Regression procedure (NLIN) in the SAS software was used to estimate the proposed model coefficients (*10*). This procedure was used because the proposed safety prediction model is both nonlinear and discontinuous. The loss function associated with NLIN was specified to equal the log likelihood function for the negative binomial distribution. Equation 18 was used to define the variance function for the multiple-vehicle and single-vehicle crashes. Equation 18, with the segment length term L removed, was used to define the variance function for the driveway-related crashes. The procedure was set up to estimate model coefficients based on maximum-likelihood methods.

One disadvantage of the NLIN procedure is that it is not able to compute the best-fit value of k for the calibrated prediction model. This disadvantage is overcome by using the Generalized Modeling (GENMOD) procedure in SAS. GENMOD automates the over-dispersion factor estimation process using maximum-likelihood methods. Initially, NLIN is used to calibrate the prediction model. Then, GENMOD is used to regress the relationship between the reported and predicted crash frequencies (where the natural log of the predicted values is specified as an offset variable, and the log link function is used) to obtain a best-fit over-dispersion estimate a using the internal variance function ($V[X] = E[N] + a\,E[N]^2$). The estimate a from GENMOD is then multiplied by the average segment length and inverted to estimate a new value of k (i.e., $k_{new} = 1/[a\,L.]$). This new value is then used in a second application of NLIN and the process is repeated until convergence is achieved between the k value used in NLIN and that obtained from GENMOD. Convergence is typically achieved in two iterations. As a last step, the value of k obtained from this iterative method was used in Equation 19 to estimate the true over-dispersion parameter.

Model Calibration

The safety prediction model calibration process consisted of the simultaneous calibration of all three component models (i.e., multiple-vehicle nondriveway, single-vehicle, and driveway-related models) and AMFs using the aggregate model represented by Equation 20. The simultaneous calibration approach was needed because the AMFs in Equation 20 were common to all three component models. The database assembled for this process included three replications of the original database. However, the dependent variable in the first replication was set equal to the multiple-vehicle crashes. The dependent variable in the second replication was set equal to the single-vehicle crashes and that for the third replication was set equal to the driveway-related crashes.

The process of replicating the original database resulted in the variability in the data being underestimated because the random error in each set of three observations per site is correlated. The NLIN procedure does not directly accommodate the analysis of repeated measures but the GENMOD procedure does support this analysis. Thus, to estimate the true variability in each regression coefficient estimate from NLIN, the GENMOD procedure was used to regress the relationship between the reported and predicted crash frequencies. GENMOD was instructed to use generalized estimating equations to estimate the standard deviation of the intercept term accounting for the repeated measures. The ratio of this standard deviation to that from a similar model that assumed independent errors was found to be 1.23. The standard deviation for each regression coefficient obtained from NLIN was multiplied by this ratio to estimate its true standard deviation.

For each segment, the predicted crash frequency from each component model was totaled and compared with the total reported crash frequency. The difference between the two totals was then summed for all segments. This sum was found to be very small (i.e., less than 0.5 percent of the total reported crash frequency), so it was concluded that there was no bias in the component models in terms of their ability to predict total crash frequency.

The calibration coefficients for the aggregate safety prediction model are described in the next subsection. The subsequent three subsections collectively describe each of the three component

models and their fit to the data. The fit statistics were separately computed using the calibrated component model and an analysis of its residuals.

Aggregate Model

The results of the aggregated safety prediction model calibration are presented in Table 3-10. Calibration of this model focused on injury (plus fatal) crash frequency. The Pearson χ^2 statistic for the model is 2751, and the degrees of freedom are 2786 (= $n - p$ = [3 × 942] − 40). As this statistic is less than $\chi^2_{0.05, 2786}$ (= 2910), the hypothesis that the model fits the data cannot be rejected.

The t-statistics listed in the last column of Table 3-10 indicate a test of the hypothesis that the coefficient value is equal to 0.0. Those t-statistics with an absolute value that is larger than 2.0 indicate that the hypothesis can be rejected with the probability of error in this conclusion being less than 0.05. For those few variables where the absolute value of the t-statistic is smaller than 2.0, it was decided that the variable was essential to the model and was rationalized as being reasonable in magnitude (even if the specific value was not known with a great deal of certainty as applied to this database).

Adjustments to the 2N Component Models

An examination of the predicted crash frequency obtained for the two-lane-street-with-a-nonrestrictive-median-type (2N) model indicates inconsistent results. Specifically, the predicted 2N crash frequency is consistently lower than that of all the other models, for a common traffic demand. A detailed examination indicates that its prediction of multiple-vehicle and driveway-related crashes are particularly low, relative to the other models of similar crash type. The likely reason for this inconsistency is that the few segments that represent this cross section in the database (as noted in the discussion associated with Table 3-7) are not representative of typical 2N segments.

The model developed by Harwood et al. (*2*), as shown in Figure 3-2, indicates that the 2N cross section should exhibit only about 10 percent fewer crashes than the 2U cross section. The analysis of before-after data for the conversion of 2U to 2N cross sections described in Reference *3* indicates a 13 percent reduction in crashes when the ADT ranges from 8000 to 16,000 veh/d. Based on this analysis, the multiple-vehicle and the driveway-related component models associated with the 2N cross section cannot be recommended for practical application if the calibration coefficients in Table 3-10 are used. Recognizing that a model for this cross section is needed, the b_{mv2n} and b_{dw2n} coefficients in Table 3-10 were adjusted proportionally such that the total predicted crash frequency is 85 to 90 percent of that obtained from the 2U component models for ADTs in the range of 8000 to 16,000 veh/d. The best-fit values of the two adjusted coefficients are:

- b_{mv2n} = -4.46
- b_{dw2n} = -2.27

Table 3-10. Urban and Suburban Arterial Model Statistical Description–Aggregate Model.

Model Statistics		Value		
Scale Parameter φ:		0.99		
Pearson χ^2:		2751 ($\chi^2_{0.05,\,2786}$ = 2910)		
Observations n_o:		942 segments (7860 injury + fatal crashes in 3 years)		
Calibrated Coefficient Values				
Variable	**Inferred Effect of...[1]**	**Value**	**Std. Dev.**	**t-statistic**
b_{sw}	Shoulder width	-0.0317	0.0129	-2.5
b_{nmw}	Nonrestrictive median width	-0.0255	0.0220	-1.2
b_{tk}	Truck presence	-0.0068	0.0062	-1.1
b_D	Driveway spacing	0.518	0.087	5.9
b_{mv2u}	2 lane, undivided on m.v. crashes in U-R	-5.620	0.651	-8.6
b_{mv2n}	2 lane, nonrestrictive median on m.v. crashes in U-R	-5.831	0.743	-7.9
b_{mv4u}	4 lane, undivided on m.v. crashes in U-R	-5.973	0.751	-8.0
b_{mv4n}	4 lane, nonrestrictive median on m.v. crashes in U-R	-5.044	0.377	-13.4
b_{mv4r}	4 lane, restrictive median on m.v. crashes in U-R	-3.744	0.460	-8.1
b_{mv6n}	6 lane, nonrestrictive median on m.v. crashes in U-R	-5.246	0.447	-11.7
b_{mv6r}	6 lane, restrictive median on m.v. crashes in U-R	-3.925	0.519	-7.6
b_{mvu}	ADT for undivided on m.v. crashes in U-R	2.309	0.280	8.3
b_{mvn}	Added ADT effect for nonrest. median on m.v. crashes	-0.486	0.303	-1.6
b_{mvr}	Added ADT effect for rest. median on m.v. crashes	-0.925	0.316	-2.9
b_{sv2u}	2 lane, undivided on s.v. crashes in U-R	-3.222	0.613	-5.3
b_{sv2n}	2 lane, nonrestrictive median on s.v. crashes in U-R	-2.659	0.648	-4.1
b_{sv4u}	4 lane, undivided on s.v. crashes in U-R	-3.746	0.743	-5.0
b_{sv4n}	4 lane, nonrestrictive median on s.v. crashes in U-R	-3.077	0.467	-6.6
b_{sv4r}	4 lane, restrictive median on s.v. crashes in U-R	-1.643	0.560	-2.9
b_{sv6n}	6 lane, nonrestrictive median on s.v. crashes in U-R	-2.798	0.564	-5.0
b_{sv6r}	6 lane, restrictive median on s.v. crashes in U-R	-1.413	0.635	-2.2
b_{svu}	ADT for undivided on s.v. crashes in U-R	1.058	0.281	3.8
b_{svn}	Added ADT effect for nonrest. median on s.v. crashes	-0.428	0.320	-1.3
b_{svr}	Added ADT effect for rest. median on s.v. crashes	-0.857	0.337	-2.5
b_{vi}	Added effect for industrial land use	0.419	0.285	1.5
b_{vb}	Added effect for business land use	0.895	0.118	7.6
b_{vo}	Added effect for office land use	0.225	0.272	0.8
b_{dw2u}	2 lane, undivided on dw. crashes in U-R	-2.120	0.476	-4.5
b_{dw2n}	2 lane, nonrestrictive median on dw. crashes in U-R	-4.458	1.119	-4.0
b_{dw4u}	4 lane, undivided on dw. crashes in U-R	-2.287	0.427	-5.4
b_{dw4n}	4 lane, nonrestrictive median on dw. crashes in U-R	-2.604	0.412	-6.3
b_{dw4r}	4 lane, restrictive median on dw. crashes in U-R	-2.411	0.421	-5.7
b_{dw6n}	6 lane, nonrestrictive median on dw. crashes in U-R	-2.612	0.428	-6.1
b_{dw6r}	6 lane, restrictive median on dw. crashes in U-R	-2.722	0.408	-6.7
b_{dwu}	ADT for undivided on dw. crashes in U-R	1.038	0.222	4.7
b_{dwn}	Added ADT effect for nonrest. median on dw. crashes	0.247	0.255	1.0
b_{dwr}	Added ADT effect for rest. median on dw. crashes	0.217	0.300	0.7
b_{dwi}	Added effect of industrial driveway on dw. crashes	0.274	0.580	0.5
b_{dwb}	Added effect of business driveway on dw. crashes	1.413	0.315	4.5
b_{dwo}	Added effect of office driveway on dw. crashes	1.068	0.481	2.2

Note: 1 - m.v.: multiple-veh. nondriveway; s.v.: single-veh.; dw.: driveway-related; U-R: undeveloped or single-family residential.

Multiple-Vehicle Nondriveway Crashes

The results of the multiple-vehicle safety prediction model calibration are presented in Table 3-11. Calibration of this model focused on injury (plus fatal) crash frequency. The Pearson χ^2 statistic for the model is 885, and the degrees of freedom are 926 (= $n - p$ = 942 – 16). As this statistic is less than $\chi^2_{0.05, 926}$ (= 998), the hypothesis that the model fits the data cannot be rejected; however, it is noted that there is slightly less variability in the data than suggested by Equation 18. The R^2 for the model is 0.79. An alternative measure of model fit that is better suited to the negative binomial distribution is R_k^2, as developed by Miaou (*11*). The R_k^2 for the calibrated model is 0.93. This statistic indicates that about 93 percent of the variability due to systematic sources is explained by the model.

Table 3-11. Prediction Model Statistical Description–Multiple-Vehicle Model.

Model Statistics	Value
R^2 (R_k^2):	0.79 (0.93)
Scale Parameter φ:	0.94
Pearson χ^2:	885 ($\chi^2_{0.05, 926}$ = 998)
Over-Dispersion Parameter k:	4.98 mi^{-1}
Observations n_o:	942 segments (3745 injury + fatal crashes in 3 years)
Standard Deviation s_e:	±1.54 crashes/yr

The coefficients in Table 3-10 were combined with Equations 23 to 29 to obtain the calibrated safety prediction models for multiple-vehicle nondriveway crashes. The form of each model is:

$$E[N]_{mv2u} = e^{-5.620 + 2.309\,Ln(ADT/1000) + Ln(L) + 0.419\,P_{ind} + 0.895\,P_{bus} + 0.225\,P_{off}} \quad \textbf{(49)}$$

$$E[N]_{mv2n} = e^{-4.46 + 1.823\,Ln(ADT/1000) + Ln(L) + 0.419\,P_{ind} + 0.895\,P_{bus} + 0.225\,P_{off}} \quad \textbf{(50)}$$

$$E[N]_{mv4u} = e^{-5.973 + 2.309\,Ln(ADT/1000) + Ln(L) + 0.419\,P_{ind} + 0.895\,P_{bus} + 0.225\,P_{off}} \quad \textbf{(51)}$$

$$E[N]_{mv4n} = e^{-5.044 + 1.823\,Ln(ADT/1000) + Ln(L) + 0.419\,P_{ind} + 0.895\,P_{bus} + 0.225\,P_{off}} \quad \textbf{(52)}$$

$$E[N]_{mv4r} = e^{-3.744 + 1.384\,Ln(ADT/1000) + Ln(L) + 0.419\,P_{ind} + 0.895\,P_{bus} + 0.225\,P_{off}} \quad \textbf{(53)}$$

$$E[N]_{mv6n} = e^{-5.246 + 1.823\,Ln(ADT/1000) + Ln(L) + 0.419\,P_{ind} + 0.895\,P_{bus} + 0.225\,P_{off}} \quad \textbf{(54)}$$

$$E[N]_{mv6r} = e^{-3.925 + 1.384\,Ln(ADT/1000) + Ln(L) + 0.419\,P_{ind} + 0.895\,P_{bus} + 0.225\,P_{off}} \quad \textbf{(55)}$$

The fit of the calibrated models is shown in Figure 3-8. This figure compares the predicted and reported crash frequency in the calibration database. The trend line shown represents a "$y = x$" line. A data point would lie on this line if its predicted and reported crash frequency were equal. The data points shown represent the observed multiple-vehicle nondriveway crash frequency for the segments used to calibrate the corresponding component model.

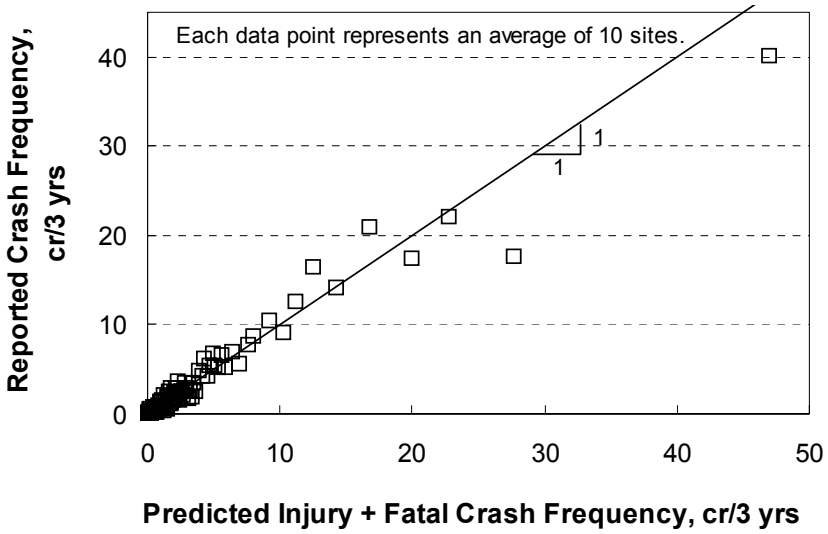


Figure 3-8. Predicted vs. Reported Multiple-Vehicle Crashes.

Each data point shown in Figure 3-8 represents the average predicted and average reported crash frequency for a group of 10 segments. The data were sorted by predicted crash frequency to form groups of segments with similar crash frequency. The purpose of this grouping was to reduce the number of data points shown in the figure and, thereby, to facilitate an examination of trends in the data. The individual segment observations were used for model calibration. In general, the data shown in the figure indicate that the model provides an unbiased estimate of expected crash frequency for segments experiencing up to 25 multiple-vehicle crashes in a three-year period.

Single-Vehicle Crashes

The results of the single-vehicle safety prediction model calibration are presented in Table 3-12. Calibration of this model also focused on injury (plus fatal) crash frequency. The Pearson χ^2 statistic for the model is 955, and the degrees of freedom are 926 (= $n - p$ = 942 – 16). As this statistic is less than $\chi^2_{0.05,\,926}$ (= 998), the hypothesis that the model fits the data cannot be rejected. The R^2 for the model is 0.46, and the R_k^2 for the calibrated model is 0.93. This statistic indicates that about 93 percent of the variability due to systematic sources is explained by the model.

Table 3-12. Prediction Model Statistical Description–Single-Vehicle Model.

Model Statistics	Value
R^2 (R_k^2):	0.46 (0.93)
Scale Parameter φ:	1.01
Pearson χ^2:	955 ($\chi^2_{0.05,\,926}$ = 998)
Over-Dispersion Parameter k:	6.12 mi^{-1}
Observations n_o:	942 segments (783 injury + fatal crashes in 3 years)
Standard Deviation s_e:	±0.36 crashes/yr

The coefficients in Table 3-10 were combined with Equations 31 to 37 to obtain the calibrated safety prediction models for single-vehicle crashes. The form of each model is:

$$E[N]_{sv2u} = e^{-3.222 + 1.058\,Ln(ADT/1000) + Ln(L) + 0.419\,P_{ind} + 0.895\,P_{bus} + 0.225\,P_{off}} \quad \textbf{(56)}$$

$$E[N]_{sv2n} = e^{-2.659 + 0.630\,Ln(ADT/1000) + Ln(L) + 0.419\,P_{ind} + 0.895\,P_{bus} + 0.225\,P_{off}} \quad \textbf{(57)}$$

$$E[N]_{sv4u} = e^{-3.746 + 1.058\,Ln(ADT/1000) + Ln(L) + 0.419\,P_{ind} + 0.895\,P_{bus} + 0.225\,P_{off}} \quad \textbf{(58)}$$

$$E[N]_{sv4n} = e^{-3.077 + 0.630\,Ln(ADT/1000) + Ln(L) + 0.419\,P_{ind} + 0.895\,P_{bus} + 0.225\,P_{off}} \quad \textbf{(59)}$$

$$E[N]_{sv4r} = e^{-1.643 + 0.201\,Ln(ADT/1000) + Ln(L) + 0.419\,P_{ind} + 0.895\,P_{bus} + 0.225\,P_{off}} \quad \textbf{(60)}$$

$$E[N]_{sv6n} = e^{-2.798 + 0.630\,Ln(ADT/1000) + Ln(L) + 0.419\,P_{ind} + 0.895\,P_{bus} + 0.225\,P_{off}} \quad \textbf{(61)}$$

$$E[N]_{sv6r} = e^{-1.413 + 0.201\,Ln(ADT/1000) + Ln(L) + 0.419\,P_{ind} + 0.895\,P_{bus} + 0.225\,P_{off}} \quad \textbf{(62)}$$

The fit of the calibrated models is shown in Figure 3-9. This figure compares the predicted and reported crash frequency in the calibration database. Each data point shown represents the average predicted and average reported crash frequency for a group of 10 segments. In general, the data shown in the figure indicate that the model provides an unbiased estimate of expected crash frequency for segments experiencing up to four single-vehicle crashes in a three-year period.

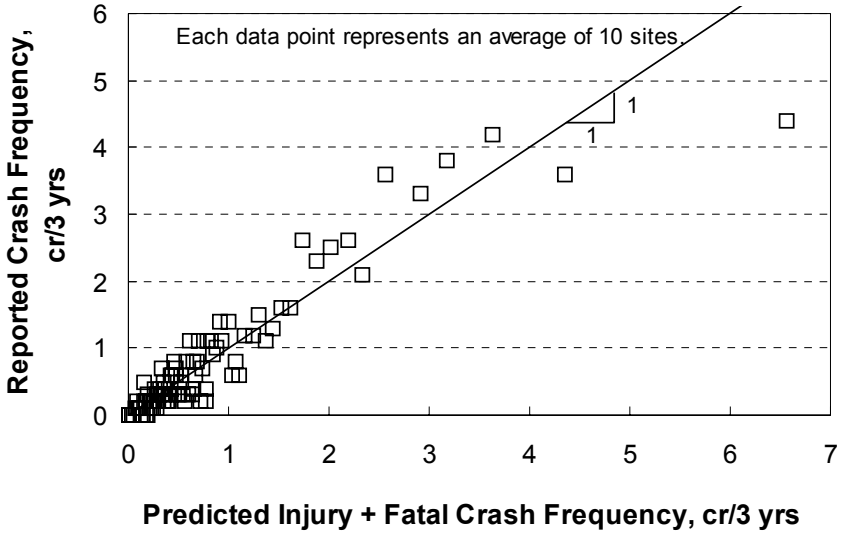


Figure 3-9. Predicted vs. Reported Single-Vehicle Crashes.

Driveway-Related Crashes

The results of the driveway-related safety prediction model calibration are presented in Table 3-13. Calibration of this model also focused on injury (plus fatal) crash frequency. The Pearson χ^2 statistic for the model is 912, and the degrees of freedom are 925 (= $n - p$ = 942 – 17). As this statistic is less than $\chi^2_{0.05,\,925}$ (= 997), the hypothesis that the model fits the data cannot be rejected. The R^2 for the model is 0.70, and the R_k^2 for the calibrated model is 0.80. This statistic indicates that about 80 percent of the variability due to systematic sources is explained by the model.

Table 3-13. Prediction Model Statistical Description–Driveway-Related Model.

Model Statistics	Value
R^2 (R_k^2):	0.70 (0.80)
Scale Parameter φ:	0.97
Pearson χ^2:	912 ($\chi^2_{0.05,\,925}$ = 997)
Over-Dispersion Parameter k:	1.89
Observations n_o:	942 segments (3332 injury + fatal crashes in 3 years)
Standard Deviation s_e:	±1.36 crashes/yr

The coefficients in Table 3-10 were combined with Equations 39 to 45 to obtain the calibrated safety prediction models for driveway-related crashes. The form of each model is:

$$E[N]_{dw2u} = (n_{res} + n_{ind}\,e^{0.274} + n_{bus}\,e^{1.413} + n_{off}\,e^{1.068})\,e^{-2.120 + 1.038\,Ln(ADT/15000) + 0.518\,Ln(S_d)} \quad (63)$$

$$E[N]_{dw2n} = (n_{res} + n_{ind}\,e^{0.274} + n_{bus}\,e^{1.413} + n_{off}\,e^{1.068})\,e^{-2.27 + 1.285\,Ln(ADT/15000) + 0.518\,Ln(S_d)} \quad (64)$$

$$E[N]_{dw4u} = (n_{res} + n_{ind}\,e^{0.274} + n_{bus}\,e^{1.413} + n_{off}\,e^{1.068})\,e^{-2.287 + 1.038\,Ln(ADT/15000) + 0.518\,Ln(S_d)} \quad (65)$$

$$E[N]_{dw4n} = (n_{res} + n_{ind}\,e^{0.274} + n_{bus}\,e^{1.413} + n_{off}\,e^{1.068})\,e^{-2.604 + 1.285\,Ln(ADT/15000) + 0.518\,Ln(S_d)} \quad (66)$$

$$E[N]_{dw4r} = (n_{res} + n_{ind}\,e^{0.274} + n_{bus}\,e^{1.413} + n_{off}\,e^{1.068})\,e^{-2.411 + 1.255\,Ln(ADT/15000) + 0.518\,Ln(S_d)} \quad (67)$$

$$E[N]_{dw6n} = (n_{res} + n_{ind}\,e^{0.274} + n_{bus}\,e^{1.413} + n_{off}\,e^{1.068})\,e^{-2.612 + 1.285\,Ln(ADT/15000) + 0.518\,Ln(S_d)} \quad (68)$$

$$E[N]_{dw6r} = (n_{res} + n_{ind}\,e^{0.274} + n_{bus}\,e^{1.413} + n_{off}\,e^{1.068})\,e^{-2.722 + 1.255\,Ln(ADT/15000) + 0.518\,Ln(S_d)} \quad (69)$$

The fit of the calibrated models is shown in Figure 3-10. This figure compares the predicted and reported crash frequency in the calibration database. Each data point shown represents the average predicted and average reported crash frequency for a group of 10 segments. In general, the data shown in the figure indicate that the model provides an unbiased estimate of expected crash frequency for segments experiencing up to 25 driveway-related crashes in a three-year period.

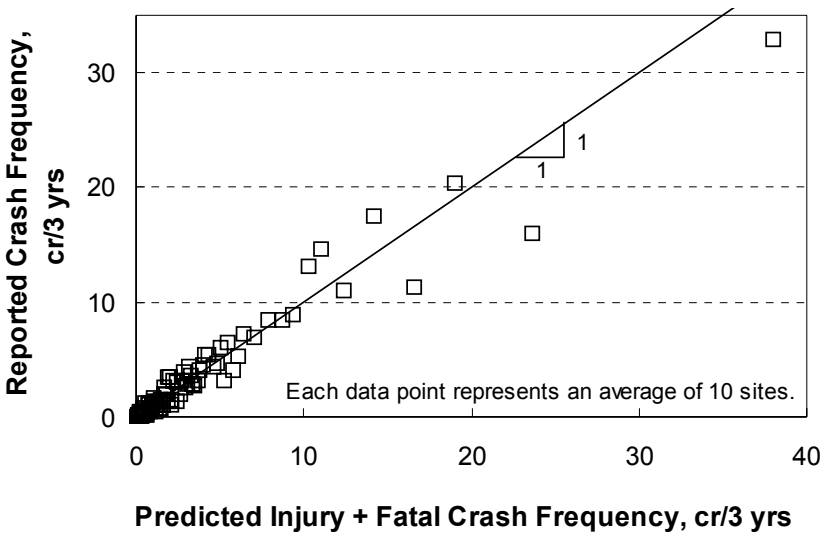


Figure 3-10. Predicted vs. Reported Driveway-Related Crashes.

Calibrated AMFs

Three AMFs were calibrated in conjunction with the safety prediction model calibration. All three were calibrated using injury (plus fatal) crash data. The three AMFs describe the relationship between shoulder width, nonrestrictive median width, and truck presence on crash frequency.

Shoulder Width AMF. The calibrated shoulder width AMF is:

$$AMF_{sw,i} = \left(e^{-0.0317\,(W_s - 1.5)} - 1.0 \right) \frac{P_i}{0.11} + 1.0 \qquad \textbf{(70)}$$

The shoulder width AMF is shown in Figure 3-11. It is limited to shoulder widths up to 5 ft. Three lines are used to illustrate this AMF for the range of proportions associated with the crash types most influenced by shoulder width. These proportions were previously listed in Table 3-5.

The thin dashed line in Figure 3-11 is reproduced from the *Workbook* to illustrate the existing shoulder width AMF trend for four-lane cross sections and six-lane raised-curb cross sections. This thin dashed line is comparable to the thick dashed line representing the proposed AMF for similar cross sections. It indicates that the proposed AMF indicates a slightly more significant reduction in crash frequency with an increase in shoulder width. Similar trends are found for the other cross section types. It should be noted that 80 percent of the segments used to calibrate this AMF had a curb present but that the specific effect of curb presence was undetectable.

Nonrestrictive Median Width AMF. The calibrated nonrestrictive median width AMF is:

$$AMF_{nmw} = e^{-0.0255\,(W_{nm} - 12)} \qquad \textbf{(71)}$$

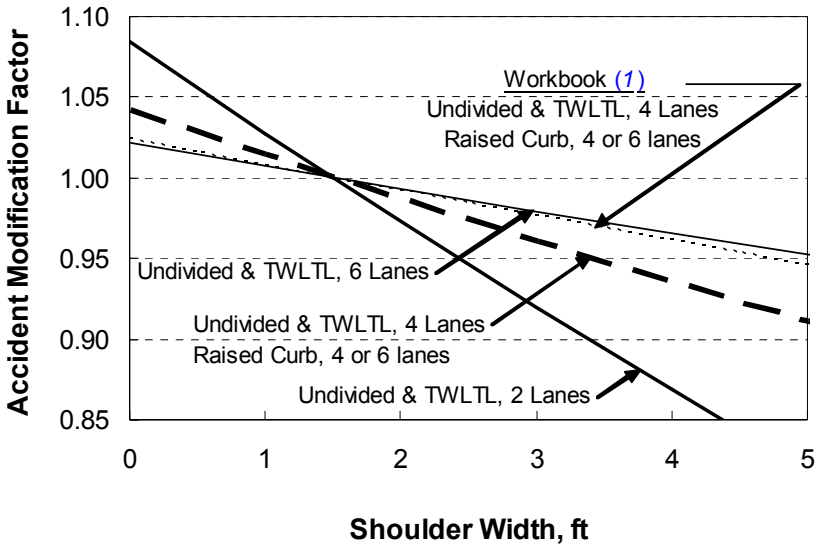


Figure 3-11. Calibrated Shoulder Width AMF.

The nonrestrictive median width AMF is shown in Figure 3-12 using a thick trend line. It is limited to nonrestrictive median widths in the range of 10 to 17 ft. The (restrictive) median width AMF described in the *Workbook* is also shown in the figure. A comparison of the two trend lines indicates that a change in restrictive median width results in a smaller change in the AMF value (and safety) than the same change in the nonrestrictive median width. This trend is logical given the heavy use of nonrestrictive medians by turning vehicles and the nonrestrictive median's lack of a raised-curb (or other restrictive element) to provide some type of positive separation between opposing traffic streams.

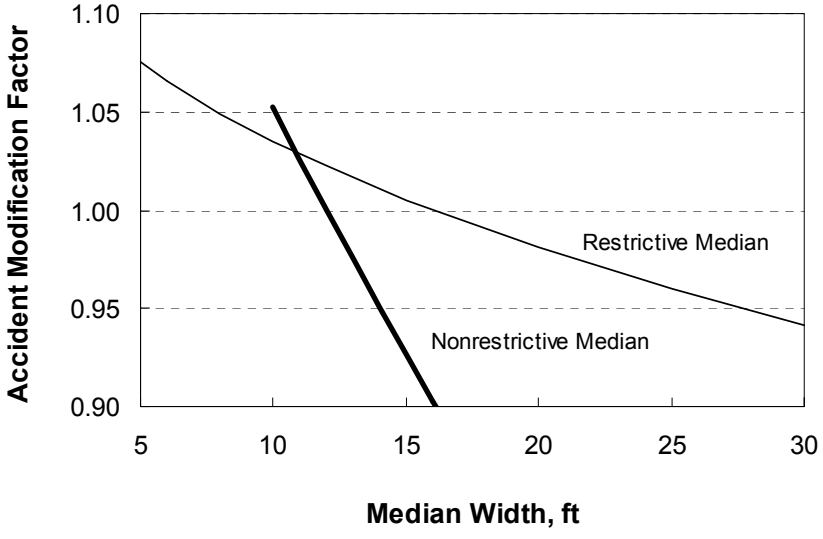


Figure 3-12. Calibrated Nonrestrictive Median Width AMF.

Truck Presence AMF. The calibrated truck presence AMF is:

$$AMF_{tk} = e^{-0.0068\,(P_t - 6.0)} \quad (72)$$

This AMF is shown in Figure 3-13 using a thick trend line. It is limited to truck percentages in the range of 0 to 20 percent. The truck presence AMF described in the *Workbook* is also shown in the figure. It indicates a more significant influence of truck presence than is evident in the data used to calibrate the proposed AMF.

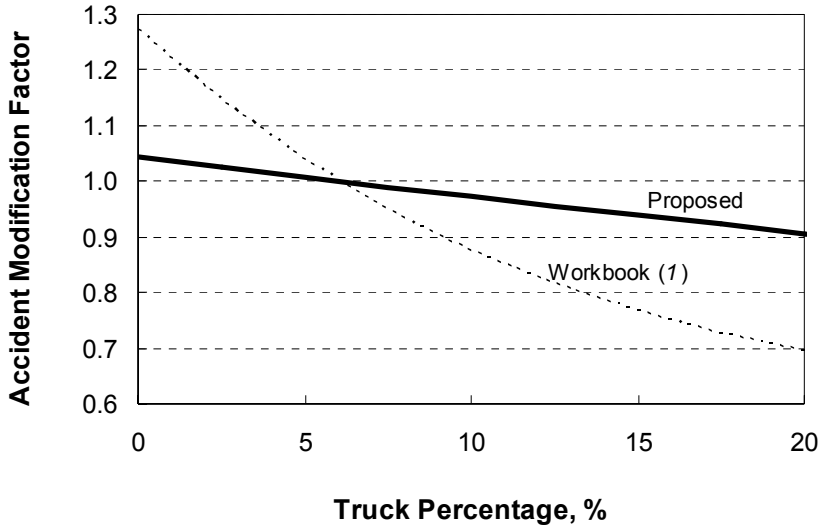
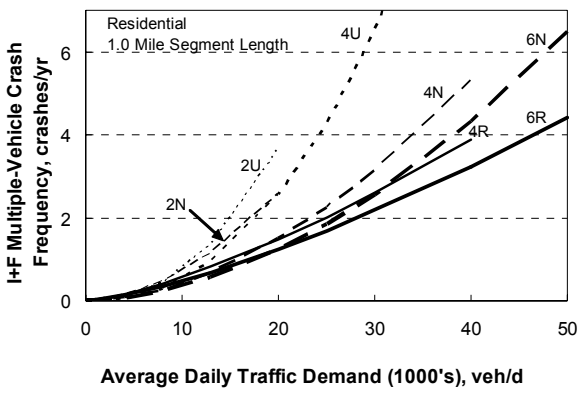


Figure 3-13. Calibrated Truck Presence AMF.

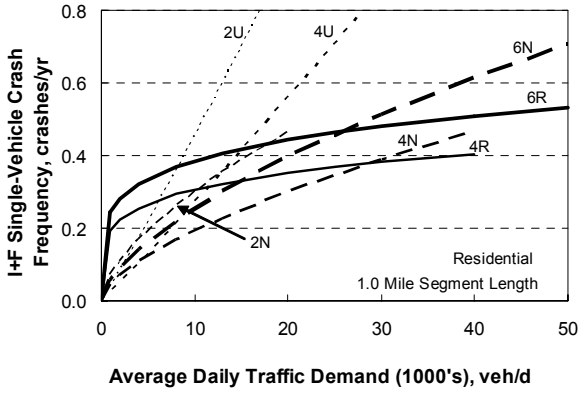
The proposed AMF shown in Figure 3-13 indicates that streets with higher percentages of trucks are associated with fewer crashes. This trend may suggest that drivers are more cautious when there are many trucks present in the traffic stream. It may also suggest that streets with higher truck percentages are more likely to have design dimensions that are "generous" (relative to the design control limits) and thus provide some additional safety benefit.

Sensitivity Analysis

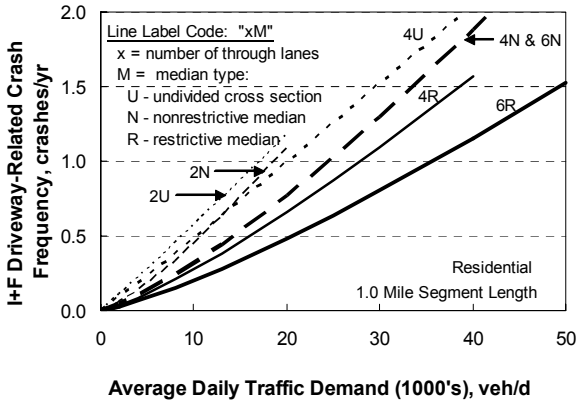
The relationship between crash frequency and traffic demand, as obtained from calibrated Equation 20, is illustrated in Figure 3-14 for a 1-mile street segment. The individual component models are illustrated in Figures 3-14a, 3-14b, and 3-14c. Figure 3-14c illustrates the expected crash frequency when 30 residential driveways are present. These driveways were assumed to be "full" driveways for the undivided cross section and nonrestrictive median type. They were assumed to be "partial" driveways for the restrictive median type. The sum of the individual component crash frequencies is illustrated in Figure 3-14d. The trends in this figure are comparable to those in Figures 3-1a and 3-2 since they are also based on residential land use.



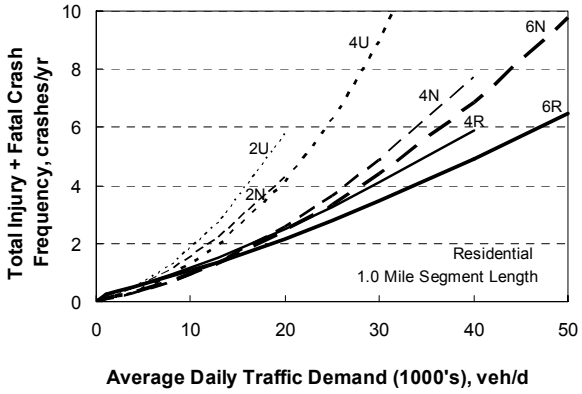
a. Multiple-Vehicle Crashes.



b. Single-Vehicle Crashes.



c. Driveway-Related Crashes.



d. Total Crash Frequency.

Figure 3-14. Proposed Crash Prediction Model.

For an ADT of 20,000 veh/d, the trends in Figure 3-14d indicate that a four-lane undivided street in a residential area has an expected crash frequency of 4.1 crashes/yr (11.7 crashes/yr if located in a business area). In contrast, a four-lane divided street in a residential area has an expected crash frequency of 2.5 crashes/yr (7.2 crashes/yr if located in a business area). When compared with the trends in Figure 3-1a, the expected crash frequency in Figure 3-14d is between 85 and 155 percent of that from Figure 3-1a for any given volume. More generally, the expected crash frequency obtained from Equation 20 is within 20 percent of that obtained from the model in the *Workbook* for similar traffic, land use, and geometric conditions.

Model Extensions

Components of the proposed safety prediction model are calibrated for segments that have a nonrestrictive median type. This median type includes both the TWLTL and the flush-paved median. The majority of the segments in the database have the TWLTL median; however, the safety

experience of the two median types is similar, and the calibrated model can be applied to either type with equal accuracy.

Components of the proposed safety prediction model are also calibrated for segments that have a restrictive median type. This median type includes both the raised-curb and the depressed median. The majority of the segments in the database have the raised-curb median; however, the safety experience of the two median types is similar, and the calibrated model can be applied to either type with equal accuracy.

The proposed safety prediction model was calibrated using a limited amount of data for eight-lane segments. In fact, no data were available for eight-lane segments with a nonrestrictive median. The safety experience of eight-lane cross sections is not believed to be the same as that of four- or six-lane cross sections. Thus, the calibrated model should *not* be applied to arterial streets with eight or more lanes.

The component model for predicting driveway-related crashes is intended for the evaluation of segments with varying numbers of driveways and associated land uses. This model is not sufficiently refined to be able to estimate the effect of installing or removing a single driveway from a street segment. If the analyst desires to conduct this type of evaluation, then the driveway should be evaluated as an intersection using the intersection safety prediction models.

REFERENCES

1. Bonneson, J.A., K. Zimmerman, and K. Fitzpatrick. *Interim Roadway Safety Design Workbook*. FHWA/TX-06/0-4703-P4, Texas Department of Transportation, Austin, Texas, April 2006.

2. Harwood, D., K. Bauer, K. Richard, D. Gilmore, J. Graham, I. Potts, D. Torbic, and E. Hauer. *Methodology to Predict the Safety Performance of Urban and Suburban Arterials.* Final Report. NCHRP Project 17-26. National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C., March 2007.

3. Bonneson, J.A., K. Zimmerman, and K. Fitzpatrick. *Roadway Safety Design Synthesis.* FHWA/TX-05/0-4703-P1. Texas Department of Transportation, Austin, Texas, November 2005.

4. Bonneson, J., D. Lord, K. Zimmerman, K. Fitzpatrick, and M. Pratt. *Development of Tools for Evaluating the Safety Implications of Highway Design Decisions.* FHWA/TX-07/0-4703-4. Texas Department of Transportation, Austin, Texas, September 2006.

5. Harwood, D.W., F.M. Council, E. Hauer, W.E. Hughes, and A. Vogt. *Prediction of the Expected Safety Performance of Rural Two-Lane Highways.* FHWA-RD-99-207. Federal Highway Administration, Washington, D.C., 2000.

6. Hauer, E., F.M. Council, and Y. Mohammedshah. "Safety Models for Urban Four-Lane Undivided Road Segments." *Transportation Research Record 1897*. Transportation Research Board, Washington, D.C., 2004.

7. Bonneson, J., M. Pratt, J. Miles, and P. Carlson. *Development of Guidelines for Establishing Effective Curve Advisory Speeds.* FHWA/TX-07/0-5439-1. Texas Department of Transportation, Austin, Texas, 2007.

8. Hauer, E. *Safety of Horizontal Curves.* Unpublished paper. Published: March 24, 2000. Accessed at: http://www.roadsafetyresearch.com/, March 2008.

9. Lord, D. "Modeling Motor Vehicle Crashes Using Poisson-Gamma Models: Examining the Effects of Low Sample Mean Values and Small Sample Size on the Estimation of the Fixed Dispersion Parameter." *Accident Analysis & Prevention*, Vol. 38, No. 4, Elsevier Ltd., Oxford, Great Britain, July 2006, pp. 751-766.

10. *SAS OnlineDoc, Version 8*, SAS Institute, Inc., Cary, North Carolina, 1999.

11. Miaou, S.P. *Measuring the Goodness-of-Fit of Accident Prediction Models.* FHWA-RD-96-040. Federal Highway Administration, Washington, D.C., 1996.

CHAPTER 4. RURAL MULTILANE HIGHWAY SEGMENTS

OVERVIEW

This chapter describes the activities undertaken to calibrate safety prediction models for rural multilane highways in Texas. Each model provides an estimate of the expected crash frequency for a highway segment with a specified length, traffic volume, and geometric design. A series of accident modification factors were also developed as part of the calibration activity. Three models were calibrated and are recommended for use in Texas. Collectively, these models address the following cross section combinations:

- undivided cross section with four through lanes,
- divided cross section with a nonrestrictive median and four through lanes, and
- divided cross section with a restrictive median and four through lanes.

Equivalent models are described in the *Interim Roadway Safety Design Workbook* (*1*). These models were developed using data from locations outside of Texas. They are replaced by the models described in this chapter, all of which have been calibrated using data specific to Texas.

The objective of this research was to calibrate models for each of the aforementioned cross section combinations using crash data for Texas highways. Highways with two-way traffic flow were the focus of the calibration activity. As part of the calibration process, several AMFs were also developed. These AMFs can be used with the calibrated models to estimate the change in safety associated with a specified change in geometry.

This chapter is divided into four parts. The first part provides a review of the literature on the topic of safety prediction models for rural highway segments. The second part describes a general AMF model framework for extending selected AMFs to a range of cross sections. The third part describes the method used to calibrate the proposed models. The last part describes the analysis and findings from the calibration process.

LITERATURE REVIEW

This part of the chapter describes a framework for safety prediction model calibration. It consists of five sections. The first section identifies the various median types found on rural highways and establishes a vocabulary in this regard. The second section provides a brief overview of safety prediction and the role of "base" prediction models. The third section summarizes the base models described in the *Workbook*. The fourth section summarizes the models developed by Lord et al. for NCHRP Project 17-29 (*2*). The last section reviews the data available in TxDOT's road inventory database.

Median Types in Rural Areas

A detailed review of about 2600 miles of rural multilane highway segments in Texas indicates that about 80 percent of these miles have a divided cross section with one of the following median types:

- two-way left-turn lane (TWLTL) median,
- flush-paved median,
- positive (rigid) barrier median,
- raised-curb median, or
- depressed median.

The first two median types listed are described herein as "nonrestrictive" median types. They describe the type of median used on about 5 percent of the highway mileage reviewed. Of those highways with a nonrestrictive median, about 97 percent have a TWLTL median.

The last three median types listed are described herein as "restrictive" median types. They describe the type of median used on about 75 percent of the highway mileage reviewed. Of those highways with a restrictive median, about 96 percent have a depressed median and 2.5 percent have a positive barrier median. It is noted that the positive barrier in the database described in a subsequent section is a "rigid" barrier. The database does not include highways with a cable barrier.

Undivided cross sections have no median. This cross section was found on about 20 percent of the highway-miles reviewed.

Safety Prediction Models

The expected crash frequency for a highway segment with specified attributes is computed using a safety prediction model. This model represents the combination of a "base" model and one or more AMFs. The base model is used to estimate the expected crash frequency for a typical segment. The AMFs are used to adjust the base estimate when the attributes of the specific segment are not considered typical. The general form of the typical safety prediction model is shown below as Equation 1, and its base model component is shown as Equation 2.

with,

$$E[N] = E[N]_b \times AMF_{lw} \times AMF_{mw} \ldots \quad \textbf{(1)}$$

$$E[N]_b = a\ ADT^{\,b}\ L \quad \textbf{(2)}$$

where,

$E[N]$ = expected crash frequency, crashes/yr;
$E[N]_b$ = expected base crash frequency, crashes/yr;
AMF_{lw} = lane width accident modification factor;
AMF_{mw} = (restrictive) median width accident modification factor;
a, b = calibration coefficients;
ADT = average daily traffic volume, veh/d; and
L = segment length, mi.

When the coefficient *b* for the *ADT* variable is equal to 1.0, then the coefficient *a* in Equation 2 is effectively equal to the crash rate, with units of "crashes per million vehicle-miles."

Interim Workbook Base Models

This section describes base models derived by Bonneson et al. (*3*) for rural highway segments. These models were calibrated using data and models described in the research literature. They were subsequently documented in the *Workbook* in the form of a base model and associated base crash rates. The base model in the *Workbook* is:

$$E[N]_b = 0.000365\ Base\ ADT\ L \quad \textbf{(3)}$$

where,

Base = injury (plus fatal) crash rate (from Table 4-1), crashes/million vehicle-miles.

A comparison of Equation 3 with Equation 2 indicates that the constant *a* is equal to the product of the base crash rate *Base* and "0.000365." The constant represents a conversion factor to convert the ADT into "millions of vehicles per year." Also, the ADT exponent *b* is equal to 1.0 in Equation 3.

The crash rate used in Equation 3 is provided in Table 4-1. They are based on a synthesis of the trends reported by several researchers in their examination of rural highway crashes (*3*). Rates were not derived for two-lane highways with depressed median or six-lane undivided highways.

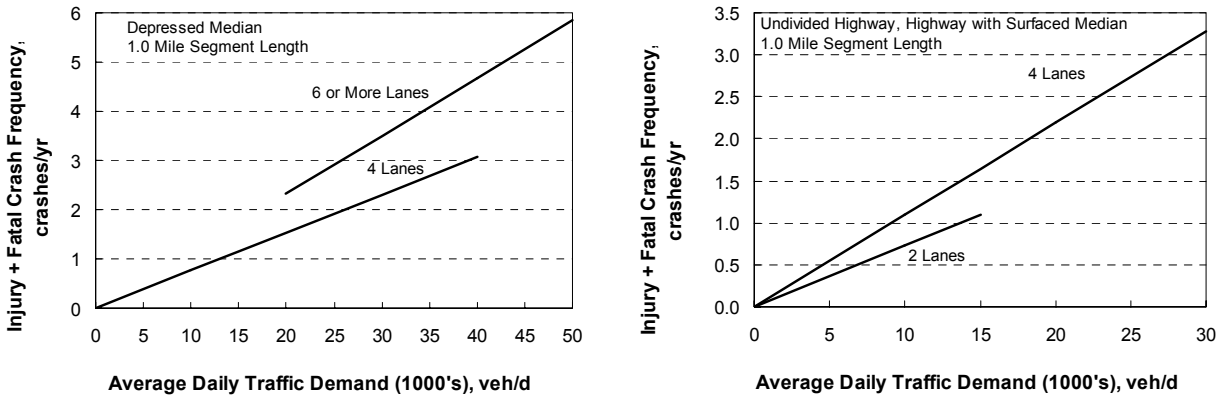
Table 4-1. Base Crash Rates for Rural Highway Segments.

Median Type	Attributes	Base Crash Rate, severe crashes/mvm [1]		
	Through Lanes:	2	4	6
Undivided or Surfaced [2]		0.20	0.30	data not available
Depressed		data not available	0.21	0.32

Notes:
1 - mvm: million vehicle miles.
2 - Surfaced: flush paved or TWLTL. Rates for the TWLTL must be adjusted using the TWLTL median type AMF.

The relationship between crash frequency and traffic demand (represented by Equation 3 and the rates in Table 4-1) is shown in Figure 4-1 for a 1-mile highway segment. The trends indicate that undivided cross sections have about 40 percent more crashes than divided cross sections for four-lane highways. The trends also indicate that undivided four-lane highways have more crashes than undivided two-lane highways, for the same daily traffic demand. A similar trend is shown when comparing six-lane highways to four-lane highways with a depressed median.

For an ADT of 10,000 veh/d, the trends in Figure 4-1 indicate that a four-lane undivided highway has an expected crash frequency of 1.1 crashes/yr. In contrast, a four-lane divided highway (with a depressed median) has an expected crash frequency of 0.77 crashes/yr.



a. Depressed Median. **b. Undivided Highway or Surfaced Median.**

Figure 4-1. Crash Prediction Model Reported in the *Workbook*.

The base model in Equation 3 is applicable to highway segments that have typical geometric design conditions. These conditions are identified in the *Workbook* and are restated in column two of Table 4-2. The base conditions listed are representative of roads in Texas that are classified as rural multilane highways. The base model is combined with the AMFs described in the *Workbook* to yield the safety prediction model for highway segments.

Table 4-2. Base Conditions for Rural Highway Segments.

Characteristic	Base Condition	Sensitivity of Continuous Variables
Horizontal curve radius	tangent (no curve)	5000 ft radius or flatter
Grade	flat (0% grade)	-3 to +3%
Lane width	12 ft	11 to 12 ft
Paved outside shoulder width	8 ft	7 to 9 ft
Paved inside shoulder width	4 ft	2 to 6 ft
Restrictive median width	76 ft	60 to 92 ft
Nonrestrictive median width	16 ft	8 to 28 ft
Shoulder rumble strips	not present	--
Centerline rumble strips	not present	--
Superelevation	not deficient	0 to 1% deficiency
Horizontal clearance	30 ft	20 ft or more
Side slope	1V:4H	1V:2.5H or flatter
Utility pole density and offset	25 poles/mi; 30 ft average offset	Any density; 20 ft offset or more
Driveway density	5 driveways/mi	0 to 12 driveways/mi

The characteristics in column 1 of Table 4-2 are each associated with an AMF to account for segment conditions that deviate from the base conditions. Each AMF was evaluated for a range of values to examine its sensitivity to a change in the associated variable. The results of this sensitivity analysis are listed in column 3 of Table 4-2. The ranges listed in these columns equate to AMF values between 0.95 and 1.05. For example, an outside shoulder width between 7 and 9 ft yields AMF values of about 1.05 and 0.95, respectively. The ranges listed provide some indication of the correlation between a change in the characteristic and the change in crash risk.

NCHRP Project 17-29 Base Models

This section describes base models derived by Lord et al. (*2*) for four-lane rural multilane highway segments. The model for undivided segments is based on data from Washington and Texas. The model for divided segments is based on data from California and Texas. The base model they developed is described using the following equations:

$$E[N]_{un} = -11.4448\ ADT^{1.2870}\ L \quad \textbf{(4)}$$

$$E[N]_{div} = -10.2326\ ADT^{1.1714}\ L \quad \textbf{(5)}$$

where,
$E[N]_{un}$ = expected total crash frequency for rural four-lane undivided highway, crashes/yr; and
$E[N]_{div}$ = expected total crash frequency for rural four-lane divided highway, crashes/yr.

The two models represented by Equations 4 and 5 represent base conditions. They were developed to estimate "total" crash frequency (i.e., K, A, B, C, and property-damage-only crashes). For consistency in the discussion of models presented in this chapter, each model was multiplied by 0.38 to convert total crash frequency to "injury plus fatal" (K,A,B,C) crash frequency. The constant of 0.38 represents the ratio of KABC to total crashes for Texas data.

The relationship between crash frequency and traffic demand, as obtained from Equations 4 and 5, is illustrated in Figure 4-2 for a 1-mile highway segment. The trends in Figure 4-2 are comparable to those in Figure 4-1 for four-lane highways.

The trends in Figure 4-2 indicate that undivided cross sections have about 14 percent *fewer* crashes than divided cross sections. This relationship is contrary to that found in Figure 4-1, where undivided cross sections were found to have 40 percent *more* crashes than divided highways.

For an ADT of 10,000 veh/d, the trends in Figure 4-2 indicate that a four-lane undivided highway has an expected crash frequency of 0.56 crashes/yr. In contrast, a four-lane divided highway has an expected crash frequency of 0.65 crashes/yr. When compared with the trends in Figure 4-1 for undivided cross sections, the expected crash frequency in Figure 4-2 is about 50 percent of that from Figure 4-1. That is, the expected crash frequency obtained from Equation 4 is notably lower than the crash frequency obtained from the models used to derive Figure 4-1 (*3*). The expected crash frequency from Equation 5 is about 85 percent of that for divided highways in Figure 4-1. These trends may be partially due to differences in the base conditions for each model. Regardless, they are evidence of the need to calibrate crash prediction models to "local" conditions.

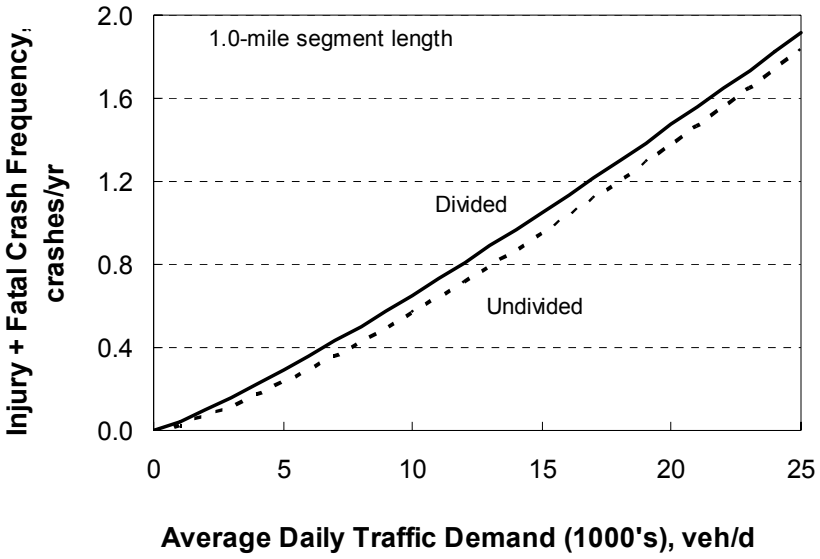


Figure 4-2. Crash Prediction Model Reported by Lord et al.

Texas Highway Data

The geometric and traffic attributes for the Texas state highway system were obtained from the Texas Reference Marker (TRM) system highway database. This database is maintained by TxDOT and contains data for several thousand highway segments, each of which is described in terms of its geometry, traffic, and location attributes. In fact, about 140 attributes are used to describe each highway segment in TRM. However, only about 20 attributes are used to describe segment geometry or traffic characteristics, and about 10 more attributes are used to describe road name and physical location. The remaining attributes describe administrative designations and road-management-related data that are not particularly relevant to safety analysis.

A highway segment in TRM is defined as a length of road along which no one attribute changes. A change in one attribute dictates the end of one segment and the start of a new segment. The average tangent (i.e., not curved) segment length in TRM is about 0.4 mi.

Several segment characteristics are needed in the calibration database. They are identified in Table 4-3. Column 3 of this table indicates whether the corresponding characteristic is available in the TRM database.

The information in Table 4-3 indicates that data for several segment characteristics are not available in the TRM database. However, it was possible to obtain lane width, median type, median width, and driveway location data using high-resolution aerial photography available from the Internet. Utility pole density, grade, rumble strip presence, and side slope data could not be reliably obtained from any source. It was judged that their base condition values would be sufficiently representative of the collective set of segments in the database. This assumption allowed the corresponding AMF value to be set to 1.0.

Table 4-3. Rural Highway Segment Calibration Data.

Attribute Category	Attribute	Availability in TRM Database
General	Area type (e.g., urban, rural)	Yes
Traffic control	Two-way vs. one-way operation	Yes
Traffic characteristics	ADT (by specified year)	Yes
	Truck percentage	Yes
Geometry	Horizontal curve radius	Yes
	Number of lanes	Yes
	Lane width	Yes for restrictive median types.
	Inside & outside shoulder width	Yes
	Median type	Partially, data does not distinguish between no median, TWLTL median, and flush-paved median.
	Median width	Yes for restrictive median types.
	Utility pole density & offset	No
	Curb presence	Yes
	Driveway location and use	No
	Grade	No
	Rumble strip presence	No
	Side slope	No
	Horizontal clearance	No

For restrictive medians, median width was measured as the distance between the near edges of the traveled way for the two opposing travel directions. Thus, median width includes the width of both inside shoulders.

AMF DEVELOPMENT

This part of the chapter describes the development of a general AMF model for incorporating the influence of median type and number-of-lanes into those AMFs that are associated with only specific crash types (e.g., run-off-the-road). The general model was first introduced in Chapter 3 of the *Synthesis* (*3*) and is derived in Chapter 3. It is extended to intersection applications in Chapter 2.

The general model is appropriate when a treatment applied to a segment (e.g., change shoulder width) has its most significant effect on specific crash types. The magnitude of this effect can be quantified using an AMF that is based on just the crashes of the specified types (i.e., AMF_{spc}). Alternatively, it can also be quantified using an AMF based on all crash types (i.e., AMF_{seg}). The general AMF model is:

$$AMF_{seg,j} = \left(AMF_{seg,i} - 1\right)\frac{P_{spc,j}}{P_{spc,i}} + 1.0 \qquad (6)$$

where,

$AMF_{seg,j}$ = AMF for a specified treatment to a segment with cross section type *j*, quantified in terms of all crash types;
$AMF_{seg,i}$ = AMF for the same treatment to a segment with cross section type *i*, quantified in terms of all crash types;
$P_{spc,j}$ = proportion of specific crash types on a segment with cross section *j*; and
$P_{spc,i}$ = proportion of same crash types on a segment with cross section *i*.

The proportion of specific crash types used in Equation 6 is provided in the *Workbook* (*1*) for several AMFs. These proportions were based on Texas crash data for the years 1999 and 2000. Since their derivation, crash data were made available for year 2001. These data were used to update the proportions to reflect a three-year average (i.e., 1999, 2000, and 2001). The results of this analysis are shown in Table 4-4. The proportions shown in each row are based on the total listed in the last column.

Table 4-4. Rural Highway Crash Distribution.

Median Type [1]	Through Lanes	Proportion of Injury + Fatal Crashes					Total Crash Frequency [2], crashes per 3 years
		Single-Vehicle		Multiple-Vehicle		Single-Vehicle with Pole	
		Either Side	Right Side	Opp. Dir.	Sideswipe		
Undivided or Nonrestrictive	2	0.5152	0.2830	0.0872	0.0084	0.0330	34,493
	4	0.3234	0.2082	0.0783	0.0343	0.0350	6,117
Restrictive	4	0.5102	0.3029	0.0284	0.0473	0.0409	8,394
	6	0.4170	0.2326	0.0298	0.0468	0.0298	235

Notes:
1 - Nonrestrictive: TWLTL or flush-paved median. Restrictive: raised-curb, positive barrier, or depressed median.
2 - Injury (plus fatal) crashes reported during the years 1999, 2000, and 2001.

The proportions listed in Table 4-4 were combined as needed for three AMFs in the *Workbook*. The combinations reflect the types of crashes that were found to be influenced by a change in the corresponding geometric element, as a function of median type and number of lanes. The resulting proportions to be used in Equation 6 are listed in Table 4-5. The values listed in this table are very similar to those reported in the *Workbook*.

METHODOLOGY

This part of the chapter describes the methodology used to calibrate the multilane highway segment safety prediction model. It is divided into three sections. The first section describes the database used to calibrate the model. The second section describes the technique used to quantify the calibration coefficients. The last section describes the technique used to adjust the over-dispersion parameter for bias due to a small sample size.

Table 4-5. Rural Highway Crash Distribution for Selected AMFs.

Median Type	Through Lanes	Proportion of Injury + Fatal Crashes				
		Lane Width AMF	Outside Shoulder AMF [1]	Inside Shoulder AMF	Shoulder Rumble Strip	Utility Pole Offset AMF
Undivided or Nonrestrictive	2	0.61 [a]	0.52 [b]	not applicable	0.28 [c]	0.033 [d]
	4	0.44 [a]	0.32 [b]		0.21 [c]	0.035 [d]
Restrictive	4	0.59 [a]	0.30 [c]	0.24 [e]	0.51 [b]	0.041 [d]
	6	0.49 [a]	0.23 [c]	0.21 [e]	0.42 [b]	0.030 [d]

Notes:
a - Single-vehicle run-off-road, same-direction sideswipe, and multiple-vehicle opposite direction crashes.
b - Single-vehicle run-off-road crashes, either side.
c - Single-vehicle run-off-road crashes, right side.
d - Single-vehicle-with-pole crashes.
e - Multiple-vehicle opposite direction and single-vehicle-run-off-road crashes, left side.
1 - Also applies to side slope AMF and horizontal clearance AMF.

Database

A database was assembled from the TRM database for segment calibration. Variables associated with the AMFs described in the *Workbook* (*1*) were included in the database to ensure that their relationship with crash occurrence was included in the analysis. However, as previously noted in the discussion associated with Table 4-3, utility pole, grade, superelevation rate, rumble strip, and side slope data are not in the TRM database. Moreover, measurement of these characteristics in the field was not a feasible option. A review of the aerial photos for each segment indicated (to the extent permitted by photo resolution) that the base condition values listed in Table 4-2 for these variables were representative of the collective set of segments in the database.

Adjacent Land Use

A database was assembled to allow for the exploration of correlation between adjacent land use and crash frequency. Land use was previously found to be correlated with crash frequency on urban and suburban arterial segments. Land use was characterized as residential or undeveloped, industrial, commercial business, and office. Table 4-6 identifies the characteristics of each land use. These characteristics were used to determine the land use adjacent to, and on each side of, the subject segment. The characterization was focused on the land use within a 100-ft wide zone adjacent to the segment, as measured from the outside edge of the near shoulder and away from the segment.

For each highway segment, land use was measured as a length along the segment. It was separately measured on each side of the segment. Driveways and public roadway approaches to the segment were included in the measured length and assigned a land use equivalent to that of the property to which they were adjacent. With this method of measurement, the total length of land use on a segment equaled twice the segment's length.

Table 4-6. Adjacent Land Use Characteristics.

Land Use	Characteristics	Examples
Residential or Undeveloped	• Buildings are small • A small percentage of the land is paved • If driveways exist, they have very low volume • Ratio of land-use acreage to parking stalls is large	• Single-family home • Undeveloped property • Graveyard • Farmland • Park or green-space recreation area
Industrial	• Buildings are large and production oriented • Driveways and parking may be designed to accommodate large trucks • Driveway volume is moderate at shift change times and is low throughout the day • Ratio of land-use acreage to parking stalls is moderate	• Factory • Warehouse • Storage tanks • Farmyard with barns and machinery
Commercial Business	• Buildings are larger and separated by convenient parking between building and roadway • Driveway volume is moderate from mid-morning to early evening • Ratio of land-use acreage to parking stalls is small	• Strip commercial • Shopping mall • Apartment complex • Trailer park • Airport • Gas station • Restaurant
Office	• Buildings typically have two or more stories • Most parking is distant from the building or behind it • Driveway volume is high at morning and evening peak traffic hours; otherwise, it is low • Ratio of land-use acreage to parking stalls is small	• Office tower • Public building • Church • School • Clubhouse (buildings at a park) • Parking lot for "8 to 5" workers

Driveway Count

Driveways along a segment were counted and categorized according to the land use they served. Public roadway approaches and driveways served by a signalized intersection were not included in the count of driveways. The following two types of driveways were identified:

- Full driveway - a driveway that allowed the following four turn movements for access: right-in, right-out, left-in, and left-out.
- Partial driveway - a driveway that allowed right-in and right-out movements but prevented left-turns with channelizing islands, a restrictive median, or one-way driveway operation.

In the count of driveways, a full driveway was counted as "1" driveway, and a partial driveway was counted as "0.5" driveways. This distinction was rationalized as appropriate given the difference in the number of conflicting paths present with each driveway type and the correlation between the number of these paths and driveway crash frequency. Partial driveways were most frequently found on segments with a restrictive median.

Driveways that were apparently unused were not counted. Similarly, driveways leading into fields, small utility installations (e.g., cellular phone tower), and abandoned buildings were not counted. A circular driveway at a residence was counted as one driveway even though both ends of

the driveway intersected the subject segment. A similar approach was used when a small business (e.g., gas station) had two curb cuts separated by only 10 or 20 ft.

Curve Radius and Length

The segments included in the database were selected from the TRM database such that they did not have curves with a radius of less than 5000 ft. This approach was undertaken to allow the calibration activity to focus on the calibration of models for tangent highway segments. However, when reviewing the aerial photograph of each segment, a small percentage of segments were found to have moderate curvature. In this situation, the curve's length and radius were recorded and added to the database.

The curve radius AMF described in Chapter 3 (i.e., Equation 16) was used to account for the presence of curvature. The following equation was then used to compute the aggregate AMF for segments with combined curve and tangent portions:

$$AMF_{cr|agg} = \frac{(L - L_c)\,1.0 - L_c\,AMF_{cr}}{L} \quad (7)$$

where,

$AMF_{cr|agg}$ = aggregated curve radius AMF for a segment with both tangent and curved portions;
AMF_{cr} = horizontal curve radius accident modification factor;
L_c = length of horizontal curve (= $I_c \times R$ / 5280 / 57.3), mi;
I_c = curve deflection angle, degrees; and
R = curve radius, ft.

In those few cases where a curve was found on the segment, it was often located entirely within the segment. However, sometimes only a portion of the curve would be located on the segment. In this situation, the length of curve recorded in the database represented only that portion of the curve that was within the segment boundaries.

Lateral Clearance Distances

Three lateral clearance distances were measured. Horizontal clearance was measured from the outside edge of the traveled way to the nearest continuous obstruction (e.g., fence line, utility poles, etc.). If a barrier is present on the roadside for only a portion of the segment, horizontal clearance was measured only for that portion without the barrier.

Also measured was the "inside clearance distance" and the "outside clearance distance." For nonrestrictive medians, the inside clearance distance is equal to the median width.

For segments with a restrictive median, the following rules were used to compute the inside clearance distance. When a barrier is not present in the median, the clearance distance represents the distance between the near edges of the inside shoulder for the two opposing travel directions (i.e., it equals the median width less the width of the two inside shoulders). When a barrier is present in

the median, an inside clearance distance is computed for both travel directions, as measured from the edge of the inside shoulder to the nearest barrier face (as defined in the next section).

A similar approach was used to measure "outside clearance distance." When a barrier is not present on the roadside, the outside clearance distance represents the distance between the edge of the outside shoulder and the nearest continuous obstruction (i.e., it equals the horizontal clearance less the width of the outside shoulder). When a barrier is present on the roadside, the outside clearance distance represents the distance between the edge of the outside shoulder and the nearest barrier face.

Longitudinal Barriers

Longitudinal barriers (i.e., concrete barrier, guardrail, or bridge rail) were noted when present on a segment. Attributes used to quantify barrier presence include barrier location, offset, length, and width. Barrier location indicates whether the barrier is parallel to the inside edge of the roadway or the outside edge of the roadway. Barrier offset, length, and width were each measured separately for inside and outside locations (e.g., inside barrier offset, outside barrier offset, etc.).

Barrier offset represents a lateral distance measured from the near edge of the shoulder to the face of the barrier. Barrier length represents the length of lane paralleled by a barrier; it is a total for both travel directions (e.g., if the outside barrier extends for the length of the roadway on both sides of the roadway, the outside barrier length equals twice the segment length). Barrier width represents either the physical width of the barrier if only one barrier is used, or the lateral distance between barrier "faces" if two parallel barriers are provided in the median area. A barrier face is the side of the barrier that is exposed to traffic.

Median Openings

Median openings were counted along the length of the segment. Openings associated with the intersection of a public roadway were not included in the count.

Calibration Technique

The calibration activity used statistical analysis software that employs maximum likelihood methods based on a negative binomial distribution of crash frequency. The calibration model consisted of Equation 1, with the following model form used to estimate the base crash frequency:

with,

$$E[N]_b = E[N]_{mv} + E[N]_{sv} + E[N]_{dw} \quad \textbf{(8)}$$

$$E[N]_{mv} = a_{mv}\, ADT^{\,b_{mv}}\, L \quad \textbf{(9)}$$

$$E[N]_{sv} = a_{sv}\, ADT^{\,b_{sv}}\, L \quad \textbf{(10)}$$

$$E[N]_{dw} = \sum_i a_{dw,i}\, (ADT/15000)^{b_{dw}}\, n_i \quad \textbf{(11)}$$

where,

$E[N]_{mv}$ = expected multiple-vehicle nondriveway crash frequency for base conditions, crashes/yr;
$E[N]_{sv}$ = expected single-vehicle crash frequency for base conditions, crashes/yr;
$E[N]_{dw}$ = expected driveway-related crash frequency for base conditions, crashes/yr;
a_{mv}, b_{mv} = calibration coefficients for multiple-vehicle nondriveway crashes;
a_{sv}, b_{sv} = calibration coefficients for single-vehicle crashes;
$a_{dw,i}$, b_{dw} = calibration coefficients for driveway-related crashes for driveway type i; and
n_i = number of driveways of type i.

The form of this model was suggested by Harwood et al. (*4*). It was subsequently used in Chapter 3 for the development of the urban and suburban arterial street segment models.

The models represented by Equations 9, 10, and 11 are referred to as "component models." Equations 1 and 8 through 11 represent one "aggregate model" and were simultaneously analyzed to find the best-fit calibration coefficients. In this manner, the component models were calibrated to compute the expected frequency of the corresponding crash type for a common set of AMFs (as shown in Equation 1). The analysis approach is described in more detail in the next part of this chapter.

Over-Dispersion Parameter Adjustment

It was assumed that crash frequency on a segment is Poisson distributed, and that the distribution of the mean crash frequency for a group of similar intersections is gamma distributed. In this manner, the distribution of crashes for a group of similar intersections can be described by the negative binomial distribution. The variance of this distribution is:

$$V[X] = y\,E[N] + \frac{(y\,E[N])^2}{k\,L} \quad \textbf{(12)}$$

where,

$V[X]$ = crash frequency variance for a group of similar locations, crashes^2;
X = reported crash count for y years, crashes;
y = time interval during which X crashes were reported, yr; and
k = over-dispersion parameter, mi^{-1}.

In recognition of the concerns expressed by Lord (*5*), the over-dispersion parameter k obtained from the regression analysis is adjusted downward, such that it yields a more reliable estimate of the variance of the crash distribution. The derivation of the adjustment technique used is described in Reference *6*. The following equation was used to estimate the true over-dispersion parameter, given the estimated over-dispersion parameter obtained from the regression analysis.

$$k_t = \frac{\sqrt{(n-p)^2\,m^2 + 69\,k_r\,L.\,(n-p)\,m} - (n-p)\,m}{34.5\,L.} \quad \textbf{(13)}$$

where,

k_t = true over-dispersion parameter;
k_r = estimated over-dispersion parameter obtained from database analysis;
n = number of observations (i.e., segments in database);

p = number of model variables;
L. = average segment length for all n observations (= 1.0 for intersection applications); and
m = average number of crashes per observation (= total crashes in database / n).

Equation 13 was used to estimate the true over-dispersion parameter for each of the models described in this chapter. All subsequent references to the over-dispersion parameter k in this chapter denote the estimated true parameter obtained from Equation 13 (i.e., hereafter, $k = k_t$).

CALIBRATION DATA AND ANALYSIS

This part of the chapter describes the calibration of the aggregate rural multilane highway segment safety prediction model. Initially, the process used to select the segments is identified. Then, the characteristics of the segments are summarized. Next, the techniques used to assemble the crash data segment are described. Finally, the findings from the model calibration are discussed. In this section, a "site" is used to refer to one highway segment.

Site Selection and Data Collection

The highway segments used to calibrate the safety prediction model were required to satisfy the following criteria:

- Area type: rural;
- Functional class: arterial or collector;
- Segment length: 0.1 mi or more; and
- Curve presence: no curves with a radius less than 5000 ft.

All segments were required to have a minimum length of 0.1 mi. The value of 0.1 mi reflects a recognition of the precision of crash location in TxDOT's state-wide crash database.

The TRM database was screened to exclude segments that had one or more curves with a radius less than 5000 ft. This criterion was intended to allow the calibration to focus on tangent segments. However, a visual inspection of aerial photography for each segment revealed the presence of a curve on some segments. Those segments with flatter curves were kept in the database, and the effect of curvature was modeled using Equation 7. Segments with sharp curvature were removed from the database.

Research by Harwood et al. (*4*) and the synthesis by Bonneson et al. (*3*) indicate that segment crash frequency is significantly influenced by cross section (i.e., median type and number or lanes). This influence is so significant and varied that a separate model should be calibrated for each unique combination of median type and lanes that exist in the database. A review of the TRM database indicated that the following three cross section combinations are most common in Texas:

- undivided cross section with four through lanes (4U),
- divided cross section with a nonrestrictive median and four through lanes (4N), and
- divided cross section with a restrictive median and four through lanes (4R).

Thus, in addition to the aforementioned site selection criteria, effort was expended during the site selection process to target enough segments for each combination such that separate models could be accurately calibrated. To achieve this goal, a minimum sample size criterion based on total exposure was established for each combination.

Site Characteristics

Traffic and geometry data were identified for several thousand highway segments using the TRM database for year 2003, the first year for which a complete database was available. All total, 562 segments satisfied the site selection criteria. These segments represent highways located in 23 of the 25 TxDOT districts. Selected characteristics are provided for these segments in Table 4-7.

Table 4-7. Summary Characteristics for Rural Multilane Highway Segments.

Median Type	Through Lanes	Total Segments	Segment Length, mi	Seg. Length Range, mi		Volume Range, veh/d	
				Minimum	Maximum	Minimum	Maximum
Undivided (U)	4	181	158.6	0.10	6.6	980	20,200
Nonrestrictive (N)	4	92	48.3	0.12	3.5	1,120	39,300
Restrictive (R)	4	289	365.0	0.11	6.5	1,550	39,200
Overall:		562	571.9	0.10	6.6	980	39,300

The barrier and driveway characteristics for rural multilane highway segments are listed in Table 4-8. This table also lists the average inside and outside clearance distances represented in the database. The clearance distances for segments with a restrictive median were the most generous at about 54 ft. The segments were found to have 2.2 to 5.5 percent of their length with an outside barrier, depending on median type.

Table 4-8. Barrier and Driveway Characteristics for Rural Multilane Highway Segments.

Median Type	Through Lanes	Lateral Clearance		Proportion Barrier		Number of Driveways			
		Inside [1]	Outside [2]	Inside [3]	Outside [3]	U-R [4]	Industrial	Business	Office
Undivided	4	n.a.	33.2	n.a.	0.033	468	270	246	18
Nonrestrictive	4	15.1	31.0	n.a.	0.022	207	177	461	11
Restrictive	4	54.2	54.8	0.033	0.055	250.0	213.5	134.5	15.5
Overall:		34.7	39.7	0.033	0.037	925	660	842	45

Notes:
1 - For restrictive medians, inside clearance equals the median width less the width of the two inside shoulders. For nonrestrictive medians, inside clearance equals the median width.
2 - Outside clearance represents the distance between the outside shoulder and the nearest continuous obstruction.
3 - Proportion of segment length with rigid barrier. Inside barrier is in the median. Outside barrier is on the roadside.
4 - U-R.: undeveloped or single-family residential property.
n.a.: not applicable.

The number of driveways associated with each cross section combination is also shown in Table 4-8. Driveways serving the residential/undeveloped category are most common. Driveway counts to the nearest 0.5 driveway are shown for the restrictive median types because this median type was frequently associated with "partial" driveways (as defined in a previous section). All total, 2472 driveways were found on the 562 segments. This count translates into a driveway density of 4.3 driveways/mile.

About 85 percent of the segment length in the database is adjacent to single-family-residential or undeveloped property. The commercial business land use is the second most commonly found land use; it averages about 8 percent of the land use. Industrial and office land uses comprise about 6 and 1 percent, respectively, of all adjacent land use. These percentages did not vary significantly among the various median types.

Data Collection

Crash data were identified for each segment using the TxDOT crash database. Three years of crash data, corresponding to years 1999, 2000, and 2001, were identified for each segment. The ADT for each of these three years was obtained from the TRM database and averaged to obtain one ADT for each segment. Crash data prior to 1999 were not used due to the increase in speed limit that occurred on many Texas highways in 1997 and 1998. Crashes that were associated with intersections were excluded from the database.

The crash records associated with each segment were categorized in terms of whether they were multiple-vehicle nondriveway crashes, single-vehicle crashes, or driveway-related crashes. Only crashes that were identified as associated with one or more injuries or fatalities (i.e., K, A, B, C severity rating) were included in the database.

Data Analysis

This section is divided into two subsections. The first subsection summarizes the crash data at the selected study sites. The second subsection describes the formulation of the calibration model and summarizes the statistical analysis methods used to calibrate it.

Database Summary

The crash data for the 562 segments are summarized in Table 4-9. These segments were collectively associated with 1376 injury or fatal crashes. The trends in crash rate are provided in the last column of the table. Table 4-1 indicates that the injury (plus fatal) crash rate for four-lane highway segments is 0.30 and 0.21 crashes/mvm for undivided and depressed (i.e., restrictive) median types, respectively. These rates are consistent with the injury (plus fatal) crash rates in Table 4-9.

Table 4-9. Crash Data Summary for Rural Multilane Highway Segments.

Median Type	Through Lanes	Exposure,[1] mvm	Injury + Fatal Crashes / 3 years				Crash Rate, cr/mvm
			Multiple-Vehicle[2]	Single-Vehicle	Driveway-Related	Total	
Undivided	4	434.5	125	208	63	396	0.30
Nonrestrictive	4	196.8	74	58	87	219	0.37
Restrictive	4	1245.3	186	534	41	761	0.20
	Overall:	1876.6	385	800	191	1376	0.24

Notes:
1 - mvm: million vehicle-miles.
2 - Multiple-vehicle crashes do not include driveway-related crashes.

Model Development and Statistical Analysis Methods

This section describes the proposed rural multilane highway segment safety prediction model and the methods used to calibrate it. The form of this model is:

$$E[N] = E[N]_b \times AMF_{lw} \times AMF_{osw} \times AMF_{isw} \times AMF_{cr|agg} \times AMF_{ic|agg} \times AMF_{oc|agg} \times AMF_{tk} \quad \textbf{(14)}$$

with,

$$E[N]_b = E[N]_{mv} I_{mv} + E[N]_{sv} I_{sv} + E[N]_{dw} I_{dw} \quad \textbf{(15)}$$

$$E[N]_{mv} = E[N]_{mv4u} I_{4u} + E[N]_{mv4n} I_{4n} + E[N]_{mv4r} I_{4r} \quad \textbf{(16)}$$

$$E[N]_{mv4u} = e^{b_{mv4u} + b_{mvu} Ln(ADT/1000) + Ln(L)} \quad \textbf{(17)}$$

$$E[N]_{mv4n} = e^{b_{mv4n} + b_{mvn} Ln(ADT/1000) + Ln(L)} \quad \textbf{(18)}$$

$$E[N]_{mv4r} = e^{b_{mv4r} + b_{mvr} Ln(ADT/1000) + Ln(L)} \quad \textbf{(19)}$$

$$E[N]_{sv} = E[N]_{sv4u} I_{4u} + E[N]_{sv4n} I_{4n} + E[N]_{sv4r} I_{4r} \quad \textbf{(20)}$$

$$E[N]_{sv4u} = e^{b_{sv4u} + b_{svu} Ln(ADT/1000) + Ln(L)} \quad \textbf{(21)}$$

$$E[N]_{sv4n} = e^{b_{sv4n} + b_{svn} Ln(ADT/1000) + Ln(L)} \quad \textbf{(22)}$$

$$E[N]_{sv4r} = e^{b_{sv4r} + b_{svr} Ln(ADT/1000) + Ln(L)} \quad \textbf{(23)}$$

$$E[N]_{dw} = E[N]_{dw4u} I_{4u} + E[N]_{dw4n} I_{4n} + E[N]_{dw4r} I_{4r} \quad (24)$$

$$E[N]_{dw4u} = (n_{res} + n_{ind} e^{b_{dwi}} + n_{bus} e^{b_{dwb}} + n_{off} e^{b_{dwo}}) e^{b_{dw4u} + b_{dwu} Ln(ADT/15000)} \quad (25)$$

$$E[N]_{dw4n} = (n_{res} + n_{ind} e^{b_{dwi}} + n_{bus} e^{b_{dwb}} + n_{off} e^{b_{dwo}}) e^{b_{dw4n} + b_{dwn} Ln(ADT/15000)} \quad (26)$$

$$E[N]_{dw4r} = (n_{res} + n_{ind} e^{b_{dwi}} + n_{bus} e^{b_{dwb}} + n_{off} e^{b_{dwo}}) e^{b_{dw4r} + b_{dwr} Ln(ADT/15000)} \quad (27)$$

$$AMF_{ic|agg} = (1.0 - P_{ib})\ AMF_{ic|is} + P_{ib}\ AMF_{ic|ib}\ AMF_{b|ib} \quad (28)$$

$$AMF_{oc|agg} = (1.0 - P_{ob})\ AMF_{oc|os} + P_{ob}\ AMF_{oc|ob}\ AMF_{b|ob} \quad (29)$$

$$AMF_{tk} = e^{b_{tk}(P_t - 16)} \quad (30)$$

with,

For raised-curb and positive barrier medians:

$$AMF_{ic|is} = e^{b_{icr}([W_m - 2W_{is}]^{0.5} - 8^{0.5})} \quad (31)$$

$$AMF_{ic|ib} = e^{b_{icr}([W_m - 2W_{is} - W_{ib}]^{0.5} - 8^{0.5})} \quad (32)$$

For depressed medians:

$$AMF_{ic|is} = e^{b_{icr}([W_m - 2W_{is}]^{0.5} - 68^{0.5})} \quad (33)$$

$$AMF_{ic|ib} = e^{b_{icr}([W_m - 2W_{is} - W_{ib}]^{0.5} - 68^{0.5})} \quad (34)$$

For nonrestrictive medians (with $P_{ib} = 0.0$):

$$AMF_{ic|is} = e^{b_{nmw}(W_m - 16)} \quad (35)$$

For undivided medians (with $P_{ib} = 0.0$):

$$AMF_{ic|is} = 1.0 \quad (36)$$

For all median types:

$$AMF_{b|ib} = e^{2\,b_{bar}/(W_m - 2W_{is} - W_{ib})} \quad (37)$$

$$AMF_{b|ob} = e^{b_{bar}/W_{ocb}} \quad (38)$$

$$AMF_{oc|os} = (e^{-0.014([W_{hc} - W_s] - 22)} - 1.0) P_{hc} + 1.0 \quad \geq 1.0 \quad (39)$$

$$AMF_{oc|ob} = (e^{-0.014(W_{ocb} - 22)} - 1.0) P_{hc} + 1.0 \quad \geq 1.0 \quad (40)$$

where,

AMF_{lw} = lane width accident modification factor;
AMF_{osw} = outside shoulder width accident modification factor;
AMF_{isw} = inside shoulder width accident modification factor;
$AMF_{cr|agg}$ = aggregated curve radius accident modification factor (Equation 7);
$AMF_{ic|agg}$ = aggregated inside clearance accident modification factor;
$AMF_{oc|agg}$ = aggregated outside clearance accident modification factor;
AMF_{tk} = truck presence accident modification factor;
I_{mv} = crash indicator variable (= 1.0 if multiple-vehicle nondriveway crash data, 0.0 otherwise);
I_{sv} = crash indicator variable (= 1.0 if single-vehicle crash data, 0.0 otherwise);
I_{dw} = crash indicator variable (= 1.0 if driveway-related crash data, 0.0 otherwise);
$E[N]_{mv4Y}$ = expected multiple-vehicle nondriveway crash frequency for four lanes and median type Y, (Y = u: undivided, n: nonrestrictive, r: restrictive); crashes/yr;
I_{4Y} = cross section indicator variable (= 1.0 if cross section has four lanes and median type Y, 0.0 otherwise);
$E[N]_{sv4Y}$ = expected single-vehicle crash frequency for four lanes and median type Y; crashes/yr;
$E[N]_{dw4Y}$ = expected driveway-related crash frequency for four lanes and median type Y; crashes/yr;
n_{res} = number of driveways serving residential land uses;
n_{ind} = number of driveways serving industrial land uses;
n_{bus} = number of driveways serving business land uses;
n_{off} = number of driveways serving office land uses;
P_t = percent trucks represented in ADT, %;
P_{ib} = proportion of segment length with a barrier present in the median (= 0.5 L_{ib} /L);
L_{ib} = length of inside lane paralleled by a barrier in median, total of both travel directions, mi;
P_{ob} = proportion of segment length with a barrier present on the roadside (i.e., outside) (= 0.5 L_{ob} /L);
L_{ob} = length of outside lane paralleled by a barrier on roadside, total of both travel directions, mi;
P_{hc} = proportion of single-vehicle run-off-road crashes;
$AMF_{ic|is}$ = inside clearance accident modification factor based on distance between inside shoulders (applies to segment portion *without* a barrier in median);
$AMF_{ic|ib}$ = inside clearance accident modification factor based on clearance distance to inside barrier (applies to segment portion *with* a barrier in median);
$AMF_{b|ib}$ = rigid barrier accident modification factor based on clearance distance to inside barrier;
$AMF_{oc|os}$ = outside clearance accident modification factor based on distance from outside shoulder to nearest continuous obstruction (applies to segment portion *without* a roadside barrier);
$AMF_{oc|ob}$ = outside clearance accident modification factor based on clearance distance to outside barrier (applies to segment portion *with* a roadside barrier);
$AMF_{b|ob}$ = rigid barrier accident modification factor based on clearance distance to outside barrier;
W_m = median width (measured from near edges of traveled way in both travel directions), ft;
W_{is} = inside shoulder width, ft;
W_s = outside shoulder width, ft;

W_{ib} = inside barrier width (measured from barrier face to barrier face), ft;
W_{hc} = horizontal clearance, ft;
W_{ocb} = outside clearance distance when a barrier is present on the roadside, ft; and
b_i = calibration coefficient for condition i (see Table 4-10).

It is assumed that any inside barrier (i.e., rigid barrier in the median) is centered in the median. Inside barrier width W_{ib} is measured from the outer face of the barrier nearest to one travel direction to the outer face of the barrier nearest to the other travel direction. This width should be no less than the width of a single barrier. Also, it must not exceed the inside clearance distance (i.e., W_m - 2 W_{is}). Design guidelines typically require a 2 ft offset between the face of the barrier and the near edge of shoulder. Therefore, as a practical matter, the inside barrier width should not exceed the inside clearance distance less 4 ft (such that W_m - 2 W_{is} - $W_{ib} \geq 4.0$).

The value of "8" in Equations 31 and 32 represents the base inside clearance for raised-curb or positive barrier medians. It is computed using a base median width of 16 ft and base inside shoulder width of 4 ft (8 = 16 - [2 × 4]). The value of "68" in Equations 33 and 34 represents the base inside clearance for the depressed median type. It is computed using a base median width of 76 ft and base inside shoulder width of 4 ft (68 = 76 - [2 × 4]). The value of "16" in Equation 35 represents the base median width for nonrestrictive medians on rural multilane highways.

Horizontal clearance is measured from the outside edge of the traveled way to the nearest continuous obstruction on the roadside. It is measured for both travel directions. Two values of horizontal clearance are obtained for a segment. These two values are then averaged to obtain the horizontal clearance W_{hc} used in the equations above. Given recognized variations in horizontal clearance along most segments, the computed value should be determined such that it is representative of the entire length of the segment. By definition, horizontal clearance W_{hc} must equal or exceed the width of the outside shoulder.

Equations 39 and 40 replicate the horizontal clearance AMF in the *Workbook* (*1*). The value of "22" in these equations represents the base outside clearance distance. It is computed as the difference between the base horizontal clearance and the base outside shoulder width (i.e., 22 = 30 - 8). The horizontal clearance AMF is based on horizontal clearances ranging from 0 to 30 ft. In recognition of this restriction, the AMF value is limited to 1.0 for horizontal clearances that exceed 30 ft (i.e., an outside clearance distance of 22 ft).

Inside and outside clearance distances were used in the model instead of median width and horizontal clearance. This approach was followed because the two "clearance distance" variables do not include shoulder width in their dimension. AMFs for inside and outside shoulder width are already included in the model. Thus, the direct use of median width or horizontal clearance AMFs would "double count" the effect of a change in shoulder width.

The mathematical relationships for each of the first three AMFs identified in the variable list are provided in the *Workbook*. A preliminary examination of the data indicated that these AMFs were accurate estimators of the relationship between lane width, outside shoulder width, or inside shoulder width and multiple-vehicle, single-vehicle, or driveway-related crashes. For this calibration

activity, the proportions in Table 4-5 were used with the AMFs from the *Workbook* (as opposed to the proportions provided in the *Workbook*).

The mathematical relationships for the median width AMF and the truck presence AMF identified in the *Workbook* were not used in the proposed safety prediction model. A preliminary examination of the data indicated that these AMFs did not accurately replicate the trends in the crash data. For this reason, it was determined that these AMFs should be calibrated using the assembled database. An examination of their effect on multiple-vehicle, single-vehicle, and driveway-related crashes did not reveal a compelling reason to develop separate AMFs for each crash type.

The utility pole density AMF identified in the *Workbook* is not included in the proposed model because pole count and location data were not readily available. It is assumed that the collective set of segments in the database is represented by the base condition of a 30 ft pole offset and 25 poles/mi. As such, the model should not be biased as a result of excluding this AMF. A similar situation was present for grade, superelevation rate, rumble strip, and side slope data. It was assumed that base conditions for these variables were represented in the collective set of segments.

The driveway density AMF and the TWLTL median type AMF identified in the *Workbook* are not included in the proposed model because their correlation with safety is represented by alternative model components. Specifically, driveway crashes and TWLTL median type are explicitly represented by the component models of the proposed safety prediction model.

Preliminary versions of the model included variables to quantify the relationship between median opening density and crash frequency. However, their correlation was negligible and the variables were removed from the model.

The Nonlinear Regression procedure (NLIN) in the SAS software was used to estimate the proposed model coefficients (7). This procedure was used because the proposed safety prediction model is both nonlinear and discontinuous. The loss function associated with NLIN was specified to equal the log likelihood function for the negative binomial distribution. Equation 12 was used to define the variance function for the multiple-vehicle and single-vehicle crashes. Equation 12, with the segment length term L removed, was used to define the variance function for the driveway-related crashes. The procedure was set up to estimate model coefficients based on maximum-likelihood methods.

One disadvantage of the NLIN procedure is that it is not able to compute the best-fit value of k for the calibrated prediction model. This disadvantage is overcome by using the Generalized Modeling (GENMOD) procedure in SAS. GENMOD automates the over-dispersion factor estimation process using maximum-likelihood methods. Initially, NLIN is used to calibrate the prediction model. Then, GENMOD is used to regress the relationship between the reported and predicted crash frequencies (where the natural log of the predicted values is specified as an offset variable, and the log link function is used) to obtain a best-fit over-dispersion estimate a using the internal variance function ($V[X] = E[N] + a\,E[N]^2$). The estimate a from GENMOD is then multiplied by the average segment length and inverted to estimate a new value of k (i.e., $k_{new} = 1/[a\,L.]$). This new value is then used in a second application of NLIN and the process is repeated

until convergence is achieved between the *k* value used in NLIN and that obtained from GENMOD. Convergence is typically achieved in two iterations. As a last step, the value of *k* obtained from this iterative method was used in Equation 13 to estimate the true over-dispersion parameter.

Model Calibration

The safety prediction model calibration process consisted of the simultaneous calibration of all three component models (i.e., multiple-vehicle nondriveway, single-vehicle, and driveway-related models) and AMFs using the aggregate model represented by Equation 14. The simultaneous calibration approach was needed because the AMFs in Equation 14 were common to all three component models. The database assembled for this process included three replications of the original database. However, the dependent variable in the first replication was set equal to the multiple-vehicle crashes. The dependent variable in the second replication was set equal to the single-vehicle crashes and that for the third replication was set equal to the driveway-related crashes.

The process of replicating the original database resulted in the variability in the data being underestimated because the random error in each set of three observations per site is correlated. The NLIN procedure does not directly accommodate the analysis of repeated measures but the GENMOD procedure does support this analysis. Thus, to estimate the true variability in each regression coefficient estimate from NLIN, the GENMOD procedure was used to regress the relationship between the reported and predicted crash frequencies. GENMOD was instructed to use generalized estimating equations to estimate the standard deviation of the intercept term accounting for the repeated measures. The ratio of this standard deviation to that from a similar model that assumed independent errors was found to be 1.03. The standard deviation for each regression coefficient obtained from NLIN was multiplied by this ratio to estimate its true standard deviation.

For each segment, the predicted crash frequency from each component model was totaled and compared with the total reported crash frequency. The difference between the two totals was then summed for all segments. This sum was found to be very small (i.e., less than 0.5 percent of the total reported crash frequency), so it was concluded that there was no bias in the component models in terms of their ability to predict total crash frequency.

The calibration coefficients for the aggregate safety prediction model are described in the next subsection. The subsequent three subsections collectively describe each of the three component models and their fit to the data. The fit statistics were separately computed using the calibrated component model and an analysis of its residuals.

Aggregate Model

The results of the aggregated safety prediction model calibration are presented in Table 4-10. Calibration of this model focused on injury (plus fatal) crash frequency. The Pearson χ^2 statistic for the model is 1447, and the degrees of freedom are 1661 ($= n - p = [3 \times 562] - 25$). As this statistic is less than $\chi^2_{0.05, 1661}$ (= 1757), the hypothesis that the model fits the data cannot be rejected.

Table 4-10. Rural Multilane Highway Model Statistical Description–Aggregate Model.

Model Statistics		Value		
Scale Parameter φ:		0.87		
Pearson χ^2:		1447 ($\chi^2_{0.05,\,1661}$ = 1757)		
Observations n_o:		562 segments (1376 injury + fatal crashes in 3 years)		
Calibrated Coefficient Values				
Variable	**Inferred Effect of...[1]**	**Value**	**Std. Dev.**	**t-statistic**
b_{tk}	Truck presence	-0.0057	0.0044	-1.3
b_{icr}	Inside clearance distance	-0.0296	0.0293	-1.0
b_{nmw}	Nonrestrictive median width	-0.0253	0.0373	-0.7
b_{bar}	Rigid barrier presence	0.7574	0.3661	2.1
b_{mv4u}	4 lane, undivided on m.v. crashes	-4.894	0.563	-8.7
b_{mv4n}	4 lane, nonrestrictive median on m.v. crashes	-5.246	0.902	-5.8
b_{mv4r}	4 lane, restrictive median on m.v. crashes	-5.204	0.389	-13.4
b_{mvu}	ADT for undivided on m.v. crashes	1.631	0.251	6.5
b_{mvn}	ADT for nonrestrictive median on m.v. crashes	1.802	0.342	5.3
b_{mvr}	ADT for restrictive median on m.v. crashes	1.486	0.151	9.8
b_{sv4u}	4 lane, undivided on s.v. crashes	-2.218	0.355	-6.2
b_{sv4n}	4 lane, nonrestrictive median on s.v. crashes	-2.556	0.760	-3.4
b_{sv4r}	4 lane, restrictive median on s.v. crashes	-2.241	0.227	-9.9
b_{svu}	ADT for undivided on s.v. crashes	0.631	0.171	3.7
b_{svn}	ADT for nonrestrictive median on s.v. crashes	0.667	0.306	2.2
b_{svr}	ADT for restrictive median on s.v. crashes	0.707	0.095	7.4
b_{dw4u}	4 lane, undivided on dw. crashes in U-R	-4.080	0.437	-9.3
b_{dw4n}	4 lane, nonrestrictive median on dw. crashes in U-R	-4.073	0.456	-8.9
b_{dw4r}	4 lane, restrictive median on dw. crashes in U-R	-4.184	0.420	-10.0
b_{dwu}	ADT for undivided on dw. crashes in U-R	0.738	0.328	2.3
b_{dwn}	ADT effect for nonrestrictive median on dw. crashes	1.443	0.377	3.8
b_{dwr}	ADT for restrictive median on dw. crashes	1.038	0.357	2.9
b_{dwi}	Added effect of industrial driveway on dw. crashes	0.985	0.507	1.9
b_{dwb}	Added effect of business driveway on dw. crashes	0.847	0.500	1.7
b_{dwo}	Added effect of office driveway on dw. crashes	2.278	0.588	3.9

Note:
1 - m.v.: multiple-vehicle nondriveway; s.v.: single-vehicle; dw.: driveway-related; U-R: undeveloped or single-family residential.

The t-statistics listed in the last column of Table 4-10 indicate a test of the hypothesis that the coefficient value is equal to 0.0. Those t-statistics with an absolute value that is larger than 2.0 indicate that the hypothesis can be rejected with the probability of error in this conclusion being less than 0.05. For those variables where the absolute value of the t-statistic is smaller than 2.0, it was decided that the variable was essential to the model and was rationalized as being reasonable in magnitude (even if the specific value was not known with a great deal of certainty as applied to this database). Specifically, the t-statistics for the first three variables listed are less than 2.0. However, as is shown in a subsequent section, the coefficient values yield very similar AMF values to those reported in the literature or calibrated in previous technical memoranda.

Multiple-Vehicle Nondriveway Crashes

The results of the multiple-vehicle safety prediction model calibration are presented in Table 4-11. Calibration of this model focused on injury (plus fatal) crash frequency. The Pearson χ^2 statistic for the model is 518, and the degrees of freedom are 552 (= $n - p$ = 562 – 10). As this statistic is less than $\chi^2_{0.05, 552}$ (= 608), the hypothesis that the model fits the data cannot be rejected; however, it is noted that there is slightly less variability in the data than suggested by Equation 12. The R^2 for the model is 0.44. An alternative measure of model fit that is better suited to the negative binomial distribution is R_k^2, as developed by Miaou (*8*). The R_k^2 for the calibrated model is 0.86. This statistic indicates that about 86 percent of the variability due to systematic sources is explained by the model.

Table 4-11. Prediction Model Statistical Description–Multiple-Vehicle Model.

Model Statistics	Value
R^2 (R_k^2):	0.44 (0.86)
Scale Parameter φ:	0.92
Pearson χ^2:	518 ($\chi^2_{0.05, 552}$ = 608)
Over-Dispersion Parameter k:	3.08 mi^{-1}
Observations n_o:	562 segments (385 injury + fatal crashes in 3 years)
Standard Deviation s_e:	±0.35 crashes/yr

The coefficients in Table 4-10 were combined with Equations 17 to 19 to obtain the calibrated safety prediction models for multiple-vehicle nondriveway crashes. The form of each model is:

$$E[N]_{mv4u} = e^{-4.894 + 1.631\,Ln(ADT/1000) + Ln(L)} \quad \textbf{(41)}$$

$$E[N]_{mv4n} = e^{-5.246 + 1.802\,Ln(ADT/1000) + Ln(L)} \quad \textbf{(42)}$$

$$E[N]_{mv4r} = e^{-5.204 + 1.486\,Ln(ADT/1000) + Ln(L)} \quad \textbf{(43)}$$

The fit of the calibrated models is shown in Figure 4-3. This figure compares the predicted and reported crash frequency in the calibration database. The trend line shown represents a "$y = x$" line. A data point would lie on this line if its predicted and reported crash frequency were equal. The data points shown represent the observed multiple-vehicle nondriveway crash frequency for the segments used to calibrate the corresponding component model.

Each data point shown in Figure 4-3 represents the average predicted and average reported crash frequency for a group of 10 segments. The data were sorted by predicted crash frequency to form groups of segments with similar crash frequency. The purpose of this grouping was to reduce the number of data points shown in the figure and, thereby, to facilitate an examination of trends in the data. The individual segment observations were used for model calibration. In general, the data

shown in the figure indicate that the model provides an unbiased estimate of expected crash frequency for segments experiencing up to five multiple-vehicle crashes in a three-year period.

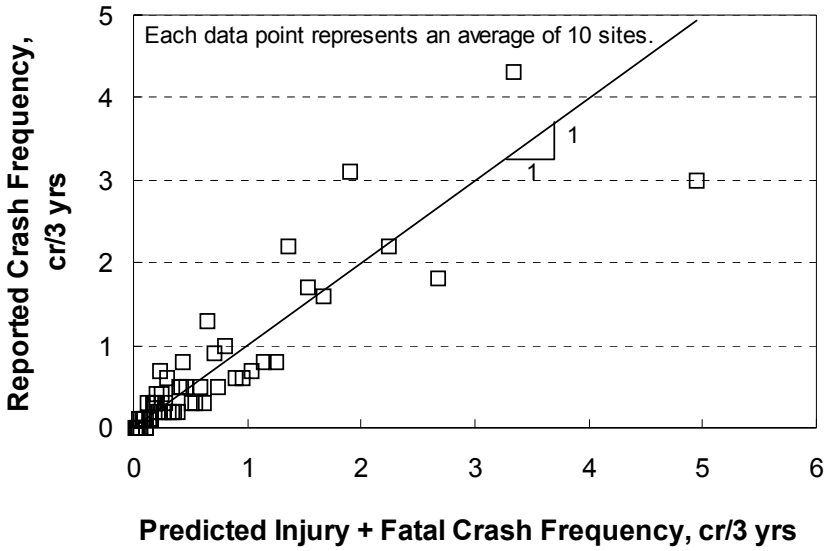


Figure 4-3. Predicted vs. Reported Multiple-Vehicle Crashes.

Single-Vehicle Crashes

The results of the single-vehicle safety prediction model calibration are presented in Table 4-12. Calibration of this model also focused on injury (plus fatal) crash frequency. The Pearson χ^2 statistic for the model is 532, and the degrees of freedom are 552 (= $n - p$ = 652 – 10). As this statistic is less than $\chi^2_{0.05, 552}$ (= 608), the hypothesis that the model fits the data cannot be rejected. The R^2 for the model is 0.56, and the R_k^2 for the calibrated model is 0.87. This statistic indicates that about 87 percent of the variability due to systematic sources is explained by the model.

Table 4-12. Prediction Model Statistical Description–Single-Vehicle Model.

Model Statistics	Value
R^2 (R_k^2):	0.56 (0.87)
Scale Parameter φ:	0.95
Pearson χ^2:	532 ($\chi^2_{0.05, 552}$ = 608)
Over-Dispersion Parameter k:	4.30 mi^{-1}
Observations n_o:	562 segments (800 injury + fatal crashes in 3 years)
Standard Deviation s_e:	±0.49 crashes/yr

The coefficients in Table 4-10 were combined with Equations 21 to 23 to obtain the calibrated safety prediction models for single-vehicle crashes. The form of each model is:

$$E[N]_{sv4u} = e^{-2.218 + 0.631\,Ln(ADT/1000) + Ln(L)} \quad \textbf{(44)}$$

$$E[N]_{sv4n} = e^{-2.556 + 0.667\,Ln(ADT/1000) + Ln(L)} \quad \textbf{(45)}$$

$$E[N]_{sv4r} = e^{-2.241 + 0.707\,Ln(ADT/1000) + Ln(L)} \quad \textbf{(46)}$$

The fit of the calibrated models is shown in Figure 4-4. This figure compares the predicted and reported crash frequency in the calibration database. Each data point shown represents the average predicted and average reported crash frequency for a group of 10 segments. In general, the data shown in the figure indicate that the model provides an unbiased estimate of expected crash frequency for segments experiencing up to seven single-vehicle crashes in a three-year period.

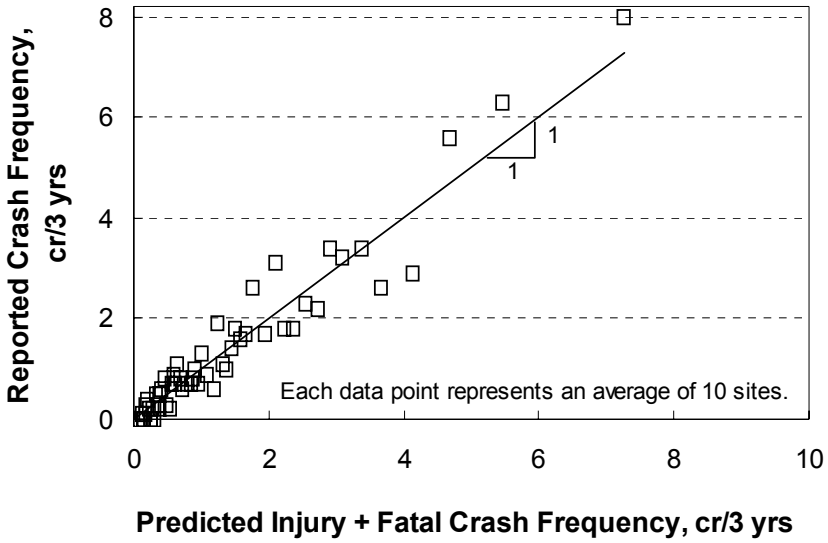


Figure 4-4. Predicted vs. Reported Single-Vehicle Crashes.

Driveway-Related Crashes

The results of the driveway-related safety prediction model calibration are presented in Table 4-13. Calibration of this model also focused on injury (plus fatal) crash frequency. The Pearson χ^2 statistic for the model is 397, and the degrees of freedom are 549 (= $n - p$ = 562 – 13). As this statistic is less than $\chi^2_{0.05,\,549}$ (= 605), the hypothesis that the model fits the data cannot be rejected. The R^2 for the model is 0.47, and the R_k^2 for the calibrated model is 0.87. This statistic indicates that about 87 percent of the variability due to systematic sources is explained by the model.

Table 4-13. Prediction Model Statistical Description–Driveway-Related Model.

Model Statistics	Value
R^2 (R_k^2):	0.47 (0.87)
Scale Parameter φ:	0.71
Pearson χ^2:	397 ($\chi^2_{0.05,549}$ = 605)
Over-Dispersion Parameter k:	1.11
Observations n_o:	562 segments (191 injury + fatal crashes in 3 years)
Standard Deviation s_e:	±0.29 crashes/yr

The coefficients in Table 4-10 were combined with Equations 25 to 27 to obtain the calibrated safety prediction models for driveway-related crashes. The form of each model is:

$$E[N]_{dw4u} = (n_{res} + n_{ind}\, e^{0.985} + n_{bus}\, e^{0.847} + n_{off}\, e^{2.278})\, e^{-4.080 + 0.738\, Ln(ADT/15000)} \quad \textbf{(47)}$$

$$E[N]_{dw4n} = (n_{res} + n_{ind}\, e^{0.985} + n_{bus}\, e^{0.847} + n_{off}\, e^{2.278})\, e^{-4.073 + 1.443\, Ln(ADT/15000)} \quad \textbf{(48)}$$

$$E[N]_{dw4r} = (n_{res} + n_{ind}\, e^{0.985} + n_{bus}\, e^{0.847} + n_{off}\, e^{2.278})\, e^{-4.184 + 1.038\, Ln(ADT/15000)} \quad \textbf{(49)}$$

The fit of the calibrated models is shown in Figure 4-5. This figure compares the predicted and reported crash frequency in the calibration database. Each data point shown represents the average predicted and average reported crash frequency for a group of 10 segments. In general, the data shown in the figure indicate that the model provides an unbiased estimate of expected crash frequency for segments experiencing up to two driveway-related crashes in a three-year period.

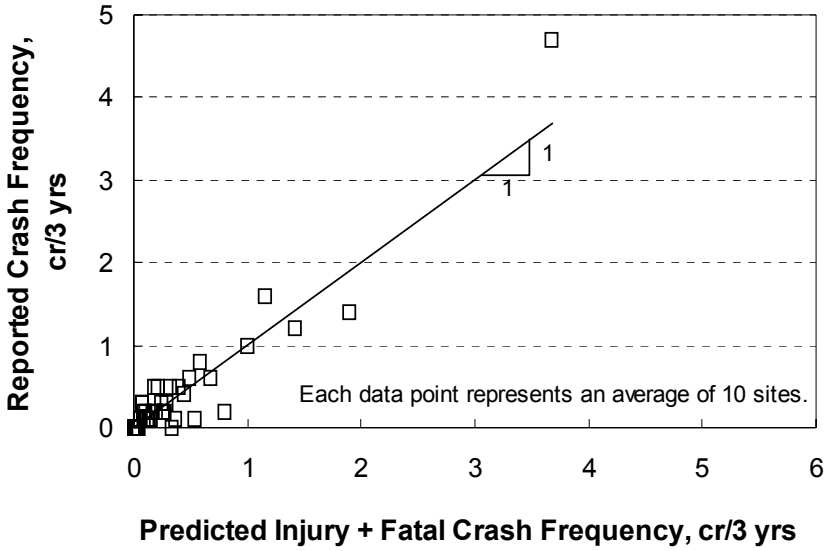


Figure 4-5. Predicted vs. Reported Driveway-Related Crashes.

Calibrated AMFs

Several AMFs were calibrated in conjunction with the safety prediction model calibration. All of them were calibrated using injury (plus fatal) crash data. Collectively, they describe the relationship between barrier presence, inside clearance distance (for restrictive medians), outside clearance distance, nonrestrictive median width, and truck presence on crash frequency.

Rigid Barrier AMF. The calibrated rigid barrier AMF has two forms, depending on whether it is used to evaluate the presence of an inside or outside barrier. The AMF for an inside barrier is:

$$AMF_{b|ib} = e^{2 \times 0.757/(W_m - 2W_{is} - W_{ib})} \quad \textbf{(50)}$$

The AMF for an outside barrier is:

$$AMF_{b|ob} = e^{0.757/W_{ocb}} \quad \textbf{(51)}$$

The two equations above are effectively the same. The width term in the first equation (i.e., $[W_m - 2\ W_{is} - W_{ib}]/2$) equates to the inside clearance distance. The width term in the second equation explicitly represents the outside clearance distance.

The rigid barrier AMF is shown in Figure 4-6. It is limited to clearance distances of 2 ft or more. This AMF is used with the aggregated inside clearance AMF in Equation 28 or the aggregated outside clearance AMF in Equation 29 to estimate the combined effect of clearance distance and barrier presence. The trends in Figure 4-6 indicate that the presence of a barrier within 2 to 4 ft of the shoulder tends to be associated with a 20 to 45 percent increase in crash frequency. In contrast, the presence of a barrier 30 ft from the shoulder is associated with only a 2 or 3 percent increase.

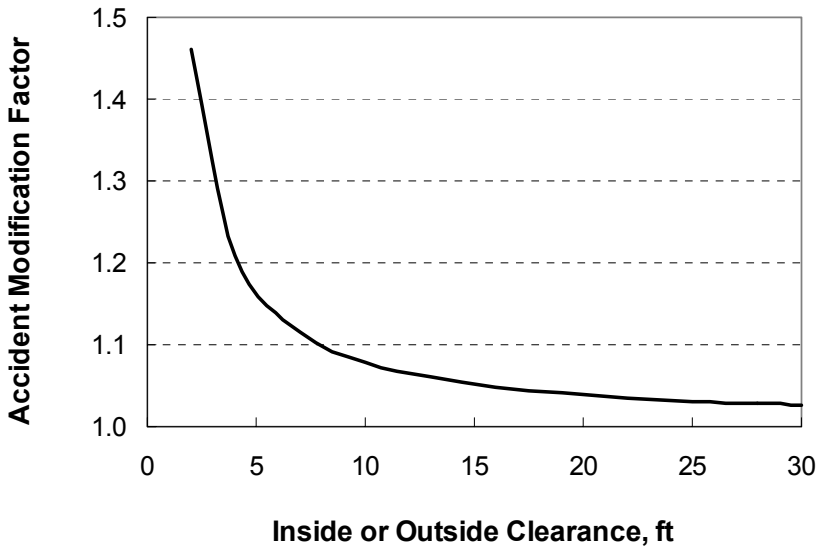


Figure 4-6. Calibrated Rigid Barrier AMF.

Inside Clearance AMF. This AMF was separately calibrated for use with restrictive medians. An AMF for nonrestrictive median width is described in the next subsection. The calibrated inside clearance AMF has two forms, depending on whether it is used to evaluate inside clearance in the presence of a rigid barrier. The AMF when an inside barrier is not present is:

$$AMF_{ic|is} = e^{-0.0296\,([W_m - 2W_{is}]^{0.5} - W_b^{0.5})} \quad \textbf{(52)}$$

The AMF when an inside barrier is present is:

$$AMF_{ic|ib} = e^{-0.0296\,([W_m - 2W_{is} - W_{ib}]^{0.5} - W_b^{0.5})} \quad \textbf{(53)}$$

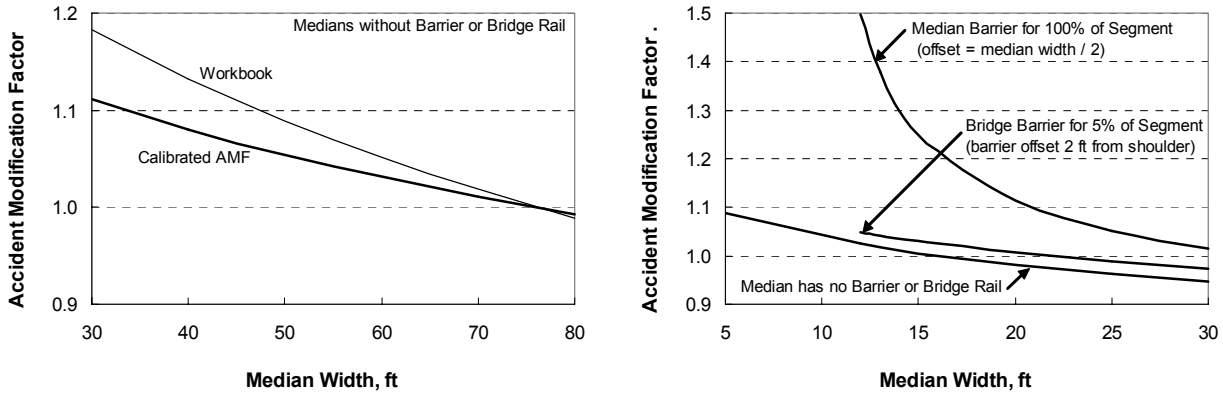
where,

W_b = base clearance distance (= 8 ft for raised-curb or positive barrier; 68 ft for depressed), ft.

When these two equations are used to evaluate raised-curb or positive barrier medians, the base width W_b is equal to 8 ft (= 16 ft median − 2 × 4 ft inside shoulder). When they are used to evaluate depressed medians, the base width W_b is equal to 68 ft (= 76 ft median − [2 × 4 ft inside shoulder]).

Equations 52 and 53 are used with Equation 28 to estimate the aggregated inside clearance AMF for segments that have a mixture of barriers and no barriers in the median area.

The calibrated inside clearance AMF when an inside barrier is not present (i.e., Equation 52) is shown in Figure 4-7a. A base clearance distance of 68 ft is used for this evaluation (it corresponds to a base median width of 76 ft). Also shown is the median width AMF from the *Workbook* (*1*). The trends in Figure 4-7a indicate that wider medians tend to be associated with a reduction in crash frequency.



a. No Rigid Barrier Present. **b. Aggregate AMF.**

Figure 4-7. Calibrated Inside Clearance AMF.

Figure 4-7b illustrates how Equation 28 can be used to combine the inside clearance and rigid barrier AMFs to the evaluation of a segment with a mixture of barriers and no barriers. The lower trend line in Figure 4-7b represents a highway with no median barrier and is used for reference only. A second trend line in Figure 4-7b applies when a bridge-related, inside barrier is present for 5 percent of the segment length. This barrier is offset 2 ft from the edge of the inside shoulder, regardless of the width of the median. The bridge barrier is shown to be associated with a 3 percent increase in crash frequency on the subject segment (relative to the "no barrier" condition).

The upper trend line in Figure 4-7b illustrates the effect of a continuous rigid barrier in the middle of the median. The effect is shown to be highly sensitive to the width of the median (and thus, the distance between the barrier and the traveled way). Figure 4-7b suggests that the installation of a median barrier in a 30 ft depressed median is likely to be associated with a 6 percent increase in crashes; however, it will also significantly reduce the number of fatal crashes (*9*). Crash data reported by Miaou et al. (*9*) indicate that median barriers increase injury (plus fatal) crashes by about 16 percent, but reduce fatal crashes by over 70 percent. The use of Equation 28 to compare medians "with" and "without" a barrier (for the same average median widths as reported by Miaou et al.) indicates a slightly more conservative crash increase of 9 percent associated with the installation of a median barrier.

Outside Clearance AMF. The calibrated outside clearance AMF has two forms, depending on whether it is used to evaluate inside clearance in the presence of a rigid barrier. The AMF when an outside barrier is not present is:

$$AMF_{oc|os} = (e^{-0.014\,([W_{hc} - W_s] - 22)} - 1.0)\,P_{hc} + 1.0 \quad \geq 1.0 \qquad \textbf{(54)}$$

The AMF when an outside barrier is present is:

$$AMF_{oc|ob} = (e^{-0.014\,(W_{ocb} - 22)} - 1.0)\,P_{hc} + 1.0 \quad \geq 1.0 \qquad \textbf{(55)}$$

A base clearance distance of 22 ft is used in Equations 54 and 55. It corresponds to a base horizontal clearance of 30 ft and a base outside shoulder width of 8 ft.

Equations 54 and 55 are used with Equation 29 to estimate the aggregated outside clearance AMF for segments that have a mixture of barriers and no barriers on the roadside.

The calibrated outside clearance AMF (when an outside barrier is not present, i.e., Equation 54) is shown as the lower trend line in Figure 4-8. This trend line is identical to that from the horizontal clearance AMF from the *Workbook* (*1*).

The upper trend line in Figure 4-8 illustrates how Equation 29 can be used to combine the outside clearance and rigid barrier AMFs to the evaluation of a segment with a mixture of barriers and no barriers. This trend line applies when a bridge-related, outside barrier is present for 5 percent of the segment length. This barrier is offset 2 ft from the edge of the outside shoulder. The bridge barrier is shown to be associated with a 4 percent increase in crash frequency on the subject segment (relative to the "no barrier" condition).

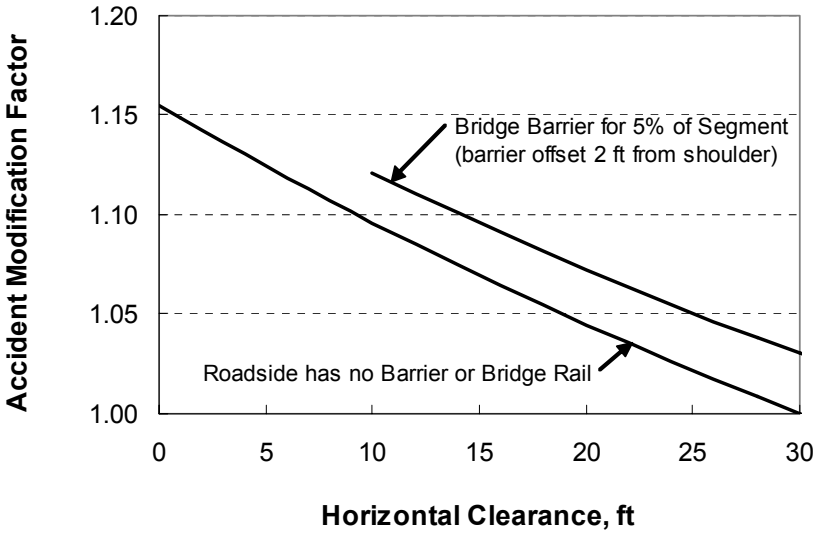


Figure 4-8. Calibrated Outside Clearance AMF.

Nonrestrictive Median Width AMF. The calibrated nonrestrictive median width AMF is:

$$AMF_{ic|is} = AMF_{nmw} = e^{-0.0253\,(W_m - 16)} \quad \textbf{(56)}$$

where,
AMF_{nmw} = nonrestrictive median width accident modification factor;

This AMF was defined in Equation 35 as being an inside clearance AMF to facilitate its calibration in the regression model (via Equation 28). However, in practice, it does not need to be used with Equation 28 because nonrestrictive medians do not have an inside barrier (i.e,. $P_{ib} = 0.0$). When used directly in Equation 14 to evaluate a highway segment, it is referred to as the nonrestrictive median width AMF (AMF_{nmw}).

The nonrestrictive median width AMF is shown in Figure 4-9 using a thick trend line. It is limited to nonrestrictive median widths in the range of 11 to 16 ft. Two additional AMFs are also shown in the figure. One AMF is the inside clearance AMF described in Equation 52. The second AMF is the nonrestrictive median width AMF for urban and suburban arterial street segments, as described in Chapter 3. The two trend lines for nonrestrictive median width AMFs are very similar and suggest the correlation between nonrestrictive median width and crash frequency is effectively the same for urban arterial streets and rural highways.

The nonrestrictive median width AMFs in Figure 4-9 indicate greater sensitivity to median width, relative to that for restrictive medians. This trend is logical given the heavy use of nonrestrictive medians by turning vehicles and the lack of a raised curb (or depressed median) to provide some type of positive separation between opposing traffic streams.

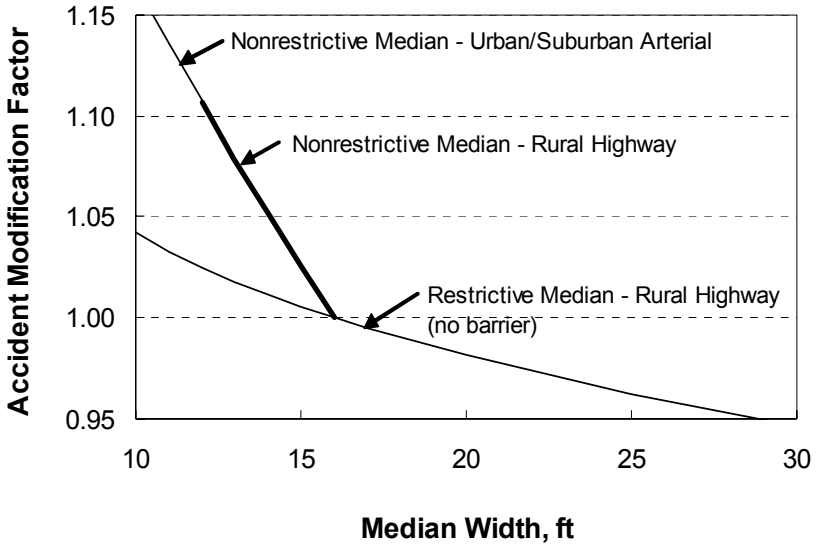


Figure 4-9. Calibrated Nonrestrictive Median Width AMF.

Truck Presence AMF. The calibrated truck presence AMF is:

$$AMF_{tk} = e^{-0.0057\,(P_t - 16)} \quad \textbf{(57)}$$

This AMF is shown in Figure 4-10 using a thick trend line. It is limited to truck percentages in the range of 0 to 25 percent. The truck presence AMF described in Chapter 3 for urban and suburban arterial streets is also shown in the figure and indicates a similar trend.

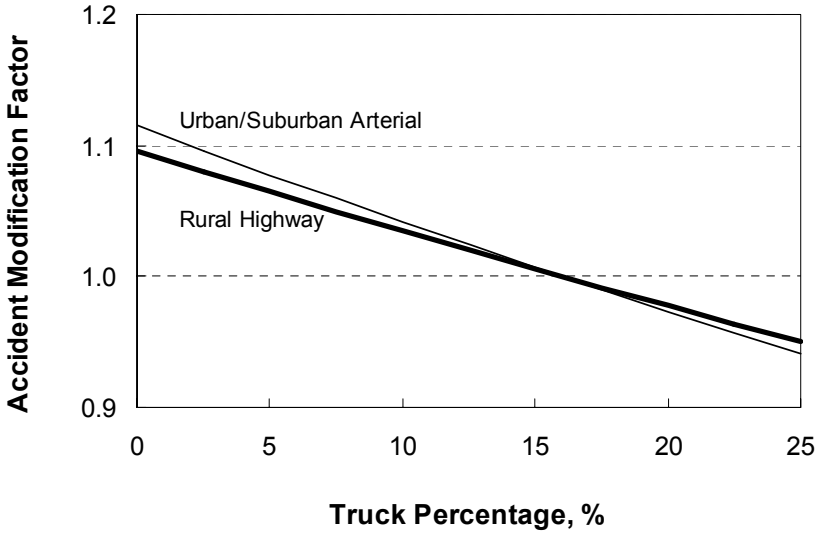


Figure 4-10. Calibrated Truck Presence AMF.

The trend in Figure 4-10 indicates that highways with higher percentages of trucks are associated with fewer crashes. It may suggest that drivers are more cautious when there are many trucks present in the traffic stream. It may also suggest that highways with higher truck percentages are more likely to have design dimensions that are "generous" (relative to the design control limits) and thus provide some additional safety benefit.

Sensitivity Analysis

The relationship between crash frequency and traffic demand, as obtained from calibrated Equation 14, is illustrated in Figure 4-11 for a 1-mile highway segment. The individual component models are illustrated in Figures 4-11a, 4-11b, and 4-11c. Figure 4-11c illustrates the expected crash frequency when 10 residential driveways are present. These driveways were assumed to be "full" driveways for the undivided cross section and nonrestrictive median type. They were assumed to be "partial" driveways for the restrictive median type. The sum of the individual component crash frequencies is illustrated in Figure 4-11d. The trends in this figure are comparable to those in Figures 4-1 and 4-2. The length of the trend lines in Figure 4-11 reflects the range of traffic demand in the data.

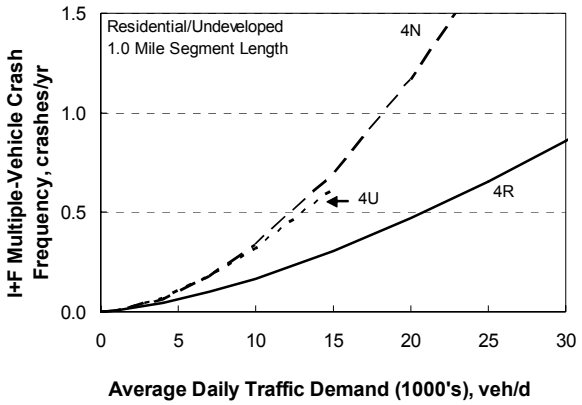
For an ADT of 10,000 veh/d, the trends in Figure 4-11d indicate that a four-lane undivided highway has an expected crash frequency of 1.2 crashes/yr. In contrast, a four-lane divided highway has an expected crash frequency of 0.86 crashes/yr. When compared with the trends in Figure 4-1, the expected crash frequency in Figure 4-11d is between 6 and 12 percent larger than that from Figure 4-1 for any given volume. More generally, the expected crash frequency obtained from Equation 14 is 6 to 12 percent larger than that obtained from the model in the *Workbook* for similar traffic, land use, and geometric conditions.

Model Extensions

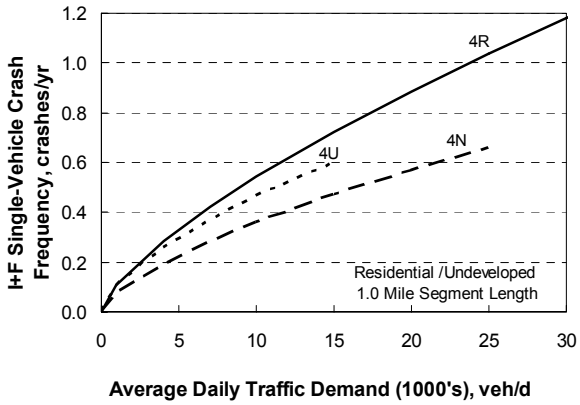
Components of the proposed safety prediction model are calibrated for segments that have a nonrestrictive median type. This median type includes both the TWLTL and the flush-paved median. The majority of the segments in the database have the TWLTL median; however, the safety experience of the two median types is similar, and the calibrated model can be applied to either type with equal accuracy.

The proposed safety prediction model was calibrated using only four-lane segments. Data for six- and eight-lane rural highways is very limited. The safety experience of the six and eight-lane cross sections is not believed to be similar to that of four-lane cross sections. Thus, the calibrated model should *not* be applied to highways with six or more lanes.

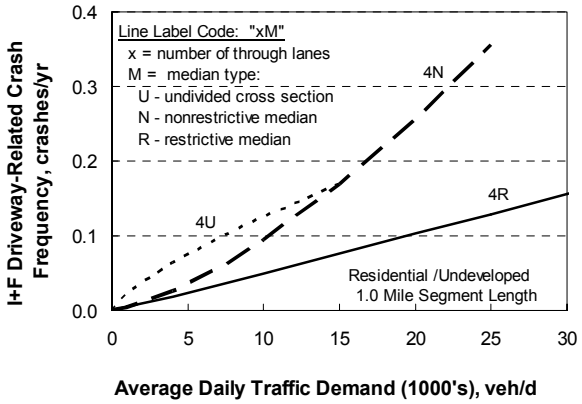
The component model for predicting driveway-related crashes is intended for the evaluation of segments with varying numbers of driveways and associated land uses. This model is not sufficiently refined to be able to estimate the effect of installing or removing a single driveway from a highway segment. If the analyst desires to conduct this type of evaluation, then the driveway should be evaluated as an intersection using the intersection safety prediction models.



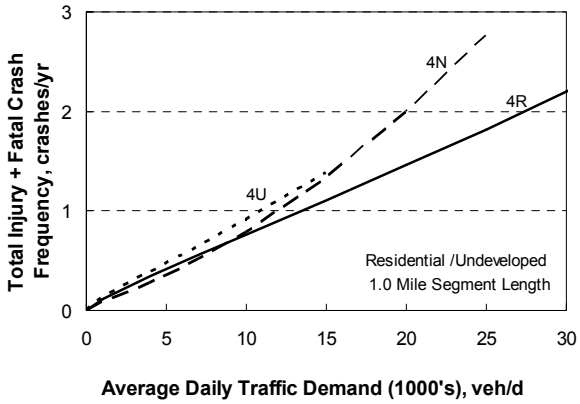
a. Multiple-Vehicle Crashes.



b. Single-Vehicle Crashes.



c. Driveway-Related Crashes.



d. Total Crash Frequency.

Figure 4-11. Proposed Crash Prediction Model.

REFERENCES

1. Bonneson, J.A., K. Zimmerman, and K. Fitzpatrick. *Interim Roadway Safety Design Workbook*. FHWA/TX-06/0-4703-P4, Texas Department of Transportation, Austin, Texas, April 2006.

2. Lord, D., B. Persaud, S. Washington, J. Ivan, I. van Schalkwyk, C. Lyon, T. Jonsson, and S. Geedipally. *Methodology to Predict the Safety Performance of Rural Multilane Highways.* Final Report. NCHRP Project 17-29. National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C., February 2008.

3. Bonneson, J.A., K. Zimmerman, and K. Fitzpatrick. *Roadway Safety Design Synthesis.* FHWA/TX-05/0-4703-P1. Texas Department of Transportation, Austin, Texas, November 2005.

4. Harwood, D., K. Bauer, K. Richard, D. Gilmore, J. Graham, I. Potts, D. Torbic, and E. Hauer. *Methodology to Predict the Safety Performance of Urban and Suburban Arterials.* Final Report. NCHRP Project 17-26. National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C., March 2007.

5. Lord, D. "Modeling Motor Vehicle Crashes Using Poisson-Gamma Models: Examining the Effects of Low Sample Mean Values and Small Sample Size on the Estimation of the Fixed Dispersion Parameter." *Accident Analysis & Prevention*, Vol. 38, No. 4, Elsevier Ltd., Oxford, Great Britain, July 2006, pp. 751-766.

6. Bonneson, J., D. Lord, K. Zimmerman, K. Fitzpatrick, and M. Pratt. *Development of Tools for Evaluating the Safety Implications of Highway Design Decisions.* FHWA/TX-07/0-4703-4. Texas Department of Transportation, Austin, Texas, September 2006.

7. *SAS OnlineDoc, Version 8*, SAS Institute, Inc., Cary, North Carolina, 1999.

8. Miaou, S.P. *Measuring the Goodness-of-Fit of Accident Prediction Models.* FHWA-RD-96-040. Federal Highway Administration, Washington, D.C., 1996.

9. Miaou, S-P., R.P. Bligh, D. Lord. "Developing Guidelines for Median Barrier Installation." *Transportation Research Record 1904.* Transportation Research Board, Washington, D.C., 2005, pp. 3-19.

CHAPTER 5. RURAL AND URBAN FREEWAY SEGMENTS

OVERVIEW

This chapter describes the activities undertaken to calibrate safety prediction models for rural and urban freeways in Texas. Each model provides an estimate of the expected crash frequency for a freeway segment with a specified length, traffic volume, and geometric design. A series of accident modification factors were also developed as part of the calibration activity. Three models were calibrated and are recommended for use in Texas. Collectively, these models address the following area type and lane combinations:

- rural freeway with four through lanes,
- rural freeway with six through lanes,
- rural freeway with eight through lanes,
- urban freeway with four through lanes,
- urban freeway with six through lanes,
- urban freeway with eight through lanes, and
- urban freeway with ten through lanes.

Equivalent models are described in the *Interim Roadway Safety Design Workbook* (*1*). These models were developed using data from locations outside of Texas. They are replaced by the models described in this chapter, all of which have been calibrated using data specific to Texas.

The objective of this research was to calibrate models for each of the aforementioned area type and lane combinations using crash data for Texas highways. Freeways with two-way traffic flow were the focus of the calibration activity. As part of the calibration process, several AMFs were also developed. These AMFs can be used with the calibrated models to estimate the change in safety associated with a specified change in geometry.

This chapter is divided into four parts. The first part provides a review of the literature on the topic of safety prediction models for rural and urban freeway segments. The second part describes a general AMF model framework for extending selected AMFs to a range of area type and lane combinations. The third part describes the method used to calibrate the proposed models. The last part describes the analysis and findings from the calibration process.

LITERATURE REVIEW

This part of the chapter describes a framework for safety prediction model calibration. It consists of four sections. The first section identifies the various median types found on freeways and establishes a vocabulary in this regard. The second section provides a brief overview of safety prediction and the role of "base" prediction models. The third section summarizes the base models described in the *Workbook*. The last section reviews the data available in TxDOT's road inventory database.

Freeway Median Types

A detailed review of about 3010 miles of rural and urban freeway segments in Texas indicates that the following median types are in use:

- positive (rigid) barrier median,
- raised-curb median, or
- depressed median.

These three median types are described herein as "restrictive" median types. About 76 percent of the freeway miles in Texas have a depressed median. Less than 1 percent of the freeway miles have a raised-curb median. Finally, about 24 percent have a positive barrier median. It is noted that the positive barrier in the database described in a subsequent section is a "rigid" barrier. The database does not include freeway segments with a cable barrier.

Safety Prediction Models

The expected crash frequency for a freeway segment with specified attributes is computed using a safety prediction model. This model represents the combination of a "base" model and one or more AMFs. The base model is used to estimate the expected crash frequency for a typical segment. The AMFs are used to adjust the base estimate when the attributes of the specific segment are not considered typical. The general form of the typical safety prediction model is shown below as Equation 1, and its base model component is shown as Equation 2.

with,

$$E[N] = E[N]_b \times AMF_{lw} \times AMF_{mw} \ldots \quad \textbf{(1)}$$

$$E[N]_b = a\ ADT^{\,b}\ L \quad \textbf{(2)}$$

where,

$E[N]$ = expected crash frequency, crashes/yr;
$E[N]_b$ = expected base crash frequency, crashes/yr;
AMF_{lw} = lane width accident modification factor;
AMF_{mw} = (restrictive) median width accident modification factor;
a, b = calibration coefficients;
ADT = average daily traffic volume, veh/d; and
L = segment length, mi.

When the coefficient b for the ADT variable is equal to 1.0, then the coefficient a in Equation 2 is effectively equal to the crash rate, with units of "crashes per million vehicle-miles."

Interim Workbook Base Models

This section describes base models derived by Bonneson et al. (*2*) for rural and urban freeway segments. These models were calibrated using data and models described in the research literature. They were subsequently documented in the *Workbook* in the form of a base model and associated base crash rates. The base model in the *Workbook* is:

$$E[N]_b = 0.000365 \; Base \; ADT \; L \quad (3)$$

where,

Base = injury (plus fatal) crash rate (from Table 5-1), crashes/million vehicle-miles.

A comparison of Equation 3 with Equation 2 indicates that the constant *a* is equal to the product of the base crash rate *Base* and "0.000365." The constant represents a conversion factor to convert the ADT into "millions of vehicles per year." Also, the ADT exponent *b* is equal to 1.0 in Equation 3.

The crash rate used in Equation 3 is provided in Table 5-1. They are based on a synthesis of the trends reported by several researchers in their examination of freeway crashes (*2*). Rates were not derived for rural freeways with 8 or 10 lanes.

Table 5-1. Base Crash Rates for Freeway Segments.

Area Type	Attributes	Base Crash Rate, severe crashes/mvm [1]		
	Through Lanes:	4	6	8 or 10
Urban		0.24	0.36	0.54
Rural		0.14	0.21	data not available

Note:
1 - mvm: million vehicle-miles.

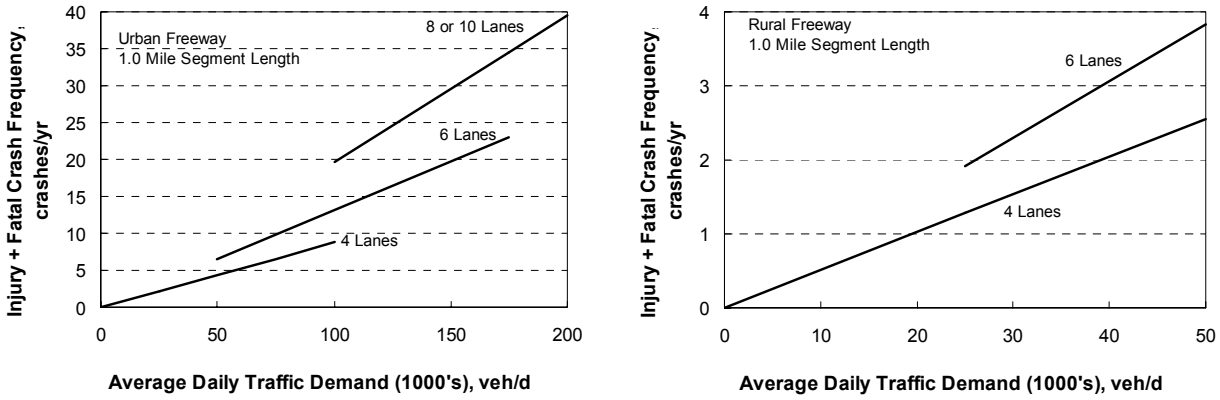
The relationship between crash frequency and traffic demand (represented by Equation 3 and the rates in Table 5-1) is shown in Figure 5-1 for a 1-mile freeway segment. The trend lines shown indicate that urban freeways have about 70 percent more crashes than rural freeways, for the same average daily traffic demand (ADT) and number of lanes. The trends also indicate that crash frequency is higher on freeways with many lanes than it is on freeways with few lanes.

For an ADT of 40,000 veh/d, the trends in Figure 5-1 indicate that a rural four-lane freeway has an expected crash frequency of 2.0 crashes/yr. In contrast, an urban four-lane freeway has an expected crash frequency of 3.50 crashes/yr.

The base model in Equation 3 is applicable to freeway segments that have typical geometric design conditions. These conditions are identified in the *Workbook* and are restated in column two of Table 5-2. The base conditions listed are representative of roads in Texas that are classified as freeway. The base model is combined with the AMFs described in the *Workbook* to yield the safety prediction model for freeway segments.

The characteristics in column 1 of Table 5-2 are each associated with an AMF to account for segment conditions that deviate from the base conditions. Each AMF was evaluated for a range of values to examine its sensitivity to a change in the associated variable. The results of this sensitivity analysis are listed in column 3 of Table 5-2. The ranges listed in these columns equate AMF values between 0.95 and 1.05. For example, an outside shoulder width between 8 and 12 ft yields AMF

values of about 1.05 and 0.95, respectively. The ranges listed provide some indication of the correlation between a change in the characteristic and the change in crash risk.



a. Urban Freeway. **b. Rural Freeway.**

Figure 5-1. Crash Prediction Model Reported in the *Workbook*.

Table 5-2. Base Conditions for Freeway Segments.

Characteristic	Base Condition	Sensitivity of Continuous Variables
Grade	flat (0% grade)	-3 to +3%
Lane width	12 ft	11 to 12 ft
Paved outside shoulder width	8 ft	8 to 12 ft
Paved inside shoulder width	4 ft (4 lane); 10 ft (6 or more lanes)	2 to 6 ft; 8 to 12 ft
Restrictive median width	24 ft (surfaced median); 76 ft (depressed median)	14 to 34 ft 60 to 92 ft
Shoulder rumble strips	not present	--
Utility pole density and offset	25 poles/mi; 30 ft average offset	Any density; 20 ft offset or more

Texas Highway Data

The geometric and traffic attributes for the Texas state highway system were obtained from the Texas Reference Marker (TRM) system highway database. This database is maintained by TxDOT and contains data for several thousand highway segments, each of which is described in terms of its geometry, traffic, and location attributes. In fact, about 140 attributes are used to describe each highway segment in TRM. However, only about 20 attributes are used to describe segment geometry or traffic characteristics, and about 10 more attributes are used to describe road name and physical location. The remaining attributes describe administrative designations and road-management-related data that are not particularly relevant to safety analysis.

A freeway segment in TRM is defined as a length of road along which no one attribute changes. A change in any attribute dictates the end of one segment and the start of a new segment. The average tangent (i.e., not curved) segment length in TRM is about 0.4 mi.

Several segment characteristics are needed in the calibration database. They are identified in Table 5-3. Column 3 of this table indicates whether the corresponding characteristic is available in the TRM database.

Table 5-3. Freeway Segment Calibration Data.

Attribute Category	Attribute	Availability in TRM Database
General	Area type (e.g., urban, rural)	Yes
Traffic control	Two-way vs. one-way operation	Yes
Traffic characteristics	ADT (by specified year)	Yes
	Truck percentage	Yes
Geometry	Horizontal curve radius	Yes
	Number of lanes	Yes
	Lane width	Yes
	Inside & outside shoulder width	Yes
	Median type	Yes
	Median width	Yes
	Utility pole density & offset	No
	Curb presence	Yes
	Ramp location and use	No
	Grade	No
	Rumble strip presence	No
	Side slope	No
	Horizontal clearance	No

The information in Table 5-3 indicates that data for several segment characteristics are not available in the TRM database. However, it was possible to obtain ramp location, ramp use (i.e., entrance or exit), and horizontal clearance data using high-resolution aerial photography available from the Internet. Utility pole density, grade, rumble strip presence, and side slope data could not be reliably obtained from any source. It was judged that their base condition values would be sufficiently representative of the collective set of segments in the database. This assumption allowed the corresponding AMF value to be set to 1.0.

Median width was measured as the distance between the near edges of the traveled way for the two opposing travel directions. Thus, median width includes the width of both inside shoulders.

AMF DEVELOPMENT

This part of the chapter describes the development of a general AMF model for incorporating the influence of area type and number-of-lanes into those AMFs that are associated with only specific crash types (e.g., run-off-the-road). The general model was first introduced in Chapter 3 of the *Synthesis* (*2*) and is derived in Chapter 3. It is extended to intersection applications in Chapter 2.

The general model is appropriate when a treatment applied to a segment (e.g., change shoulder width) has its most significant effect on specific crash types. The magnitude of this effect can be quantified using an AMF that is based on just the crashes of the specified types (i.e., AMF_{spc}). Alternatively, it can also be quantified using an AMF based on all crash types (i.e., AMF_{seg}). The general AMF model is:

$$AMF_{seg,j} = \left(AMF_{seg,i} - 1\right)\frac{P_{spc,j}}{P_{spc,i}} + 1.0 \qquad (4)$$

where,

$AMF_{seg,j}$ = AMF for a specified treatment to a segment with area type and lane combination j, quantified in terms of all crash types;
$AMF_{seg,i}$ = AMF for the same treatment to a segment with area type and lane combination i, quantified in terms of all crash types;
$P_{spc,j}$ = proportion of specific crash types on a segment with area type and lanes j; and
$P_{spc,i}$ = proportion of same crash types on a segment with area type and lanes i.

The proportion of specific crash types used in Equation 4 is provided in the *Workbook* (*1*) for several AMFs. These proportions were based on Texas crash data for the years 1999 and 2000. Since their derivation, crash data were made available for year 2001. These data were used to update the proportions to reflect a three-year average (i.e., 1999, 2000, and 2001). The results of this analysis are shown in Table 5-4. The proportions shown in each row are based on the total listed in the last column.

Table 5-4. Freeway Crash Distribution.

Area Type	Through Lanes	Proportion of Injury + Fatal Crashes					Total Crash Frequency [1], crashes per 3 years
		Single-Vehicle		Multiple-Vehicle		Single-Vehicle with Pole	
		Either Side	Right Side	Opp. Dir.	Sideswipe		
Rural	4	0.5141	0.2564	0.0393	0.0670	0.0222	8,953
	6	0.4308	0.1401	0.0277	0.1015	0.0185	650
Urban	4	0.3283	0.1469	0.0223	0.0931	0.0326	16,816
	6	0.2367	0.0888	0.0079	0.1273	0.0192	28,328
	8	0.2040	0.0655	0.0055	0.1671	0.0110	20,508
	10	0.2127	0.0708	0.0041	0.1902	0.0088	5,368

Note:
1 - Injury (plus fatal) crashes reported during the years 1999, 2000, and 2001.

The proportions listed in Table 5-4 were combined as needed for three AMFs in the *Workbook*. The combinations reflect the types of crashes that were found to be influenced by a change in the corresponding geometric element, as a function of area type and number of lanes. The resulting proportions to be used in Equation 4 are listed in Table 5-5. The values listed in this table are very similar to those reported in the *Workbook*.

Table 5-5. Freeway Crash Distribution for Selected AMFs.

Area Type	Through Lanes	Proportion of Injury + Fatal Crashes				
		Lane Width AMF	Outside Shoulder AMF [1]	Inside Shoulder AMF	Shoulder Rumble Strip	Utility Pole Offset AMF
Rural	4	0.62 [a]	0.26 [b]	0.30 [c]	0.51 [d]	0.022 [e]
	6	0.56 [a]	0.14 [b]	0.32 [c]	0.43 [d]	0.019 [e]
Urban	4	0.44 [a]	0.15 [b]	0.20 [c]	0.33 [d]	0.033 [e]
	6	0.37 [a]	0.089 [b]	0.16 [c]	0.24 [d]	0.019 [e]
	8	0.38 [a]	0.066 [b]	0.14 [c]	0.20 [d]	0.011 [e]
	10	0.41 [a]	0.071 [b]	0.15 [c]	0.21 [d]	0.0088 [e]

Notes:
a - Single-vehicle run-off-road, same-direction sideswipe, and multiple-vehicle opposite direction crashes.
b - Single-vehicle run-off-road crashes, right side.
c - Multiple-vehicle opposite direction and single-vehicle-run-off-road crashes, left side.
d - Single-vehicle run-off-road crashes, either side.
e - Single-vehicle-with-pole crashes.
1 - Also applies to the horizontal clearance AMF.

METHODOLOGY

This part of the chapter describes the methodology used to calibrate the freeway segment safety prediction model. It is divided into three sections. The first section describes the database used to calibrate the model. The second section describes the technique used to quantify the calibration coefficients. The last section describes the technique used to adjust the over-dispersion parameter for bias due to a small sample size.

Database

A database was assembled from the TRM database for segment calibration. Variables associated with the AMFs described in the *Workbook* (*1*) were included in the database to ensure that their relationship with crash occurrence was included in the analysis. However, as previously noted in the discussion associated with Table 5-3, utility pole location, grade, rumble strip presence, and side slope data are not in the TRM database. Moreover, measurement of these characteristics in the field was not a feasible option. A review of the aerial photos for each segment indicated (to the extent permitted by photo resolution) that the base condition values listed in Table 5-2 for these variables were representative of the collective set of segments in the database.

Ramp Location and Use

Ramps along a segment were counted and categorized as to whether they served traffic entering or exiting the freeway. Entrance-exit ramp pairs that formed a weaving section were also identified. In almost all cases, the ramp pairs were on the right-hand side of the roadway and joined by a common auxiliary lane. A couple of segments were found to have more complicated weaving that resulted when one of the ramps was on the left-hand side of the roadway.

The number of lanes associated with each ramp entrance was recorded in the database. Also, the number of lanes in the auxiliary lane of each weaving section was also recorded. However, almost all ramps and weaving sections were found to have only one lane. A small percentage had two lanes.

It is generally recognized that the length of the ramp entrance and the length of the weaving section have an important influence on their respective traffic capacities as well as the operation of the freeway segment. This influence stems from the interaction between freeway and ramp traffic streams as they change lanes to enter or exit the freeway. In recognition of a possible correlation between of this interaction with crash frequency, the length of the ramp entrance and the length of the weaving section were measured and included in the database. The conventions used to measure these lengths are shown in Figure 5-2.

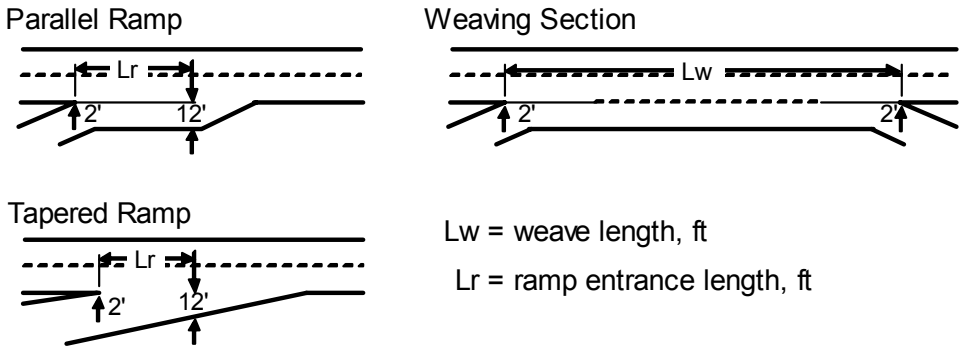


Figure 5-2. Ramp Entrance and Weaving Section Length Measurements.

Curve Radius and Length

The segments included in the database were selected from the TRM database such that they did not have curves with a radius of less than 5000 ft. This approach was undertaken to allow the calibration activity to focus on the calibration of models for tangent highway segments. However, when reviewing the aerial photograph of each segment, a small percentage of segments were found to have moderate curvature. In this situation, the curve's length and radius were recorded and added to the database.

The curve radius AMF described in Chapter 3 (i.e., Equation 16) was used account for the presence of curvature. The following equation was then used to compute the aggregate AMF for segments with combined curve and tangent portions:

$$AMF_{cr|agg} = \frac{(L - L_c)\,1.0 \; - \; L_c\,AMF_{cr}}{L} \qquad \textbf{(5)}$$

where,

$AMF_{cr|agg}$ = aggregated curve radius AMF for a segment with both tangent and curved portions;
AMF_{cr} = horizontal curve radius accident modification factor;
L_c = length of horizontal curve (= $I_c \times R$ / 5280 / 57.3), mi;
I_c = curve deflection angle, degrees; and
R = curve radius, ft.

In those few cases where a curve was found on the segment, it was often located entirely within the segment. However, sometimes only a portion of the curve would be located on the segment. In this situation, the length of curve recorded in the database represented only that portion of the curve that was within the segment boundaries.

Lateral Clearance Distances

Three lateral clearance distances were measured. Horizontal clearance was measured from the outside edge of the traveled way to either the nearest continuous obstruction (e.g., fence line, utility poles, etc.) or the near edge of the frontage road traveled way, whichever is closer to the freeway. If a barrier is present on the roadside for only a portion of the segment, horizontal clearance was measured only for that portion without the barrier.

Also measured was the "inside clearance distance" and the "outside clearance distance." When a barrier is not present in the median, the "inside clearance distance" represents the distance between the near edges of the inside shoulder for the two opposing travel directions (i.e., it equals the median width less the width of the two inside shoulders). When a barrier is present in the median, the inside clearance distance is computed for both travel directions, as measured from the edge of the inside shoulder to the nearest barrier face (as defined in the next section).

A similar approach was used to measure "outside clearance distance." When a barrier is not present on the roadside, the outside clearance distance equals the horizontal clearance less the width of the outside shoulder. When a barrier is present on the roadside, the outside clearance distance represents the distance between the edge of the outside shoulder and the nearest barrier face.

Longitudinal Barriers

Longitudinal barriers (i.e., concrete barrier, guardrail, or bridge rail) were noted when present on a segment. Attributes used to quantify barrier presence include barrier location, offset, length, and width. Barrier location indicates whether the barrier is parallel to the inside edge of the roadway or the outside edge of the roadway. Barrier offset, length, and width were each measured separately for inside and outside locations (e.g., inside barrier offset, outside barrier offset, etc.). Barrier offset

represents a lateral distance measured from the near edge of the shoulder to the face of the barrier. Barrier length represents the length of lane paralleled by a barrier; it is a total for both travel directions (e.g., if the outside barrier extends for the length of the roadway on both sides of the roadway, the outside barrier length equals twice the segment length). Barrier width represents either the physical width of the barrier if only one barrier is used, or the lateral distance between barrier "faces" if two parallel barriers are provided in the median area. A barrier face is the side of the barrier that is exposed to traffic.

Calibration Technique

The calibration activity used statistical analysis software that employs maximum likelihood methods based on a negative binomial distribution of crash frequency. The calibration model consisted of Equation 1, with the following model form used to estimate the base crash frequency:

$$E[N]_b = E[N]_{mv} + E[N]_{sv} + E[N]_{enr} + E[N]_{exr} \quad \textbf{(6)}$$

with,

$$E[N]_{mv} = a_{mv}\, ADT^{\,b_{mv}}\, L \quad \textbf{(7)}$$

$$E[N]_{sv} = a_{sv}\, ADT^{\,b_{sv}}\, L \quad \textbf{(8)}$$

$$E[N]_{enr} = a_{enr}\, (ADT/15000)^{b_{enr}}\, n_{enr} \quad \textbf{(9)}$$

$$E[N]_{exr} = a_{exr}\, (ADT/15000)^{b_{exr}}\, n_{exr} \quad \textbf{(10)}$$

where,

$E[N]_{mv}$ = expected multiple-vehicle non-ramp crash frequency for base conditions, crashes/yr;
$E[N]_{sv}$ = expected single-vehicle crash frequency for base conditions, crashes/yr;
$E[N]_{enr}$ = expected ramp-entrance-related crash frequency for base conditions, crashes/yr;
$E[N]_{exr}$ = expected ramp-exit-related crash frequency for base conditions, crashes/yr;
a_{mv}, b_{mv} = calibration coefficients for multiple-vehicle non-ramp crashes;
a_{sv}, b_{sv} = calibration coefficients for single-vehicle crashes;
a_{enr}, b_{enr} = calibration coefficients for ramp-entrance-related crashes;
a_{exr}, b_{exr} = calibration coefficients for ramp-exit-related crashes;
n_{enr} = number of ramp entrances; and
n_{exr} = number of ramp exits.

The form of this model was suggested by Harwood et al. (*3*). It was subsequently used in Chapter 3 for the development of the urban and suburban arterial street segment models. It was also used in Chapter 4 for the development of the rural multilane highway segment models.

The models represented by Equations 7, 8, 9, and 10 are referred to as "component models." Equations 1 and 6 through 10 represent one "aggregate model" and were simultaneously analyzed to find the best-fit calibration coefficients. In this manner, the component models were calibrated to compute the expected frequency of the corresponding crash type for a common set of AMFs (as

shown in Equation 1). The analysis approach is described in more detail in the next part of this chapter.

Over-Dispersion Parameter Adjustment

It was assumed that crash frequency on a segment is Poisson distributed, and that the distribution of the mean crash frequency for a group of similar intersections is gamma distributed. In this manner, the distribution of crashes for a group of similar intersections can be described by the negative binomial distribution. The variance of this distribution is:

$$V[X] = y\,E[N] + \frac{(y\,E[N])^2}{k\,L} \quad \textbf{(11)}$$

where,

$V[X]$ = crash frequency variance for a group of similar locations, crashes^2;
X = reported crash count for y years, crashes;
y = time interval during which X crashes were reported, yr; and
k = over-dispersion parameter, mi^{-1}.

In recognition of the concerns expressed by Lord (*4*), the over-dispersion parameter k obtained from the regression analysis is adjusted downward, such that it yields a more reliable estimate of the variance of the crash distribution. The derivation of the adjustment technique used is described in Reference *5*. The following equation was used to estimate the true over-dispersion parameter, given the estimated over-dispersion parameter obtained from the regression analysis.

$$k_t = \frac{\sqrt{(n-p)^2\,m^2 + 69\,k_r\,L.\,(n-p)\,m} - (n-p)\,m}{34.5\,L.} \quad \textbf{(12)}$$

where,

k_t = true over-dispersion parameter;
k_r = estimated over-dispersion parameter obtained from database analysis;
n = number of observations (i.e., segments in database);
p = number of model variables;
$L.$ = average segment length for all n observations (= 1.0 for intersection applications); and
m = average number of crashes per observation (= total crashes in database / n).

Equation 12 was used to estimate the true over-dispersion parameter for each of the models described in this chapter. All subsequent references to the over-dispersion parameter k in this chapter denote the estimated true parameter obtained from Equation 12 (i.e., hereafter, $k = k_t$).

CALIBRATION DATA AND ANALYSIS

This part of the chapter describes the calibration of the aggregate freeway segment safety prediction model. Initially, the process used to select the segments is identified. Then, the characteristics of the segments are summarized. Next, the techniques used to assemble the crash data segment are described. Finally, the findings from the model calibration are discussed. In this section, a "site" is used to refer to one freeway segment.

Site Selection and Data Collection

The freeway segments used to calibrate the safety prediction model were required to satisfy the following criteria:

- Area type: rural or urban;
- Functional class: freeway (i.e., interstate or urban principal arterial highway);
- Segment length: 0.1 mi or more; and
- Curve presence: no curves with a radius less than 5000 ft.

All segments were required to have a minimum length of 0.1 mi. The value of 0.1 mi reflects a recognition of the precision of crash location in TxDOT's state-wide crash database.

The TRM database was screened to exclude segments that had one or more curves with a radius less than 5000 ft. This criterion was intended to allow the calibration to focus on tangent segments. However, a visual inspection of aerial photography for each segment revealed the presence of a curve on some segments. Those segments with flatter curves were kept in the database, and the effect of curvature was modeled using Equation 5. Segments with sharp curvature were removed from the database.

The synthesis by Bonneson et al. (*2*) indicates that segment crash frequency is significantly influenced by area type and number or lanes. This influence is so significant and varied that a separate model should be calibrated for each unique combination of area type and lanes that exist in the database. A review of the TRM database indicated that the following combinations are most common in Texas:

- rural freeway with four through lanes,
- rural freeway with six through lanes,
- rural freeway with eight through lanes,
- urban freeway with four through lanes,
- urban freeway with six through lanes,
- urban freeway with eight through lanes, and
- urban freeway with ten through lanes.

Thus, in addition to the aforementioned site selection criteria, effort was expended during the site selection process to target enough segments for each combination such that separate models could be accurately calibrated. To achieve this goal, a minimum sample size criterion based on total exposure was established for each combination.

Site Characteristics

Traffic and geometry data were identified for several thousand freeway segments using the TRM database for year 2003, the first year for which a complete database was available. All total, 588 segments satisfied the site selection criteria. These segments represent freeways located in 16 of the 25 TxDOT districts. Selected characteristics are provided for these segments in Table 5-6.

Table 5-6. Summary Characteristics for Freeway Segments.

Area Type	Through Lanes	Total Segments	Segment Length, mi	Seg. Length Range, mi		Volume Range, veh/d	
				Minimum	Maximum	Minimum	Maximum
Rural	4	212	252.5	0.11	7.9	7,200	69,100
	6	21	26.4	0.18	5.3	31,100	72,700
	8	3	1.8	0.39	0.8	109,000	140,000
Urban	4	135	101.1	0.12	2.5	4,800	92,300
	6	118	84.3	0.06	2.7	24,300	170,000
	8	90	51.2	0.11	1.9	33,500	198,000
	10	9	6.8	0.17	1.5	152,000	234,000
Overall:		588	524.0	0.06	7.9	4,800	234,000

The barrier and ramp characteristics for freeway segments are listed in Table 5-7. This table also lists the average inside and outside clearance distances represented in the database. Of particular note is the tendency for reduced lateral clearance and increased proportion of barrier as the number of lanes increases. This trend suggests that segments with more lanes have narrower medians and are more likely to have median barrier. For the same number of lanes, urban segments typically have a larger proportion of barrier. Finally, there is a trend toward a higher proportion of the urban segment length being used for ramp entrances or weaving sections, relative to rural segments.

Table 5-7. Barrier and Ramp Characteristics for Freeway Segments.

Area Type	Through Lanes	Lateral Clearance		Proportion Barrier		Proportion Ramp		Number of Ramps	
		Inside [1]	Outside [2]	Inside [3]	Outside [3]	Entrance [4]	Weave [4]	Entrance	Exit
Rural	4	42.9	80.6	0.058	0.096	0.025	0.004	166	170
	6	21.5	143.1	0.741	0.193	0.032	0.003	28	30
	8	7.0	41.0	1.000	0.295	0.007	0.000	2	2
Urban	4	44.3	94.8	0.282	0.210	0.028	0.029	143	152
	6	21.2	124.5	0.819	0.374	0.042	0.098	131	144
	8	17.6	113.9	0.796	0.482	0.052	0.150	106	125
	10	9.9	49.1	1.000	0.616	0.062	0.107	12	15
Overall:		23.5	92.4	0.671	0.324	0.246	0.391	588	638

Notes:
1 - For restrictive medians, inside clearance equals the median width less the width of the two inside shoulders.
2 - Outside clearance represents the distance between the outside shoulder and the nearest continuous obstruction.
3 - Proportion of segment length with rigid barrier. Inside barrier is in the median. Outside barrier is on the roadside.
4 - Proportion of segment length associated with total ramp entrance length or total weave length (see Figure 5-2).

Data Collection

Crash data were identified for each segment using the TxDOT crash database. Three years of crash data, corresponding to years 1999, 2000, and 2001, were identified for each segment. The

ADT for each of these three years was obtained from the TRM database and averaged to obtain one ADT for each segment. Crash data prior to 1999 were not used due to the increase in speed limit that occurred on many Texas freeways in 1997 and 1998.

The crash records associated with each segment were categorized in terms of whether they were multiple-vehicle non-ramp crashes, single-vehicle crashes, ramp-entrance-related crashes, or ramp-exit-related crashes. Ramp-related crashes represent crashes between a vehicle in the ramp entrance (or exit) area and a vehicle on the main lanes; they do not represent crashes on the ramp proper. Only crashes that were identified as associated with one or more injuries or fatalities (i.e., K, A, B, C severity rating) were included in the database.

Data Analysis

This section is divided into two subsections. The first subsection summarizes the crash data at the selected study sites. The second subsection describes the formulation of the calibration model and summarizes the statistical analysis methods used to calibrate it.

Database Summary

The crash data for the 588 segments are summarized in Table 5-8. These segments were collectively associated with 8436 injury or fatal crashes. The trends in crash rate are provided in the last column of the table. Table 5-1 indicates that the injury (plus fatal) crash rate for rural freeways with four and six lanes is 0.14 and 0.21, respectively. It also indicates that the rate for urban freeways is 0.24 and 0.36 for four and six lanes, respectively. These rates are consistent with the injury (plus fatal) crash rates in Table 5-8. They confirm the trends found in Table 5-1 (i.e., that urban freeway segments experience more crashes than rural segments, and that crash rate increases with an increase in the number of lanes).

Table 5-8. Crash Data Summary for Freeway Segments.

Area Type	Through Lanes	Exposure,[1] mvm	Injury + Fatal Crashes / 3 years					Crash Rate, cr/mvm
			Multiple-Vehicle [2]	Single-Vehicle	Ramp Entrance	Ramp Exit	Total	
Rural	4	1904.2	413	643	5	5	1066	0.19
	6	459.1	133	142	1	5	281	0.20
	8	82.2	86	28	1	1	116	0.47
Urban	4	1468.9	732	538	16	10	1296	0.29
	6	2452.4	1630	982	42	8	2662	0.36
	8	2290.6	1661	796	23	15	2495	0.36
	10	448.4	363	149	5	3	520	0.39
Overall:		9105.8	5018	3278	93	47	8436	0.31

Notes:
1 - mvm: million vehicle-miles.
2 - Multiple-vehicle crashes do not include ramp-related crashes.

Model Development and Statistical Analysis Methods

This section describes the proposed freeway segment safety prediction model and the methods used to calibrate it. The form of this model is:

$$E[N] = E[N]_b \times AMF_{lw} \times AMF_{osw} \times AMF_{isw} \times AMF_{cr|agg} \times AMF_{ic|agg} \times AMF_{oc|agg} \times AMF_{enr|agg} \times AMF_{wev|agg} \times AMF_{tk} \quad \textbf{(13)}$$

with,

$$E[N]_b = E[N]_{mv} I_{mv} + E[N]_{sv} I_{sv} + E[N]_{enr} I_{enr} + E[N]_{exr} I_{exr} \quad \textbf{(14)}$$

$$E[N]_{mv} = E[N]_{mv4} I_4 + E[N]_{mv6} I_6 + E[N]_{mv8} I_8 + E[N]_{mv10} I_{10} \quad \textbf{(15)}$$

$$E[N]_{mv4} = e^{b_{mv4} + b_{mv} Ln(ADT/1000) + Ln(L) + b_{mv,rural} I_{rural}} \quad \textbf{(16)}$$

$$E[N]_{mv6} = e^{b_{mv6} + b_{mv} Ln(ADT/1000) + Ln(L) + b_{mv,rural} I_{rural}} \quad \textbf{(17)}$$

$$E[N]_{mv8} = e^{b_{mv8} + b_{mv} Ln(ADT/1000) + Ln(L) + b_{mv,rural} I_{rural}} \quad \textbf{(18)}$$

$$E[N]_{mv10} = e^{b_{mv10} + b_{mv} Ln(ADT/1000) + Ln(L) + b_{mv,rural} I_{rural}} \quad \textbf{(19)}$$

$$E[N]_{sv} = E[N]_{sv4} I_4 + E[N]_{sv6} I_6 + E[N]_{sv8} I_8 + E[N]_{sv10} I_{10} \quad \textbf{(20)}$$

$$E[N]_{sv4} = e^{b_{sv4} + b_{sv} Ln(ADT/1000) + Ln(L) + b_{sv,rural} I_{rural}} \quad \textbf{(21)}$$

$$E[N]_{sv6} = e^{b_{sv6} + b_{sv} Ln(ADT/1000) + Ln(L) + b_{sv,rural} I_{rural}} \quad \textbf{(22)}$$

$$E[N]_{sv8} = e^{b_{sv8} + b_{sv} Ln(ADT/1000) + Ln(L) + b_{sv,rural} I_{rural}} \quad \textbf{(23)}$$

$$E[N]_{sv10} = e^{b_{sv10} + b_{sv} Ln(ADT/1000) + Ln(L) + b_{sv,rural} I_{rural}} \quad \textbf{(24)}$$

$$E[N]_{enr} = E[N]_{enr4} I_4 + E[N]_{enr6} I_6 + E[N]_{enr8} I_8 + E[N]_{enr10} I_{10} \quad \textbf{(25)}$$

$$E[N]_{enr4} = n_{enr} \, e^{b_{enr4} + b_{enr} Ln(ADT/15000) + b_{enr,rural} I_{rural}} \quad \textbf{(26)}$$

$$E[N]_{enr6} = n_{enr} \, e^{b_{enr6} + b_{enr} Ln(ADT/15000) + b_{enr,rural} I_{rural}} \quad \textbf{(27)}$$

$$E[N]_{enr8} = n_{enr} \, e^{b_{enr8} + b_{enr} Ln(ADT/15000) + b_{enr,rural} I_{rural}} \quad \textbf{(28)}$$

$$E[N]_{enr10} = n_{enr} \, e^{b_{enr10} + b_{enr} Ln(ADT/15000) + b_{enr,rural} I_{rural}} \quad \textbf{(29)}$$

$$E[N]_{exr} = E[N]_{exr4} I_4 + E[N]_{exr6} I_6 + E[N]_{exr8} I_8 + E[N]_{exr10} I_{10} \quad (30)$$

$$E[N]_{exr4} = n_{exr} \, e^{b_{exr4} + b_{exr} Ln(ADT/15000) + b_{exr,rural} I_{rural}} \quad (31)$$

$$E[N]_{exr6} = n_{exr} \, e^{b_{exr6} + b_{exr} Ln(ADT/15000) + b_{exr,rural} I_{rural}} \quad (32)$$

$$E[N]_{exr8} = n_{exr} \, e^{b_{exr8} + b_{exr} Ln(ADT/15000) + b_{exr,rural} I_{rural}} \quad (33)$$

$$E[N]_{exr10} = n_{exr} \, e^{b_{exr10} + b_{exr} Ln(ADT/15000) + b_{exr,rural} I_{rural}} \quad (34)$$

$$AMF_{ic|agg} = (1.0 - P_{ib}) \, AMF_{ic|is} + P_{ib} \, AMF_{ic|ib} \, AMF_{b|ib} \quad (35)$$

$$AMF_{oc|agg} = (1.0 - P_{ob}) \, AMF_{oc|os} + P_{ob} \, AMF_{oc|ob} \, AMF_{b|ob} \quad (36)$$

$$AMF_{enr|agg} = (1.0 - P_{enr}) \, 1.0 + P_{enr} \, AMF_{enr} \quad (37)$$

$$AMF_{wev|agg} = (1.0 - P_{wev}) \, 1.0 + P_{wev} \, AMF_{wev} \quad (38)$$

$$AMF_{tk} = e^{b_{tk}(P_t - 20)} \quad (39)$$

and

$$AMF_{ic|is} = e^{-0.0296 \left([W_m - 2 W_{is}]^{0.5} - [W_{mb} - 2 W_{isb}]^{0.5}\right)} \quad (40)$$

$$AMF_{ic|ib} = e^{-0.0296 \left([2 W_{icb}]^{0.5} - [W_{mb} - 2 W_{isb}]^{0.5}\right)} \quad (41)$$

$$AMF_{b|ib} = e^{b_{bar}/W_{icb}} \quad (42)$$

$$AMF_{oc|os} = \left(e^{-0.014 \left([W_{hc} - W_s] - 20\right)} - 1.0\right) P_{hc} + 1.0 \quad \geq 1.0 \quad (43)$$

$$AMF_{oc|ob} = \left(e^{-0.014 (W_{ocb} - 20)} - 1.0\right) P_{hc} + 1.0 \quad \geq 1.0 \quad (44)$$

$$AMF_{b|ob} = e^{b_{bar}/W_{ocb}} \quad (45)$$

$$AMF_{enr} = e^{b_{enr}/L^*_{enr}} \quad (46)$$

$$AMF_{wev} = e^{b_{enr}/L^*_{wev}} \quad (47)$$

where,

AMF_{lw} = lane width accident modification factor;
AMF_{osw} = outside shoulder width accident modification factor;
AMF_{isw} = inside shoulder width accident modification factor;
$AMF_{cr|agg}$ = aggregated curve radius accident modification factor (Equation 5);
$AMF_{ic|agg}$ = aggregated inside clearance accident modification factor;
$AMF_{oc|agg}$ = aggregated outside clearance accident modification factor;
$AMF_{enr|agg}$ = aggregated ramp entrance accident modification factor;
$AMF_{wev|agg}$ = aggregated weaving section accident modification factor;
AMF_{tk} = truck presence accident modification factor;
I_{mv} = crash indicator variable (= 1.0 if multiple-vehicle non-ramp crash data, 0.0 otherwise);
I_{sv} = crash indicator variable (= 1.0 if single-vehicle crash data, 0.0 otherwise);
I_{enr} = crash indicator variable (= 1.0 if ramp-entrance-related crash data, 0.0 otherwise);
I_{exr} = crash indicator variable (= 1.0 if ramp-exit-related crash data, 0.0 otherwise);
$E[N]_{mvX}$ = expected multiple-vehicle non-ramp crash frequency for number-of-lanes X (X = 4, 6, 8, 10); crashes/yr;
I_X = cross section indicator variable (= 1.0 if cross section has X lanes, 0.0 otherwise);
I_{rural} = area type indicator variable (= 1.0 if area is rural, 0.0 if it is urban);
$E[N]_{svX}$ = expected single-vehicle crash frequency if cross section has X lanes; crashes/yr;
$E[N]_{enrX}$ = expected ramp-entrance-related crash frequency if cross section has X lanes; crashes/yr;
$E[N]_{exrX}$ = expected ramp-exit-related crash frequency if cross section has X lanes; crashes/yr;
n_{enr} = number of ramp entrances;
n_{exr} = number of ramp exits;
P_{ib} = proportion of segment length with a barrier present in the median (i.e., inside);
P_{ob} = proportion of segment length with a barrier present on the roadside (i.e., outside);
$AMF_{ic|is}$ = inside clearance accident modification factor based on distance between inside shoulders (applies to segment portion *without* a barrier in median);
$AMF_{ic|ib}$ = inside clearance accident modification factor based on clearance distance to inside barrier (applies to segment portion *with* a barrier in median);
$AMF_{b|ib}$ = rigid barrier accident modification factor based on clearance distance to inside barrier;
$AMF_{oc|os}$ = outside clearance accident modification factor based on distance from outside shoulder to nearest continuous obstruction (applies to segment portion *without* a roadside barrier);
$AMF_{oc|ob}$ = outside clearance accident modification factor based on clearance distance to outside barrier (applies to segment portion *with* a roadside barrier);
$AMF_{b|ob}$ = rigid barrier accident modification factor based on clearance distance to outside barrier;
P_{enr} = proportion of segment length adjacent to a ramp entrance;
AMF_{enr} = ramp entrance accident modification factor;
P_{wev} = proportion of segment length adjacent to a weaving section;
AMF_{wev} = weaving section accident modification factor;
P_t = percent trucks represented in ADT, %;
W_m = median width (measured from near edges of traveled way in both travel directions), ft;
W_{mb} = base median width (see Table 5-9), ft;

W_{is} = inside shoulder width, ft;
W_{isb} = base inside shoulder width (see Table 5-9), ft;
W_{icb} = inside clearance distance when a barrier is present in the median, ft; and
W_s = outside shoulder width, ft;
W_{ib} = inside barrier width (measured from barrier face to barrier face), ft;
W_{hc} = horizontal clearance, ft;
W_{ocb} = outside clearance distance when a barrier is present on the roadside, ft;
P_{hc} = proportion of single-vehicle run-off-road crashes;
L^*_{enr} = average ramp entrance length, ft;
L^*_{wev} = average weaving section length, ft; and
b_i = calibration coefficient for condition i (see Table 5-10).

The presence of a barrier along some portions of the segment (and not along other portions) has led to the derivation of aggregate AMFs for barrier in the median and on the roadside. The equations for computing these AMFs (i.e., Equations 35 and 36) include one or more AMFs that apply when the barrier exists for the entire length of the segment. Key variables for these AMFs (i.e., Equations 40 to 45) include the appropriate base median width and base inside shoulder width. These variables are listed in Table 5-9.

Table 5-9. Base Condition Median Width and Inside Shoulder Width.

Median Type	Through Lanes	Base Median Width (W_{mb}), ft	Base Inside Shoulder Width (W_{isb}), ft
Positive Barrier	4	36	4
	6 or more	26	10
Depressed (no barrier)	4	56	4
	6 or more	56	10

Other key variables that are needed are the inside clearance distance when a barrier is present W_{icb} and outside clearance distance when a barrier is present W_{ocb}. As indicated in Equations 42 and 45, this distance is included as a divisor in the exponential term. This relationship implies that the correlation between clearance distance and crash frequency decreases rapidly with increasing distance and converges to "no influence" at long distances. When multiple sections of barrier exist along the segment, a length-weighted average of the reciprocal of the clearance distance is computed. The length used to weight the average is the barrier length. The reciprocal of distance is used because of its use as a divisor in Equation 42 or 45.

Additional key variables include the proportion of segment length with a barrier present in the median P_{ib} and the proportion of segment length with a barrier present on the roadside P_{ob}. Equations for calculating these proportions and the aforementioned clearance distances are described in the following paragraphs.

For segments with a continuous barrier centered in the median (i.e., symmetric median barrier), the following equations should be used to estimate W_{icb} and P_{ib}:

$$W_{icb} = \frac{2\,L}{\frac{L_{ib,sh}}{2.0} + \sum \frac{L_{ib,off}}{W_{off}} + \frac{2\,L - L_{ib,sh} - L_{ib,off}}{0.5\,(W_m - 2W_{is} - W_{ib})}} \quad \textbf{(48)}$$

$$P_{ib} = 1.0 \quad \textbf{(49)}$$

where,

$L_{ib,\,sh}$ = length of inside lane paralleled by a barrier located at the edge of the inside shoulder, total of both travel directions, mi;

$L_{ib,\,off}$ = length of inside lane paralleled by a barrier located at distance W_{off} from the inside shoulder, total of both travel directions, mi; and

W_{off} = clearance distance between the edge of the shoulder and the nearest barrier face, ft.

The summation term "∑" in Equation 48 indicates that the ratio of barrier length $L_{ib,\,off}$ to clearance distance W_{off} should be computed for each individual length of barrier that is found in the median along the segment (e.g., a barrier protecting a sign support, a bridge rail, etc.). The constant "2.0" in Equation 48 (and subsequent clearance equations) represents the minimum offset between the edge of shoulder and the barrier face.

For segments with a continuous barrier adjacent to one roadbed (i.e., asymmetric median barrier), the following equations should be used to estimate W_{icb} and P_{ib}:

$$W_{icb} = \frac{2\,L}{\frac{L}{2.0} + \sum \frac{L_{ib,sh}}{2.0} + \sum \frac{L_{ib,off}}{W_{off}} + \frac{L - L_{ib,sh} - L_{ib,off}}{W_m - 2W_{is} - W_{ib}}} \quad \textbf{(50)}$$

$$P_{ib} = 1.0 \quad \textbf{(51)}$$

Similar to the previous guidance, the summation term "∑" in Equation 50 indicates that the ratio of barrier length $L_{ib,\,off}$ to clearance distance W_{off}, as well as the ratio of barrier length $L_{ib,\,sh}$ to a clearance distance of 2.0 ft, should be computed for each individual length of barrier that is found in the median along the segment (e.g., a barrier protecting a sign support, a bridge rail, etc.). Each length of barrier should be included in one of the summation terms (but not both), depending on whether the barrier is at the edge of the shoulder or offset further into the median.

For segments with a depressed median and some short sections of barrier in the median (e.g., bridge rail), the following equations should be used to estimate W_{icb} and P_{ib}:

$$W_{icb} = \frac{\sum L_{ib,sh} + \sum L_{ib,off}}{\sum \frac{L_{ib,sh}}{2.0} + \sum \frac{L_{ib,off}}{W_{off}}} \quad \textbf{(52)}$$

$$P_{ib} = \frac{\sum L_{ib,sh} + \sum L_{ib,off}}{2\,L} \quad \textbf{(53)}$$

For depressed medians without a continuous barrier or short sections of barrier in the median, the following equation should be used to estimate P_{ib}.

$$P_{ib} = 0.0 \quad \textbf{(54)}$$

As suggested by Equation 35, the calculation of W_{icb} is not required when $P_{ib} = 0.0$.

For segments with a barrier on the roadside, the following equations should be used to estimate W_{ocb} and P_{ob}:

$$W_{ocb} = \frac{\sum L_{ob,sh} + \sum L_{ob,off}}{\sum \frac{L_{ob,sh}}{2.0} + \sum \frac{L_{ob,off}}{W_{off}}} \quad \textbf{(55)}$$

$$P_{ob} = \frac{\sum L_{ob,sh} + \sum L_{ob,off}}{2L} \quad \textbf{(56)}$$

where,

$L_{ob,\,sh}$ = length of outside lane paralleled by a barrier located at the edge of the outside shoulder, total of both travel directions, mi;

$L_{ob,\,off}$ = length of outside lane paralleled by a barrier located at distance W_{off} from the outside shoulder, total of both travel directions, mi; and

W_{off} = clearance distance between the edge of the shoulder and the nearest barrier face, ft.

For segments without a barrier on the roadside, the following equation should be used to estimate P_{ob}:

$$P_{ob} = 0.0 \quad \textbf{(57)}$$

As suggested by Equation 36, the calculation of W_{ocb} is not required when $P_{ob} = 0.0$.

The presence of a ramp entrance or weaving section along some portions of the segment (and not along other portions) has led to the derivation of aggregate AMFs for ramp entrance length and weaving section length. The equations for computing these AMFs (i.e., Equations 37 and 38) include one or more AMFs that apply when the entrance or weaving section exists for the entire length of the segment. Key variables for these AMFs (i.e., Equations 46 and 47) include the average ramp entrance length L^*_{enr} and the average weaving section length L^*_{wev}.

As indicated in Equations 46 and 47, the average ramp entrance length and the average weaving section length are each included as a divisor in their respective exponential term. This relationship implies that the correlation between length and crash frequency decreases rapidly with increasing length and converges to "no influence" at long lengths. When multiple ramp entrances or weaving sections exist along the segment, a length-weighted average of the reciprocal of the ramp entrance length or weaving section length is computed. The length used to weight the average is the length of entrance (or weaving section) that exists on the subject segment.

Additional key variables include the proportion of segment length associated with one or more ramp entrances P_{enr} and the proportion of segment length associated with one or more weaving sections P_{wev}. Equations for calculating these proportions and the aforementioned clearance distances are described in the following paragraphs.

For segments with one or more ramp entrances, the following equations should be used to estimate the average ramp entrance length L^*_{enr} and the proportion of segment length adjacent to a ramp entrance P_{enr}:

$$L^*_{enr} = 5280 \frac{\sum L_{enr,seg}}{\sum \frac{L_{enr,seg}}{L_{enr}}} \quad \textbf{(58)}$$

$$P_{enr} = \frac{\sum L_{enr,seg}}{2L} \quad \textbf{(59)}$$

where,

$L_{enr,\,seg}$ = length of ramp entrance that exists within the subject segment, mi; and
L_{enr} = length of ramp entrance (may extend beyond segment boundaries), mi.

The summation term "∑" in the divisor of Equation 58 indicates that the ratio of $L_{enr,seg}$ to L_{enr} should be computed for each individual ramp entrance that is found along the segment.

For segments with one or more weaving section, the following equations should be used to estimate the average weaving section length L^*_{wev} and the proportion of segment length adjacent to a weaving section P_{wev}:

$$L^*_{wev} = 5280 \frac{\sum L_{wev,seg}}{\sum \frac{L_{wev,seg}}{L_{wev}}} \quad \textbf{(60)}$$

$$P_{wev} = \frac{\sum L_{wev,seg}}{2L} \quad \textbf{(61)}$$

where,

$L_{wev,\,seg}$ = length of weaving section that exists within the subject segment, mi; and
L_{wev} = length of weaving section (may extend beyond segment boundaries), mi.

The summation term "∑" in the divisor of Equation 60 indicates that the ratio of $L_{wev,seg}$ to L_{wev} should be computed for each individual weaving section that is found along the segment.

Horizontal clearance is measured from the outside edge of the traveled way to the nearest continuous obstruction on the roadside. It is measured for both travel directions. Two values of horizontal clearance are obtained for a segment. These two values are then averaged to obtain the horizontal clearance W_{hc} used in the equations above. Given recognized variations in horizontal

clearance along most segments, the computed value should be determined such that it is representative of the entire length of the segment. By definition, horizontal clearance W_{hc} must equal or exceed the width of the outside shoulder.

Equations 43 and 44 replicate the horizontal clearance AMF in the *Workbook* (*1*). The value of "20" in these equations represents the base outside clearance distance. It is computed as the difference between the base horizontal clearance and the base outside shoulder width (i.e., 20 = 30 - 10). The horizontal clearance AMF is based on horizontal clearances ranging from 0 to 30 ft. In recognition of this restriction, the AMF value is limited to 1.0 for horizontal clearances that exceed 30 ft (i.e., an outside clearance distance of 20 ft).

Inside and outside clearance distances were used in the model instead of median width and horizontal clearance. This approach was followed because the two "clearance distance" variables do not include shoulder width in their dimension. AMFs for inside and outside shoulder width are already included in the model. Thus, the direct use of median width or horizontal clearance AMFs would "double count" the effect of a change in shoulder width.

The mathematical relationships for each of the first three AMFs identified in the variable list are provided in the *Workbook*. A preliminary examination of the data indicated that these AMFs were accurate estimators of the relationship between lane width, outside shoulder width, or inside shoulder width and multiple-vehicle, single-vehicle, or ramp-related crashes. For this calibration activity, the proportions in Table 5-5 were used with the AMFs from the *Workbook* (as opposed to the proportions provided in the *Workbook*).

The mathematical relationship for the truck presence AMF identified in the *Workbook* was not used in the proposed safety prediction model. A preliminary examination of the data indicated that this AMF did not accurately replicate the trends in the crash data. For this reason, it was determined that this AMF should be calibrated using the assembled database. An examination of its effect on multiple-vehicle, single-vehicle, and ramp-related crashes did not reveal a compelling reason to develop separate AMFs for each crash type.

An examination of the data indicated that median width is highly correlated with barrier presence and barrier offset in the database. For this reason, it was determined that either the rigid barrier AMF or the inside clearance AMF could be calibrated using the assembled database (but not both). An examination of the data indicated that the inside clearance AMFs (i.e., $AMF_{ic|is}$, $AMF_{ic|ib}$) calibrated for rural multilane highways (see Chapter 4) were representative of the associated trends in the freeway database. Hence, these two AMFs were included in the regression model as Equations 40 and 41, thereby leaving the rigid barrier AMF to be calibrated.

The utility pole density AMF identified in the *Workbook* is not included in the proposed model because pole count and location data were not readily available. It is assumed that the collective set of segments in the database is represented by the base condition of a 30-ft pole offset and 25 poles/mi. As such, the model should not be biased as a result of excluding this AMF. A similar situation was present for grade, rumble strip presence, and side slope data. It was assumed that base conditions for these variables were represented in the collective set of segments.

The Nonlinear Regression procedure (NLIN) in the SAS software was used to estimate the proposed model coefficients (*6*). This procedure was used because the proposed safety prediction model is both nonlinear and discontinuous. The loss function associated with NLIN was specified to equal the log likelihood function for the negative binomial distribution. Equation 11 was used to define the variance function for the multiple-vehicle and single-vehicle crashes. Equation 11, with the segment length term L removed, was used to define the variance function for the ramp-related crashes. The procedure was set up to estimate model coefficients based on maximum-likelihood methods.

One disadvantage of the NLIN procedure is that it is not able to compute the best-fit value of k for the calibrated prediction model. This disadvantage is overcome by using the Generalized Modeling (GENMOD) procedure in SAS. GENMOD automates the over-dispersion factor estimation process using maximum-likelihood methods. Initially, NLIN is used to calibrate the prediction model. Then, GENMOD is used to regress the relationship between the reported and predicted crash frequencies (where the natural log of the predicted values is specified as an offset variable, and the log link function is used) to obtain a best-fit over-dispersion estimate a using the internal variance function ($V[X] = E[N] + a\,E[N]^2$). The estimate a from GENMOD is then multiplied by the average segment length and inverted to estimate a new value of k (i.e., $k_{new} = 1/[a\,L.]$). This new value is then used in a second application of NLIN and the process is repeated until convergence is achieved between the k value used in NLIN and that obtained from GENMOD. Convergence is typically achieved in two iterations. As a last step, the value of k obtained from this iterative method was used in Equation 12 to estimate the true over-dispersion parameter.

Model Calibration

The safety prediction model calibration process consisted of the simultaneous calibration of all four component models (i.e., multiple-vehicle non-ramp, single-vehicle, ramp-entrance-related, and ramp-exit-related models) and AMFs using the aggregate model represented by Equation 13. The simultaneous calibration approach was needed because the AMFs in Equation 13 were common to all four component models. The database assembled for this process included four replications of the original database. However, the dependent variable in the first replication was set equal to the multiple-vehicle crashes. The dependent variable in the second replication was set equal to the single-vehicle crashes. That for the third replication was set equal to the ramp-entrance-related crashes and that for the fourth replication was set equal to the ramp-exit-related crashes.

The process of replicating the original database resulted in the variability in the data being underestimated because the random error in each set of three observations per site is correlated. The NLIN procedure does not directly accommodate the analysis of repeated measures but the GENMOD procedure does support this analysis. Thus, to estimate the true variability in each regression coefficient estimate from NLIN, the GENMOD procedure was used to regress the relationship between the reported and predicted crash frequencies. GENMOD was instructed to use generalized estimating equations to estimate the standard deviation of the intercept term accounting for the repeated measures. The ratio of this standard deviation to that from a similar model that assumed independent errors was found to be 1.15. The standard deviation for each regression coefficient obtained from NLIN was multiplied by this ratio to estimate its true standard deviation.

For each segment, the predicted crash frequency from each component model was totaled and compared with the total reported crash frequency. The difference between the two totals was then summed for all segments. This sum was found to be very small (i.e., less than 0.5 percent of the total reported crash frequency), so it was concluded that there was no bias in the component models in terms of their ability to predict total crash frequency.

The calibration coefficients for the aggregate safety prediction model are described in the next subsection. Then, the subsequent four subsections collectively describe each of the four component models and their fit to the data. The fit statistics were separately computed using the calibrated component model and an analysis of its residuals.

Aggregate Model

The results of the aggregated safety prediction model calibration are presented in Table 5-10. Calibration of this model focused on injury (plus fatal) crash frequency. The Pearson χ^2 statistic for the model is 1835, and the degrees of freedom are 2325 (= $n - p$ = [4 × 588] − 27). As this statistic is less than $\chi^2_{0.05, 2325}$ (= 2438), the hypothesis that the model fits the data cannot be rejected.

The t-statistics listed in the last column of Table 5-10 indicate a test of the hypothesis that the coefficient value is equal to 0.0. Those t-statistics with an absolute value that is larger than 2.0 indicate that the hypothesis can be rejected with the probability of error in this conclusion being less than 0.05. For those few variables where the absolute value of the t-statistic is smaller than 2.0, it was decided that the variable was essential to the model and was rationalized as being reasonable in magnitude (even if the specific value was not known with a great deal of certainty as applied to this database). Specifically, the t-statistic for three of the "added effect of rural area type" variables is less than 2.0. However, these variables are retained in the model because collectively, they indicate a logical safety difference between urban and rural freeways.

Multiple-Vehicle Non-Ramp Crashes

The results of the multiple-vehicle safety prediction model calibration are presented in Table 5-11. Calibration of this model focused on injury (plus fatal) crash frequency. The Pearson χ^2 statistic for the model is 614, and the degrees of freedom are 579 (= $n - p$ = 588 − 9). As this statistic is less than $\chi^2_{0.05, 579}$ (= 636), the hypothesis that the model fits the data cannot be rejected. The R^2 for the model is 0.76. An alternative measure of model fit that is better suited to the negative binomial distribution is R_k^2, as developed by Miaou (7). The R_k^2 for the calibrated model is 0.89. This statistic indicates that about 89 percent of the variability due to systematic sources is explained by the model.

Table 5-10. Freeway Model Statistical Description–Aggregate Model.

Model Statistics		Value		
Scale Parameter φ:		0.79		
Pearson χ^2:		1835 ($\chi^2_{0.05,\,2325}$ = 2438)		
Observations n_o:		588 segments (8436 injury + fatal crashes in 3 years)		
Calibrated Coefficient Values				
Variable	**Inferred Effect of...[1]**	**Value**	**Std. Dev.**	**t-statistic**
b_{tk}	Truck presence	-0.0100	0.0031	-3.2
b_{bar}	Rigid barrier presence	0.8903	0.1110	8.0
b_{enr}	Length of ramp entrance or weaving section	152.9	44.4	3.4
b_{mv4}	4 lane on multiple-vehicle crashes in urban areas	-5.237	0.334	-15.7
b_{mv6}	6 lane on multiple-vehicle crashes in urban areas	-5.648	0.386	-14.6
b_{mv8}	8 lane on multiple-vehicle crashes in urban areas	-5.846	0.425	-13.8
b_{mv10}	10 lane on multiple-vehicle crashes in urban areas	-6.118	0.501	-12.2
b_{mv}	ADT on multiple-vehicle crashes	1.550	0.088	17.6
$b_{mv,\,rural}$	Added effect of rural area type on multi-veh. crashes	-0.151	0.100	-1.5
b_{sv4}	4 lane on single-vehicle crashes in urban areas	-2.013	0.255	-7.9
b_{sv6}	6 lane on single-vehicle crashes in urban areas	-2.131	0.299	-7.1
b_{sv8}	8 lane on single-vehicle crashes in urban areas	-2.184	0.329	-6.6
b_{sv10}	10 lane on single-vehicle crashes in urban areas	-2.267	0.388	-5.8
b_{sv}	ADT on single-vehicle crashes	0.646	0.069	9.4
$b_{sv,\,rural}$	Added effect of rural area type on single-veh. crashes	-0.009	0.083	-0.1
b_{enr4}	4 lane on ramp-entrance crashes in urban areas	-4.956	0.511	-9.7
b_{enr6}	6 lane on ramp-entrance crashes in urban areas	-5.237	0.675	-7.8
b_{enr8}	8 lane on ramp-entrance crashes in urban areas	-6.221	0.818	-7.6
b_{enr10}	10 lane on ramp-entrance crashes in urban areas	-6.156	1.040	-5.9
b_{enr}	ADT on ramp-entrance crashes	1.329	0.355	3.7
$b_{enr,\,rural}$	Added effect of rural area type on ramp-ent. crashes	-0.449	0.510	-0.9
b_{exr4}	4 lane on ramp-exit crashes in urban areas	-6.356	0.783	-8.1
b_{exr6}	6 lane on ramp-exit crashes in urban areas	-7.354	1.057	-7.0
b_{exr8}	8 lane on ramp-exit crashes in urban areas	-7.637	1.229	-6.2
b_{exr10}	10 lane on ramp-exit crashes in urban areas	-7.620	1.529	-5.0
b_{exr}	ADT on ramp-exit crashes	1.685	0.530	3.2
$b_{exr,\,rural}$	Added effect of rural area type on ramp-exit crashes	1.255	0.558	2.2

Table 5-11. Prediction Model Statistical Description–Multiple-Vehicle Model.

Model Statistics	Value
R^2 (R_k^2):	0.76 (0.89)
Scale Parameter φ:	1.04
Pearson χ^2:	614 ($\chi^2_{0.05,\,579}$ = 636)
Over-Dispersion Parameter k:	4.40 mi^{-1}
Observations n_o:	588 segments (5018 injury + fatal crashes in 3 years)
Standard Deviation s_e:	±2.24 crashes/yr

The coefficients in Table 5-10 were combined with Equations 16 to 19 to obtain the calibrated safety prediction models for multiple-vehicle non-ramp crashes. The form of each model is:

$$E[N]_{mv4} = e^{-5.237 + 1.550\,Ln(ADT/1000) + Ln(L) - 0.151\,I_{rural}} \quad \textbf{(62)}$$

$$E[N]_{mv6} = e^{-5.648 + 1.550\,Ln(ADT/1000) + Ln(L) - 0.151\,I_{rural}} \quad \textbf{(63)}$$

$$E[N]_{mv8} = e^{-5.846 + 1.550\,Ln(ADT/1000) + Ln(L) - 0.151\,I_{rural}} \quad \textbf{(64)}$$

$$E[N]_{mv10} = e^{-6.118 + 1.550\,Ln(ADT/1000) + Ln(L) - 0.151\,I_{rural}} \quad \textbf{(65)}$$

The fit of the calibrated models is shown in Figure 5-3. This figure compares the predicted and reported crash frequency in the calibration database. The trend line shown represents a "$y = x$" line. A data point would lie on this line if its predicted and reported crash frequency were equal. The data points shown represent the observed multiple-vehicle non-ramp crash frequency for the segments used to calibrate the corresponding component model.

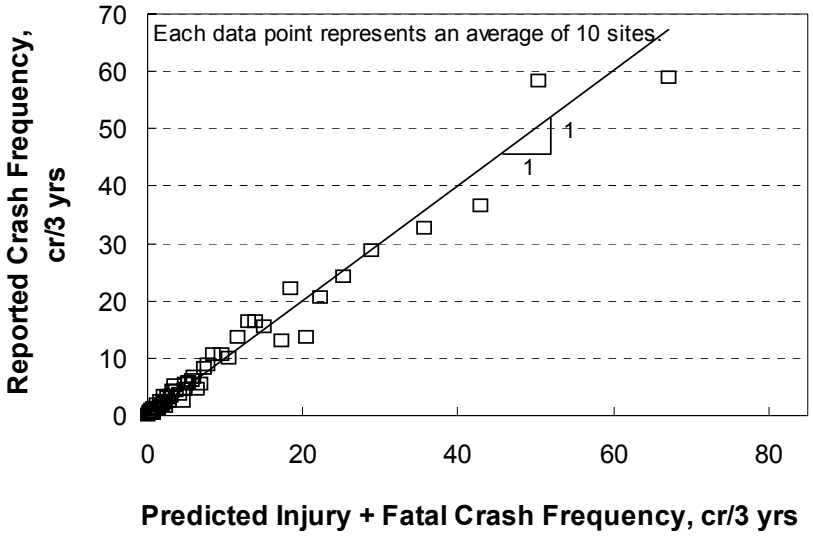


Figure 5-3. Predicted vs. Reported Multiple-Vehicle Crashes.

Each data point shown in Figure 5-3 represents the average predicted and average reported crash frequency for a group of 10 segments. The data were sorted by predicted crash frequency to form groups of segments with similar crash frequency. The purpose of this grouping was to reduce the number of data points shown in the figure and, thereby, to facilitate an examination of trends in the data. The individual segment observations were used for model calibration. In general, the data shown in the figure indicate that the model provides an unbiased estimate of expected crash frequency for segments experiencing up to 60 multiple-vehicle crashes in a three-year period.

Single-Vehicle Crashes

The results of the single-vehicle safety prediction model calibration are presented in Table 5-12. Calibration of this model also focused on injury (plus fatal) crash frequency. The Pearson χ^2 statistic for the model is 586, and the degrees of freedom are 579 (= $n - p$ = 588 – 9). As this statistic is less than $\chi^2_{0.05, 579}$ (= 636), the hypothesis that the model fits the data cannot be rejected. The R^2 for the model is 0.72, and the R_k^2 for the calibrated model is 0.90. This statistic indicates that about 90 percent of the variability due to systematic sources is explained by the model.

Table 5-12. Prediction Model Statistical Description–Single-Vehicle Model.

Model Statistics	Value
R^2 (R_k^2):	0.72 (0.90)
Scale Parameter φ:	1.00
Pearson χ^2:	586 ($\chi^2_{0.05, 579}$ = 636)
Over-Dispersion Parameter k:	9.05 mi^{-1}
Observations n_o:	588 segments (3278 injury + fatal crashes in 3 years)
Standard Deviation s_e:	±1.14 crashes/yr

The coefficients in Table 5-10 were combined with Equations 21 to 24 to obtain the calibrated safety prediction models for single-vehicle crashes. The form of each model is:

$$E[N]_{sv4} = e^{-2.013 + 0.646\,Ln(ADT/1000) + Ln(L) - 0.009\,I_{rural}} \quad \textbf{(66)}$$

$$E[N]_{sv6} = e^{-2.131 + 0.646\,Ln(ADT/1000) + Ln(L) - 0.009\,I_{rural}} \quad \textbf{(67)}$$

$$E[N]_{sv8} = e^{-2.184 + 0.646\,Ln(ADT/1000) + Ln(L) - 0.009\,I_{rural}} \quad \textbf{(68)}$$

$$E[N]_{sv10} = e^{-2.267 + 0.646\,Ln(ADT/1000) + Ln(L) - 0.009\,I_{rural}} \quad \textbf{(69)}$$

The fit of the calibrated models is shown in Figure 5-4. This figure compares the predicted and reported crash frequency in the calibration database. Each data point shown represents the average predicted and average reported crash frequency for a group of 10 segments. In general, the data shown in the figure indicate that the model provides an unbiased estimate of expected crash frequency for segments experiencing up to 30 single-vehicle crashes in a three-year period.

Ramp-Entrance-Related Crashes

The results of the ramp-entrance-related safety prediction model calibration are presented in Table 5-13. Calibration of this model also focused on injury (plus fatal) crash frequency. The Pearson χ^2 statistic for the model is 303, and the degrees of freedom are 579 (= $n - p$ = 588 – 9). As this statistic is less than $\chi^2_{0.05, 579}$ (= 636), the hypothesis that the model fits the data cannot be rejected. The R^2 for the model is 0.28, and the R_k^2 for the calibrated model is 0.95. This statistic indicates that about 95 percent of the variability due to systematic sources is explained by the model.

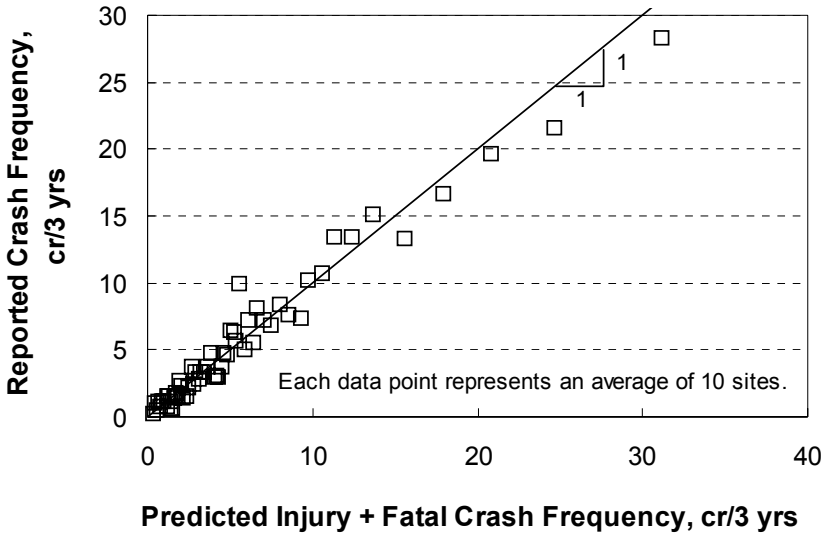


Figure 5-4. Predicted vs. Reported Single-Vehicle Crashes.

Table 5-13. Prediction Model Statistical Description–Ramp-Entrance-Related Model.

Model Statistics	Value
R^2 (R_k^2):	0.28 (0.95)
Scale Parameter φ:	0.52
Pearson χ^2:	303 ($\chi^2_{0.05,\,579}$ = 636)
Over-Dispersion Parameter k:	3.62
Observations n_o:	588 segments (93 injury + fatal crashes in 3 years)
Standard Deviation s_e:	±0.14 crashes/yr

The coefficients in Table 5-10 were combined with Equations 26 to 29 to obtain the calibrated safety prediction models for ramp-entrance-related crashes. The form of each model is:

$$E[N]_{enr4} = n_{enr}\, e^{-4.956 + 1.329\, Ln(ADT/15000) - 0.449\, I_{rural}} \quad \textbf{(70)}$$

$$E[N]_{enr6} = n_{enr}\, e^{-5.237 + 1.329\, Ln(ADT/15000) - 0.449\, I_{rural}} \quad \textbf{(71)}$$

$$E[N]_{enr8} = n_{enr}\, e^{-6.221 + 1.329\, Ln(ADT/15000) - 0.449\, I_{rural}} \quad \textbf{(72)}$$

$$E[N]_{enr10} = n_{enr}\, e^{-6.156 + 1.329\, Ln(ADT/15000) - 0.449\, I_{rural}} \quad \textbf{(73)}$$

The fit of the calibrated models is shown in Figure 5-5. This figure compares the predicted and reported crash frequency in the calibration database. Each data point shown represents the average predicted and average reported crash frequency for a group of 10 segments. In general, the data shown in the figure indicate that the model provides an unbiased estimate of expected crash frequency for segments experiencing up to 1.5 ramp-entrance-related crashes in a three-year period.

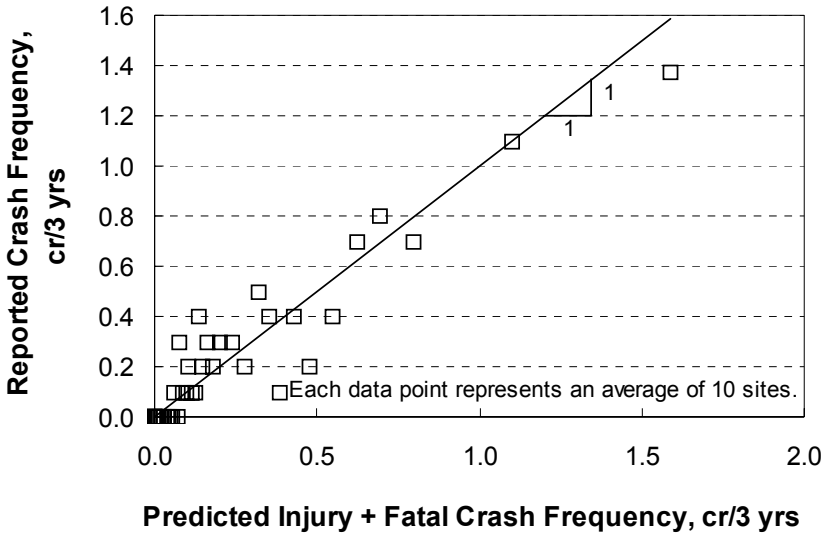


Figure 5-5. Predicted vs. Reported Ramp-Entrance-Related Crashes.

Ramp-Exit-Related Crashes

The results of the ramp-exit-related safety prediction model calibration are presented in Table 5-14. Calibration of this model also focused on injury (plus fatal) crash frequency. The Pearson χ^2 statistic for the model is 333, and the degrees of freedom are 579 (= $n - p$ = 588 – 9). As this statistic is less than $\chi^2_{0.05,\ 579}$ (= 636), the hypothesis that the model fits the data cannot be rejected. The R^2 for the model is 0.10, and the R_k^2 for the calibrated model is 0.78. This statistic indicates that about 78 percent of the variability due to systematic sources is explained by the model.

Table 5-14. Prediction Model Statistical Description–Ramp-Exit-Related Model.

Model Statistics	Value
R^2 (R_k^2):	0.10 (0.78)
Scale Parameter φ:	0.57
Pearson χ^2:	333 ($\chi^2_{0.05,\ 579}$ = 636)
Over-Dispersion Parameter k:	0.695
Observations n_o:	588 segments (47 injury + fatal crashes in 3 years)
Standard Deviation s_e:	±0.10 crashes/yr

The coefficients in Table 5-10 were combined with Equations 31 to 34 to obtain the calibrated safety prediction models for ramp-exit-related crashes. The form of each model is:

$$E[N]_{exr4} = n_{exr}\ e^{-6.356\ +\ 1.685\ Ln(ADT/15000)\ +\ 1.255\ I_{rural}} \quad \textbf{(74)}$$

$$E[N]_{exr6} = n_{exr}\ e^{-7.354\ +\ 1.685\ Ln(ADT/15000)\ +\ 1.255\ I_{rural}} \quad \textbf{(75)}$$

$$E[N]_{exr8} = n_{exr}\, e^{-7.637 + 1.685\, Ln(ADT/15000) + 1.255\, I_{rural}} \quad (76)$$

$$E[N]_{exr10} = n_{exr}\, e^{-7.620 + 1.685\, Ln(ADT/15000) + 1.255\, I_{rural}} \quad (77)$$

The fit of the calibrated models is shown in Figure 5-6. This figure compares the predicted and reported crash frequency in the calibration database. Each data point shown represents the average predicted and average reported crash frequency for a group of 10 segments. In general, the data shown in the figure indicate that the model provides an unbiased estimate of expected crash frequency for segments experiencing up to 0.7 ramp-exit-related crashes in a three-year period.

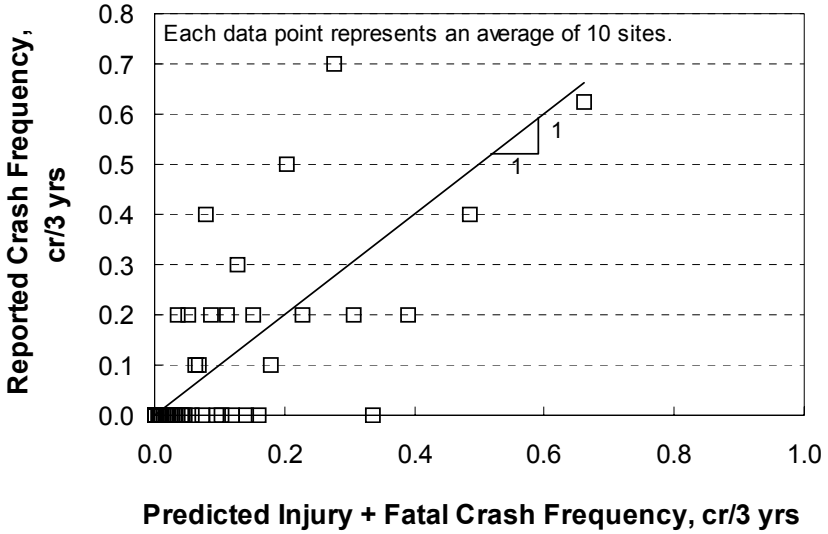


Figure 5-6. Predicted vs. Reported Ramp-Exit-Related Crashes.

Calibrated AMFs

Several AMFs were calibrated in conjunction with the safety prediction model calibration. All of them were calibrated using injury (plus fatal) crash data. Collectively, they describe the relationship between barrier presence, inside clearance distance, outside clearance distance, ramp entrance length, weaving section length, and truck presence on crash frequency.

Rigid Barrier AMF. The calibrated rigid barrier AMF has two forms, depending on whether it is used to evaluate the presence of an inside or outside barrier. The AMF for an inside barrier is:

$$AMF_{b|ib} = e^{0.890/W_{icb}} \quad (78)$$

The AMF for an outside barrier is:

$$AMF_{b|ob} = e^{0.890/W_{ocb}} \quad (79)$$

The two equations above are effectively the same. The width term in the first equation is the inside clearance distance. The width term in the second equation is the outside clearance distance. Equations for computing these distances are provided in the previous section.

The rigid barrier AMF is shown in Figure 5-7. It is limited to clearance distances of 2 ft or more. This AMF is used with the aggregated inside clearance AMF in Equation 35 or the aggregated outside clearance AMF in Equation 36 to estimate the combined effect of clearance distance and barrier presence. The trends in Figure 5-7 indicate that the presence of a barrier within 2 to 4 ft of the shoulder tends to be associated with a 25 to 50 percent increase in crash frequency. In contrast, the presence of a barrier 30 ft from the shoulder is associated with only a 3 percent increase.

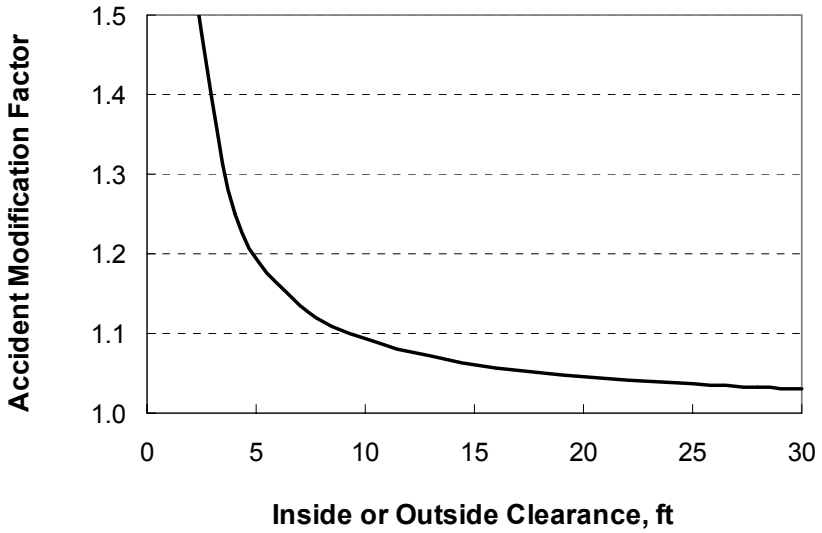


Figure 5-7. Calibrated Rigid Barrier AMF.

Inside Clearance AMF. The inside clearance AMF has two forms, depending on whether it is used to evaluate inside clearance in the presence of a rigid barrier. The AMF when an inside barrier is not present is:

$$AMF_{ic|is} = e^{-0.0296\,([W_m - 2W_{is}]^{0.5} - [W_{mb} - 2W_{isb}]^{0.5})} \quad \textbf{(80)}$$

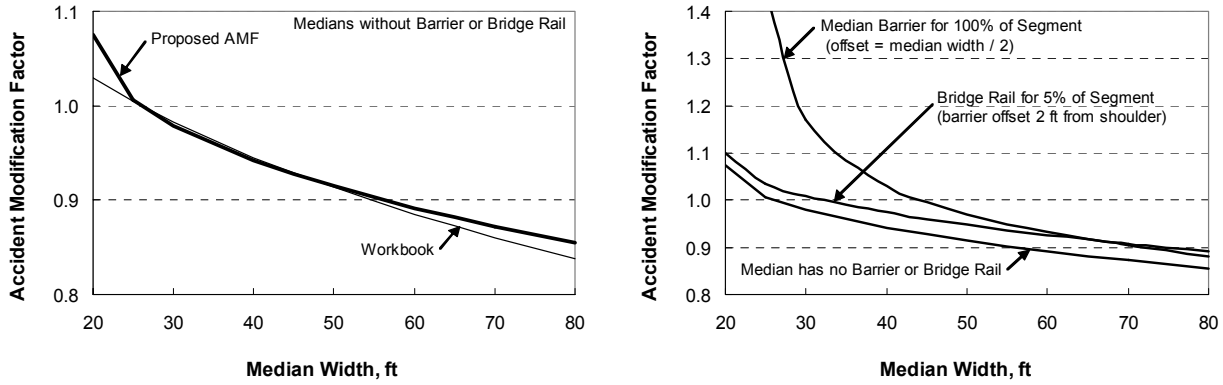
The AMF when an inside barrier is present is:

$$AMF_{ic|ib} = e^{-0.0296\,([2W_{icb}]^{0.5} - [W_{mb} - 2W_{isb}]^{0.5})} \quad \textbf{(81)}$$

The base median width and the base inside shoulder width values are provided in Table 5-9. The calibration coefficient of "-0.0296" was obtained from a similar model developed for rural multilane highways (see Chapter 4). This coefficient could not be reliably calibrated with the freeway segment database because of the significant correlation between barrier presence and median width that existed in the database.

Equations 80 and 81 are used with Equation 35 to estimate the aggregated inside clearance AMF for segments that have a mixture of barriers and no barriers in the median area.

The calibrated inside clearance AMF when an inside barrier is not present (i.e., Equation 80) is shown in Figure 5-8a. A base median width of 26 ft and a base inside shoulder width of 10 ft are used for this evaluation (they correspond to a positive barrier median and six or more through lanes, as shown in Table 5-9). Also shown is the median width AMF from the *Workbook* (*1*). The trends in Figure 5-8a indicate that wider medians tend to be associated with a reduction in crash frequency. Similar agreement was found when comparing Equation 80 with the *Workbook* AMF for base median widths of 36 and 56 ft.



a. No Rigid Barrier Present. **b. Aggregate AMF.**

Figure 5-8. Calibrated Inside Clearance AMF.

Figure 5-8b illustrates how Equation 35 can be used to combine the inside clearance and rigid barrier AMFs to the evaluation of a segment with a mixture of barriers and no barriers. The lower trend line in Figure 5-8b is the same as that shown in Figure 5-8a and is used for reference. A second trend line in Figure 5-8b applies when a bridge-related, inside barrier is present for 5 percent of the segment length. This barrier is offset 2 ft from the edge of the inside shoulder, regardless of the width of the median. The bridge barrier is shown to be associated with a 3 percent increase in crash frequency on the subject segment (relative to the "no barrier" condition).

The upper trend line in Figure 5-8b illustrates the effect of a continuous rigid barrier in the middle of the median. The effect is shown to be highly sensitive to the width of the median (and thus, the distance between the barrier and the traveled way). Figure 5-8b suggests that the installation of a median barrier in a 30 ft depressed median is likely to be associated with a 19 percent increase in crashes; however, it will also significantly reduce the number of fatal crashes (*8*). Crash data reported by Miaou et al. (*8*) indicate that median barriers increase injury (plus fatal) crashes by about 16 percent, but reduce fatal crashes by over 70 percent. The use of Equation 35 to compare medians "with" and "without" a barrier (for the same average median widths as reported

by Miaou et al.) indicates a slightly more conservative crash increase of 13 percent associated with the installation of a median barrier.

Outside Clearance AMF. The calibrated outside clearance AMF has two forms, depending on whether it is used to evaluate inside clearance in the presence of a rigid barrier. The AMF when an outside barrier is not present is:

$$AMF_{oc|os} = (e^{-0.014([W_{hc} - W_s] - 20)} - 1.0) P_{hc} + 1.0 \quad \geq 1.0 \qquad (82)$$

The AMF when an outside barrier is present is:

$$AMF_{oc|ob} = (e^{-0.014(W_{ocb} - 20)} - 1.0) P_{hc} + 1.0 \quad \geq 1.0 \qquad (83)$$

A base clearance distance of 20 ft is used in Equations 82 and 83. It corresponds to a base horizontal clearance of 30 ft and a base outside shoulder width of 10 ft.

Equations 82 and 83 are used with Equation 36 to estimate the aggregated outside clearance AMF for segments that have a mixture of barriers and no barriers on the roadside.

The calibrated outside clearance AMF (when an outside barrier is not present, i.e., Equation 82) is shown as the lower trend line in Figure 5-9. This trend line is identical to that from the horizontal clearance AMF from the *Workbook* (*1*).

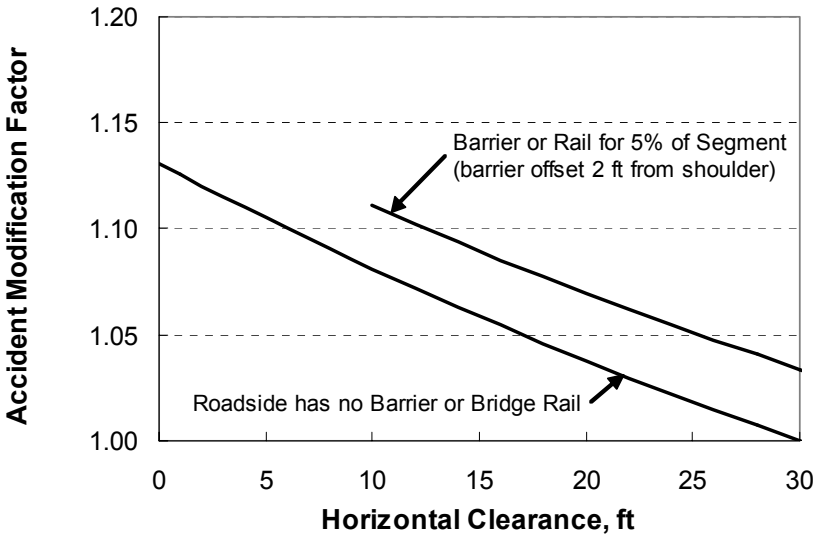


Figure 5-9. Calibrated Outside Clearance AMF.

The upper trend line in Figure 5-9 illustrates how Equation 36 can be used to combine the outside clearance and rigid barrier AMFs to the evaluation of a segment with a mixture of barriers and no barriers. This trend line applies when a bridge-related, outside barrier is present for 5 percent

of the segment length. This barrier is offset 2 ft from the edge of the outside shoulder. The bridge barrier is shown to be associated with a 5 percent increase in crash frequency on the subject segment (relative to the "no barrier" condition).

Ramp Entrance AMF. The calibrated ramp entrance AMF is:

$$AMF_{enr} = e^{152.9/L^*_{enr}} \quad \textbf{(84)}$$

The average ramp entrance length L^*_{enr} is used in this equation to estimate the AMF value. Equation 58 is provided for computing this length.

The ramp entrance AMF is shown in Figure 5-10. It is limited to entrance lengths of 100 ft or more. This AMF is used with the aggregated ramp entrance AMF in Equation 37 to estimate the ramp entrance effect on freeway segments. The trends in Figure 5-10 indicate that shorter ramp entrances tend to be associated with more frequent crashes. The rate of increase is particularly notable for ramp entrances that are less than 300 ft in length.

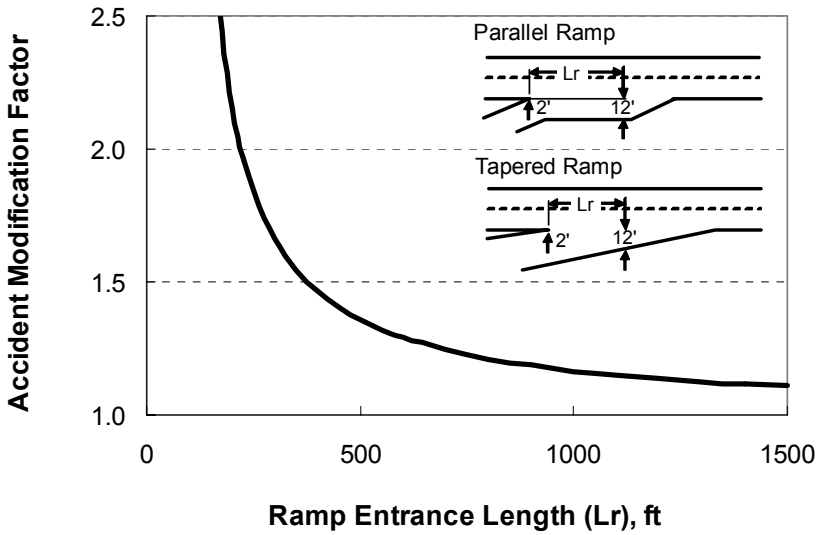


Figure 5-10. Calibrated Ramp Entrance AMF.

Weaving Section AMF. The calibrated weaving section AMF is:

$$AMF_{wev} = e^{152.9/L^*_{wev}} \quad \textbf{(85)}$$

The average weaving section length L^*_{wev} is used in this equation to estimate the AMF value. Equation 60 is provided for computing this length.

The weaving section AMF is shown in Figure 5-11. It is limited to weaving section lengths of 800 ft or more. This AMF is used with the aggregated weaving section AMF in Equation 38 to

estimate the weaving section effect on freeway segments. The trends in Figure 5-11 indicate that shorter weaving sections tend to be associated with more frequent crashes.

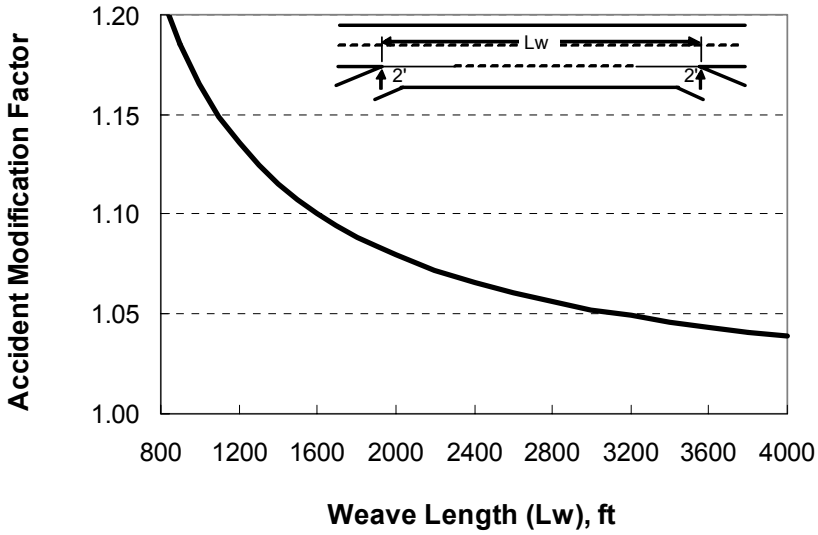


Figure 5-11. Calibrated Weaving Section AMF.

Truck Presence AMF. The calibrated truck presence AMF is:

$$AMF_{tk} = e^{-0.010\,(P_t - 20)} \quad \textbf{(86)}$$

This AMF is shown in Figure 5-12 using a thick trend line. It is limited to truck percentages in the range of 0 to 30 percent. The truck presence AMF described in Chapter 4 for rural multilane highways is also shown in the figure and indicates a similar trend.

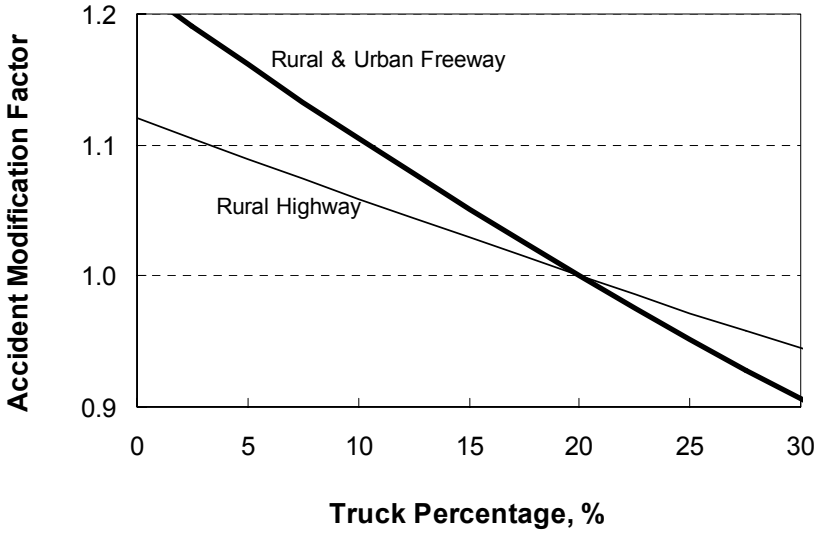
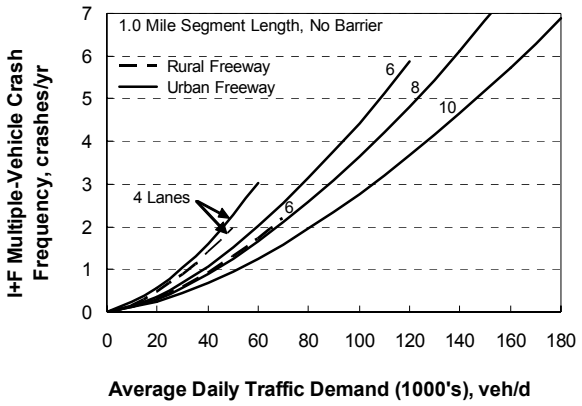


Figure 5-12. Calibrated Truck Presence AMF.

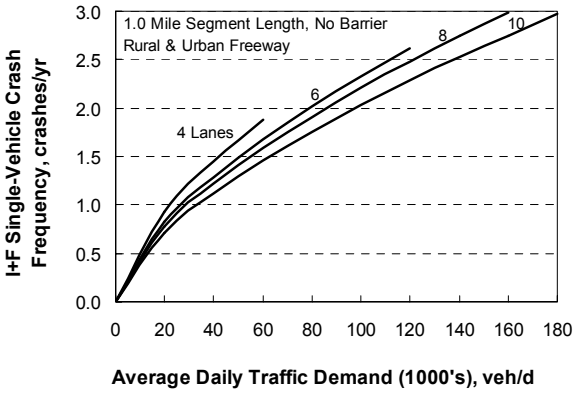
The trend in Figure 5-12 indicates that freeways with higher percentages of trucks are associated with fewer crashes. It may suggest that drivers are more cautious when there are many trucks present in the traffic stream. It may also suggest that freeways with higher truck percentages are more likely to have design dimensions that are "generous" (relative to the design control limits) and thus provide some additional safety benefit.

Sensitivity Analysis

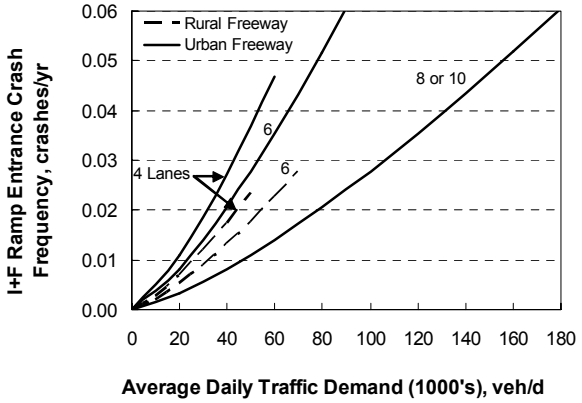
The relationship between crash frequency and traffic demand, as obtained from calibrated Equation 13, is illustrated in Figure 5-13 for a 1-mile freeway segment with two ramp entrances, two ramp exits, and no barrier. The individual component models are illustrated in Figures 5-13a, 5-13b, 5-13c, and 5-13d. The sum of the individual component crash frequencies is illustrated in Figure 5-14. The trends in this figure are comparable to those in Figure 5-1. The length of the trend lines in Figures 5-13 and 5-14 reflect the range of traffic demand in the data.



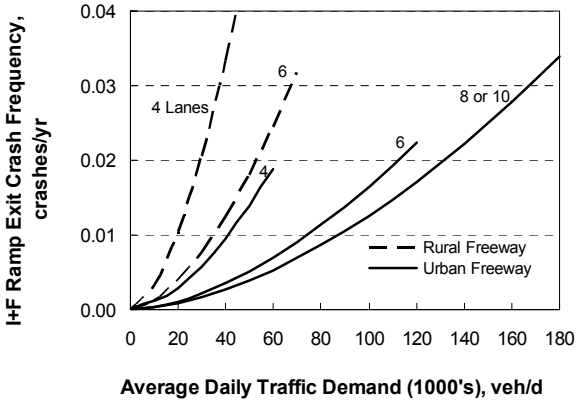
a. Multiple-Vehicle Crashes.



b. Single-Vehicle Crashes.



c. Ramp-Entrance-Related Crashes.



d. Ramp-Exit Related Crashes.

Figure 5-13. Proposed Crash Prediction Model Components.

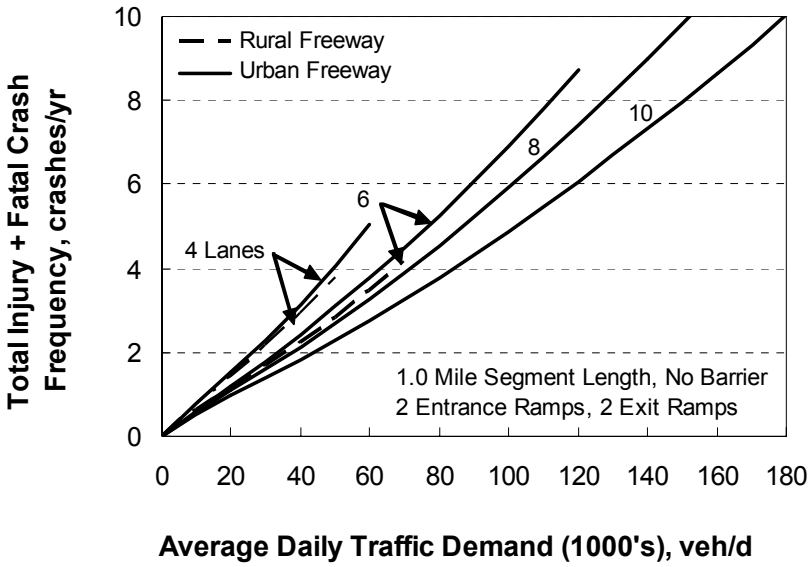


Figure 5-14. Proposed Crash Prediction Model.

The trend lines shown in Figure 5-14 indicate that urban freeways have about 7 percent more crashes than rural freeways. This finding is in contrast to the 70 percent increase noted in the *Workbook*, as shown in Figure 5-1. It is likely that the trends in Figure 5-1 reflect the influence of barrier length, ramp entrances, ramp exits, and weaving section length. As shown in Table 5-7, these influences are more prevalent on urban freeway segments. In contrast, these influences have been explicitly quantified in the proposed model such that they do not influence the trends shown in Figure 5-14. The proposed model is a more accurate illustration of differences between freeway segments in rural vs. urban areas, when the segments have the same barrier proportion, ramp entrance length, and weaving section length.

The trend lines shown in Figure 5-14 also indicate that crash frequency is *lower* on freeways with many lanes that it is on freeways with few lanes. This trend is similar to that found for urban and suburban arterial streets, but in contrast to that reported in the *Workbook* (as shown in Figure 5-1). For example, the trends in Figure 5-1 indicate that six-lane freeway segments have about 23 percent fewer crashes than a four-lane segment, while the trends in Figure 5-1 indicate that six-lane segments have 50 percent more crashes than a four-lane segment. As shown in Table 5-7, it is likely that the trends in Figure 5-1 reflect the fact that the proportion of barrier along a freeway segment typically increases (and the lateral clearance decreases) with an increase in the number of lanes. The proposed model accounts for the influence of barrier presence and lateral clearance. Therefore, it is a more accurate estimator of the relationship between number-of-lanes and freeway crash frequency.

At an ADT of 40,000 veh/d, the trends in Figure 5-14 indicate that a rural four-lane freeway has an expected crash frequency of 2.9 crashes/yr. In contrast, an urban four-lane freeway has an expected crash frequency of 3.1 crashes/yr. A comparison of the trends in Figures 5-1 and 5-14 indicates that the trends for four-lane segments are generally similar, but that those in Figure 5-1 for

six or more lanes indicate many more crashes (for the same ADT) than those in Figure 5-14. The difference between these trend lines diminishes when barrier length, ramp entrances, ramp exits, and weaving section lengths are increased in the proposed models to levels consistent with six- and eight-lane freeway segments.

Model Extensions

The proposed safety prediction model was calibrated using a limited amount of data for rural eight-lane freeway segments. Moreover, no data were available for rural ten-lane segments. As suggested by the urban segment data, the safety experience of the eight- and ten-lane cross sections is not believed to be similar to that of four- or six-lane cross sections. Thus, the calibrated model should *not* be applied to rural freeways with either eight or ten lanes.

REFERENCES

1. Bonneson, J.A., K. Zimmerman, and K. Fitzpatrick. *Interim Roadway Safety Design Workbook*. FHWA/TX-06/0-4703-P4, Texas Department of Transportation, Austin, Texas, April 2006.

2. Bonneson, J.A., K. Zimmerman, and K. Fitzpatrick. *Roadway Safety Design Synthesis.* FHWA/TX-05/0-4703-P1. Texas Department of Transportation, Austin, Texas, November 2005.

3. Harwood, D., K. Bauer, K. Richard, D. Gilmore, J. Graham, I. Potts, D. Torbic, and E. Hauer. *Methodology to Predict the Safety Performance of Urban and Suburban Arterials.* Final Report. NCHRP Project 17-26. National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C., March 2007.

4. Lord, D. "Modeling Motor Vehicle Crashes Using Poisson-Gamma Models: Examining the Effects of Low Sample Mean Values and Small Sample Size on the Estimation of the Fixed Dispersion Parameter." *Accident Analysis & Prevention*, Vol. 38, No. 4, Elsevier Ltd., Oxford, Great Britain, July 2006, pp. 751-766.

5. Bonneson, J., D. Lord, K. Zimmerman, K. Fitzpatrick, and M. Pratt. *Development of Tools for Evaluating the Safety Implications of Highway Design Decisions.* FHWA/TX-07/0-4703-4. Texas Department of Transportation, Austin, Texas, September 2006.

6. *SAS OnlineDoc, Version 8*, SAS Institute, Inc., Cary, North Carolina, 1999.

7. Miaou, S.P. *Measuring the Goodness-of-Fit of Accident Prediction Models.* FHWA-RD-96-040. Federal Highway Administration, Washington, D.C., 1996.

8. Miaou, S-P., R.P. Bligh, D. Lord. "Developing Guidelines for Median Barrier Installation." *Transportation Research Record 1904.* Transportation Research Board, Washington, D.C., 2005, pp. 3-19.