Appendix B—Specialized Procedures Common to Chapters 18 and 19

This appendix describes two specialized procedures intended for use with the predictive method presented in Chapters 18 and 19. The first procedure is used to calibrate the predictive models in Chapters 18 and 19 to local conditions. The second procedure is the empirical Bayes (EB) Method for combining observed crash frequencies with the estimate provided by the predictive models in Chapters 18 and 19. Both of these procedures are an integral part of the predictive method in Chapters 18 and 19, and are presented in this appendix only to avoid repetition across the chapters.

B.1. CALIBRATION OF THE CHAPTER 18 AND 19 PREDICTIVE MODELS

The predictive models in each of Chapters 18 and 19 consist of safety performance functions (SPFs), crash modification factors (CMFs), and a calibration factor. Each model is developed for a specific site type (i.e., segment or intersection).

The predictive models were developed from the most complete and consistent data sets available. However, the general level of crash frequencies may vary substantially from one jurisdiction to another for a variety of reasons including climate, driver populations, animal populations, crash reporting thresholds, and crash reporting system procedures. Therefore, for these predictive models to provide results that are meaningful and accurate for each jurisdiction, it is important that they be calibrated for application in the jurisdiction in which they are applied. A procedure for determining the calibration factors for the predictive models is presented in Section B.1.1.

Some HSM users may prefer to develop SPFs with data from their jurisdiction for use in the predictive models rather than calibrating the existing SPFs. Calibration of the SPFs in Chapters 18 and 19 will provide satisfactory results. However, SPFs developed directly with data for a specific jurisdiction may provide more reliable estimates for that jurisdiction than calibration of the existing SPFs. Guidance is presented in Section B.1.2 on the development of jurisdiction-specific SPFs that are suitable for use with the predictive models in Chapters 18 and 19.

The predictive method in each of Chapters 18 and 19 consists of a set of predictive models, default distributions, and severity distribution functions (SDFs). Most of the regression coefficients and distribution values used in the predictive methods have been determined through research. Therefore, modification of the regression coefficients is not recommended. However, a few specific quantities, such as the distribution of crashes by collision type, can vary substantially from jurisdiction to jurisdiction. Where local data are available, users are encouraged to replace these default values with locally derived values. A procedure for deriving jurisdiction-specific distribution values is presented in Section B.1.3. A procedure for calibrating the SDFs is described in Section B.1.4.

B.1.1. Calibration of Predictive Models

The calibration procedure is used to derive the value of the calibration factor that is included in each predictive model. A calibration factor represents the ratio of the total number of observed crashes for a selected set of sites to the total number of predicted crashes for the same sites, during the same time period, using the applicable predictive model. Thus, the nominal value of the calibration factor is 1.00 when the observed and predicted number of crashes happen to be equal. When there are more crashes observed than are predicted by the predictive model, the computed calibration factor will be greater than 1.00. When there are fewer crashes observed than are predicted by the predictive model, the computed calibration factor will be less than 1.00.

It is recommended that new values of the calibration factors be derived at least every two to three years, and some HSM users may prefer to develop calibration factors on an annual basis. The calibration factor for the most recent available period is to be used for all assessment of proposed future projects. If available, calibration factors for the specific time periods included in the evaluation period are to be used in effectiveness evaluations that use the procedures presented in Chapter 9.

If the procedure in Section B.1.3 is used to calibrate a default distribution, then the locally-calibrated values should be used in the calibration process described in this section.

The calibration procedure involves five steps. Each step is described in the following five subsections.

B.1.1.1. Step 1—Identify the predictive models to be calibrated.

Calibration is performed separately for each predictive model described in Chapters 18 and 19. Table B-1 identifies the combinations of site type and cross section or control type represented in each predictive model and for which calibration factors can be derived. The models of interest are identified by the user in this step.

Table B-1. Predictive Models in Chapters 18 and 19 that Need Calibration

Site Type and Cross Section or Control Type	Calibration Factor: Symbol	Calibration Factor: Equation Number
ROADWAY SEGMENTS		
Freeways		
Multiple-vehicle fatal-and-injury crashes, all cross sections	$C_{fs,\,ac,\,mv,\,fi}$	18-3
Multiple-vehicle property-damage-only crashes, all cross sections	$C_{fs,\,ac,\,mv,\,pdo}$	18-5
Single-vehicle fatal-and-injury crashes, all cross sections	$C_{fs,\,ac,\,sv,\,fi}$	18-4
Single-vehicle property-damage-only crashes, all cross sections	$C_{fs,\,ac,\,sv,\,pdo}$	18-6
Ramps		
Entrance ramp, multiple-vehicle fatal-and-injury crashes, all lanes	$C_{rps,\,EN,\,mv,\,fi}$	19-3
Entrance ramp, multiple-vehicle property-damage-only crashes, all lanes	$C_{rps,\,EN,\,mv,\,pdo}$	19-5
Entrance ramp, single-vehicle fatal-and-injury crashes, all lanes	$C_{rps,\,EN,\,sv,\,fi}$	19-4
Entrance ramp, single-vehicle property-damage-only crashes, all lanes	$C_{rps,\,EN,\,sv,\,pdo}$	19-6
Exit ramp, multiple-vehicle fatal-and-injury crashes, all lanes	$C_{rps,\,EX,\,mv,\,fi}$	not shown
Exit ramp, multiple-vehicle property-damage-only crashes, all lanes	$C_{rps,\,EX,\,mv,\,pdo}$	not shown
Exit ramp, single-vehicle fatal-and-injury crashes, all lanes	$C_{rps,\,EX,\,sv,\,fi}$	not shown
Exit ramp, single-vehicle property-damage-only crashes, all lanes	$C_{rps,\,EX,\,sv,\,pdo}$	not shown
C-D road, multiple-vehicle fatal-and-injury crashes, all cross sections	$C_{cds,\,ac,\,mv,\,fi}$	19-8
C-D road, multiple-vehicle property-damage-only crashes, all cross sections	$C_{cds,\,ac,\,mv,\,pdo}$	19-10
C-D road, single-vehicle fatal-and-injury crashes, all cross sections	$C_{cds,\,ac,\,sv,\,fi}$	19-9
C-D road, single-vehicle property-damage-only crashes, all cross sections	$C_{cds,\,ac,\,sv,\,pdo}$	19-11
INTERSECTIONS		
Freeway Speed-Change Lanes		
Ramp entrance speed-change lane, fatal-and-injury crashes of all types	$C_{sc,\,EN,\,at,\,fi}$	18-8
Ramp entrance speed-change lane, property-damage-only crashes of all types	$C_{sc,\,EN,\,at,\,pdo}$	18-9
Ramp exit speed-change lane, fatal-and-injury crashes of all types	$C_{sc,\,EX,\,at,\,fi}$	18-11
Ramp exit speed-change lane, property-damage-only crashes of all types	$C_{sc,\,EX,\,at,\,pdo}$	18-12
Crossroad Ramp Terminals		
One-way stop control, fatal-and-injury crashes of all types	$C_{aS,\,ST,\,at,\,fi}$	19-13
One-way stop control, property-damage-only crashes of all types	$C_{aS,\,ST,\,at,\,pdo}$	19-14
Signal control, fatal-and-injury crashes of all types	$C_{aS,\,SG,\,at,\,fi}$	19-16
Signal control, property-damage-only crashes of all types	$C_{aS,\,SG,\,at,\,pdo}$	19-17

Also established in this step is the calibration period. A calibration period longer than three years is not recommended because the expected average crash frequency is likely to change over time. The calibration period should have a duration that is a multiple of 12 months to avoid seasonal effects. For ease of application, it is recommended that the calibration periods consist of one, two, or three full calendar years. It is recommended to use the same calibration period for all sites, but exceptions may be made where necessary.

B.1.1.2. Step 2—Select sites for calibration of the predictive model.

Calibration sites are selected during this step. One set of calibration sites is assembled for each predictive model identified in Step 1. A given site may be included in more than one set *provided* that the site in question is consistent with the model's calibration factor characteristics (as listed in Table B-1). It is desirable that these sites be reasonably representative of the range of site characteristics to which the predictive model will be applied. However, no formal stratification by traffic volume or other site characteristics is needed in selecting the calibration sites. As such, the sites can be selected in a manner to make the data collection needed for Step 3 as efficient as practical.

Each calibration site should be selected without regard to the number of crashes reported during the calibration period. In other words, calibration sites should not be selected to intentionally limit the calibration database to include only sites with either high or low crash frequencies. Where practical, this may be accomplished by selecting calibration sites randomly from a larger set of candidate sites.

The desirable minimum sample size for the calibration database for one predictive model is 30 to 50 sites. For segments, each site should be between 0.1 and 1.0 mi in length. Lengths in this range should be long enough to have statistical validity and short enough to be realistically homogeneous.

For large jurisdictions, such as entire states, with a variety of topographical and climate conditions, it may be desirable to assemble a separate set of calibration sites representing two or three different conditions. In this manner, separate calibration factors are developed for each specific terrain type or geographical region for a given predictive model. For example, a state with distinct plains and mountain regions (or with distinct dry and wet regions), might choose to develop separate calibration factors for those regions. Where separate calibration factors are developed by terrain type or region, this needs to be done consistently for all predictive models applicable to those regions.

B.1.1.3. Step 3—Obtain data for each set of calibration sites for the calibration period.

This step is repeated for each predictive model identified in Step 1 and its associated set of calibration sites assembled in Step 2. For this step, a calibration database is assembled for each set of calibration sites. The calibration data are assembled for a common calibration period for all sites. The calibration database should include the following information for each site represented in the database:

- all target crashes that are reported during the calibration period, and
- site characteristics data needed to apply the predictive model for the same calibration period.

Target crashes are those crashes that are consistent with the predictive model being calibrated. For example, if the predictive model is applicable to multiple-vehicle fatal-and-injury crashes on freeway segments, then the target crashes are multiple-vehicle fatal-and-injury crashes on freeway segments.

For a given site type, the calibration database should include at least 100 target crashes per year. If this minimum is not realized, then additional sites should be added to the database following the guidelines in Step 2.

The crash data used for calibration should include all crashes related to each site selected for the calibration database. Crashes should be assigned to specific sites based on the guidelines presented in Section B.2.3.

Table B-2 identifies the site characteristics data that are needed to apply the predictive models. The table classifies each data element as either required or desirable for the calibration procedure. Data for each of the required elements are needed for calibration. For the desirable data elements, it is recommended that actual data be used if available. Assumptions are offered in the table when these data are not available.

Table B-2. Data Needs for Calibration of Chapter 18 and 19 Predictive Models

Chapter	Data Element	Data Need: Required	Data Need: Desirable	Default Assumption
ROADWAY SEGMENTS				
18—Freeways	Area type (rural or urban)	X		Need actual data
	Number of through lanes	X		Need actual data
	Segment length	X		Need actual data
	Length and radii of horizontal curves	X		Need actual data
	Lane width	X		Need actual data
	Inside and outside shoulder width (paved)	X		Need actual data
	Median width	X		Need actual data
	Length of rumble strips on inside and outside shoulders		X	Base default on agency policy
	Length of (and offset to) median barrier	X		Need actual data
	Length of (and offset to) outside barrier	X		Need actual data
	Clear zone width		X	Base default on agency policy
	AADT volume of (and distance to) nearest upstream entrance ramp	X		Need actual data
	AADT volume of (and distance to) nearest downstream exit ramp	X		Need actual data
	Presence of speed-change lane	X		Need actual data
	Presence and length of Type B weaving sections	X		Need actual data
	Proportion of AADT that occurs during hours where lane volume exceeds 1,000 veh/h/ln		X	Equation for computing default is near the end of Section 18.4.2
	Average annual daily traffic (AADT) volume	X		Need actual data

Table B-2. Data Needs for Calibration of Chapter 18 and 19 Predictive Models (continued)

Chapter	Data Element	Data Need		Default Assumption
		Required	Desirable	
ROADWAY SEGMENTS				
19—Ramps	**For ramps and collector-distributor (C-D) roads:**			
	Area type (rural or urban)	X		Need actual data
	Number of through lanes	X		Need actual data
	Segment length	X		Need actual data
	Average annual daily traffic (AADT) volume	X		Need actual data
	Length and radii of horizontal curves	X		Need actual data
	Lane width	X		Need actual data
	Left and right shoulder width (paved)	X		Need actual data
	Length of (and offset to) right side barrier	X		Need actual data
	Length of (and offset to) left side barrier	X		Need actual data
	Presence of lane add or drop		X	Assume not present
	Presence of speed-change lane	X		Need actual data
	For C-D roads only:			
	Presence and length of weaving section	X		Need actual data
ROADWAY SEGMENTS				
18—Freeways	**For freeway speed-change lanes:**			
	Area type (rural or urban)	X		Need actual data
	Number of through lanes	X		Need actual data
	Segment length	X		Need actual data
	Length and radii of horizontal curves	X		Need actual data
	Lane width	X		Need actual data
	Inside shoulder width (paved)	X		Need actual data
	Median width	X		Need actual data
	Presence of rumble strips on inside shoulder		X	Base default on agency policy
	Length of (and offset to) median barrier	X		Need actual data
	AADT volume of ramp in speed-change lane	X		Need actual data
	Presence and length of Type B weaving sections	X		Need actual data
	Proportion of AADT that occurs during hours where lane volume exceeds 1,000 veh/h/ln		X	Equation for computing default is near the end of Section 18.4.2
	AADT of freeway adjacent to speed-change lane	X		Need actual data

Table B-2. Data Needs for Calibration of Chapter 18 and 19 Predictive Models (continued)

Chapter	Data Element	Data Need: Required	Data Need: Desirable	Default Assumption
INTERSECTIONS				
	For all crossroad ramp terminals:			
	Area type (rural or urban)	X		Need actual data
	Ramp terminal configuration	X		Need actual data
	Type of traffic control	X		Need actual data
	Control for exit ramp right-turn movement	X		Need actual data
	AADT for inside and outside crossroad legs	X		Need actual data
	AADT volume for each ramp leg	X		Need actual data
	Number of through lanes on each crossroad approach	X		Need actual data
	Number of lanes on the exit ramp	X		Need actual data
	Number of crossroad approaches with left-turn lanes	X		Need actual data
	Number of crossroad approaches with right-turn lanes	X		Need actual data
	Number of unsignalized public street approaches to the crossroad leg outside of the interchange		X	Assume no public street approaches present
19—Ramps	Distance to next public street intersection		X	Assume 0.15 mi for urban areas, assume 0.20 mi for rural areas
	Distance to adjacent crossroad ramp terminal		X	Based default on terminal configuration and area type[a]
	Crossroad median width and left-turn lane width	X		Need actual data
	For signal-controlled crossroad ramp terminals only:			
	Number of unsignalized driveways on the crossroad leg outside of the interchange		X	Assume no driveways present
	Number of crossroad approaches with protected-only left-turn operation	X		Need actual data
	Number of crossroad approaches with right-turn channelization	X		Need actual data
	Presence of exit ramp right-turn channelization	X		Need actual data
	Presence of a non-ramp public street leg		X	Assume leg not present
	For one-way stop-controlled crossroad ramp terminals only:			
	Skew angle	X		Need actual data

[a] Default values by crossroad ramp terminal configuration and area type. Urban areas: *A2* = 0.17 mi, *A4* = 0.17 mi, *B2* = 0.19 mi, *B4* = 0.19 mi, *D3* = 0.13 mi, *D4* = 0.11 mi. Rural areas: *A2* = 0.20 mi, *A4* = 0.20 mi, *B2* = 0.22 mi, *B4* = 0.22 mi, *D3* = 0.16 mi, *D4* = 0.17 mi. Crossroad ramp terminal configurations are shown in Figure 19-1.

If data for some required elements are not readily available, it may be possible to select sites in Step 2 for which these data are available. For example, in calibrating the predictive models for freeway segments, if data on the radii of horizontal curves are not readily available, the calibration data set could be limited to tangent freeways. Decisions of this type should be made, as needed, to keep the effort required to assemble the calibration data set within reasonable bounds.

B.1.1.4. Step 4—Apply the applicable predictive method to estimate the predicted average crash frequency for each site during the calibration period as a whole.

This step is repeated for each predictive model identified in Step 1 and its associated set of calibration sites assembled in Step 2. The site characteristics data assembled in Step 3 are used to apply the applicable predictive method to each site in the set of calibration sites. For this application, the predictive model should be applied without using the EB Method and without employing a calibration factor (i.e., a calibration factor of 1.00 is assumed). Through this process, the predicted average crash frequency is obtained for each site in the set of calibration sites and for each year in the calibration period.

B.1.1.5. Step 5—Compute calibration factors for use in the predictive models.

The final step is to compute the calibration factor using the following equation. The appropriate subscripts for this equation are identified in Table B-1 for each predictive model.

$$C_{w,x,y,z} = \frac{\sum_{i=1}^{n_{sites}} \sum_{j=1}^{n_c} N_{o,w(i),x(i),y,z,j}}{\sum_{i=1}^{n_{sites}} \sum_{j=1}^{n_c} N_{p,w(i),x(i),y,z,j}} \tag{B-1}$$

Where:

$C_{w,x\ y,z}$ = calibration factor to adjust SPF for local conditions for site type *w*, cross section or control type *x*, crash type *y*, and severity *z*;

n_{sites} = number of sites;

n_c = number of years in the crash period (yr);

$N_{o,\,w(i),\,x(i),\,y,\,z,\,j}$ = observed crash frequency for site *i* with site type *w(i)* and year *j* (includes cross section or control type *x(i)* for crash type *y* and severity *z*) (crashes/yr); and

$N_{p,\,w(i),\,x(i),\,y,\,z,\,j}$ = predicted average crash frequency for site *i* with site type *w(i)* and year *j* (includes cross section or control type *x(i)* for crash type *y* and severity *z*) (crashes/yr).

The computation is performed separately for each predictive model identified in Step 1. The computed calibration factor is rounded to two decimal places for application in the appropriate predictive model.

B.1.2. Development of Jurisdiction-Specific Safety Performance Functions (SPFs) for Use in the Predictive Method

Satisfactory results from the Chapter 18 and 19 predictive methods can be obtained by calibrating the predictive model for each predictive model, as explained in Section B.1.1. However, some users may prefer to develop jurisdiction-specific SPFs using their agency's own data because these SPFs are likely to enhance the reliability of the predictive method. While there is no requirement that this be done, HSM users are welcome to use local data to develop their own SPFs.

Within the first two to three years after a jurisdiction-specific SPF is developed, calibration of the jurisdiction-specific SPF may not be necessary, particularly if other default values in the predictive models were also replaced with locally-derived values, as explained in Section B.1.3.

If jurisdiction-specific SPFs are used in a predictive method, they need to be developed with methods that are statistically valid and developed in such a manner that they fit into the applicable predictive method. The following guidelines for development of jurisdiction-specific SPFs that are acceptable for use in Chapters 18 and 19 include:

- In preparing the crash data to be used for development of jurisdiction-specific SPFs, crashes are assigned to roadway segments and intersections following the definitions explained in Section B.2.3.
- The jurisdiction-specific SPF should be developed with a statistical technique (such as negative binomial regression) that accounts for the overdispersion typically found in crash data, and quantifies an overdispersion parameter.
- The jurisdiction-specific SPF should use the same base conditions as the corresponding SPF in Chapter 18 or 19, or should be capable of being converted to those base conditions.
- The jurisdiction-specific SPF should include the effects of traffic volume. For segments, the average annual daily traffic volume is included. For intersections, the major- and minor-road average annual daily traffic volumes are included.
- The jurisdiction-specific SPF for any roadway segment facility type should have a functional form in which predicted average crash frequency is directly proportional to segment length.

These guidelines are not intended to stifle creativity and innovation in model development. However, a model that does not account for overdispersed data or that cannot be integrated with the rest of the predictive method will not be useful.

Two types of data sets may be used for SPF development. First, SPFs may be developed using only data that represent the base conditions, which are defined for each SPF in Chapters 18 and 19. Second, it is also acceptable to develop models using data for a broader set of conditions than the base conditions. In this approach, all variables that are part of the applicable base-condition definition, but have non-base-condition values, should be included in an initial model. Then, the initial model should be made applicable to the base conditions by substituting values that correspond to those base conditions into the model.

B.1.3. Replacement of Selected Default Values in the Predictive Methods

Table B-3 identifies the specific distributions used in the Chapter 18 and 19 predictive methods. The default distribution values provided in these tables were developed from the most complete and consistent databases available. If desired, these default values may be replaced with locally derived values. This replacement is optional, but it may yield more reliable results.

Any replacement values derived with the procedures presented in this section should be incorporated in the predictive models before the calibration described in Section B.1.1 is performed.

Table B-3. Crash Distributions in Chapter 18 and 19 Predictive Models That May Be Calibrated to Local Conditions

Chapter	Table Number	Site Type: Roadway Segments	Site Type: Intersections	Distribution That May Be Calibrated to Local Conditions
18—Freeways	Table 18-6	X		Crash type for multiple-vehicle crashes
	Table 18-8	X		Crash type for single-vehicle crashes
	Table 18-10		X	Crash type for ramp-entrance-related crashes
	Table 18-12		X	Crash type for ramp-exit-related crashes
19—Ramps	Table 19-6	X		Crash type for multiple-vehicle crashes
	Table 19-9	X		Crash type for single-vehicle crashes
	Table 19-16		X	Crash type for signal-controlled ramp terminal crashes
	Table 19-21		X	Crash type for one-way stop-controlled ramp terminal crashes
	Table 19-45		X	Crash type for all-way stop-controlled ramp terminal crashes

B.1.3.1. Replacement of Default Values for Freeways

Four default distributions for freeways may be updated with locally-derived replacement values. Procedures to develop each of these replacement values are described in the following subsections.

Crash Type for Multiple-Vehicle Crashes

Table 18-6 presents the distribution of multiple-vehicle crashes by crash type for freeway segments. The distribution is categorized by two crash severity levels and two area types. If sufficient data are available, the values in Table 18-6 may be updated. This table represents a joint distribution of two variables for each area type. Therefore, for a given area type, sufficient data for calibrating the distribution requires a set of freeway segments that have collectively experienced at least 200 multiple-vehicle crashes during a recent one- to three-year period (i.e., 200 crashes in the entire time period).

Crash Type for Single-Vehicle Crashes

Table 18-8 presents the distribution of single-vehicle crashes by crash type for freeway segments. The distribution is categorized by two crash severity levels and two area types. If sufficient data are available, the values in Table 18-8 may be updated. This table represents a joint distribution of two variables for each area type. Therefore, for a given area type, sufficient data for calibrating the distribution requires a set of freeway segments that have collectively experienced at least 200 single-vehicle crashes during a recent one- to three-year period (i.e., 200 crashes in the entire time period).

Crash Type for Ramp-Entrance-Related Crashes

Table 18-10 presents the distribution of ramp-entrance-related crashes by crash type for freeway ramp entrances (and adjacent freeway lanes). The distribution is based on ramp-entrance speed-change lane crashes. It does not include crashes associated with a ramp entrance that adds a lane to the cross section. The distribution is categorized by two crash severity levels and two area types. If sufficient data are available, the values in Table 18-10 may be updated. This table represents a joint distribution of two variables for each area type. Therefore, for a given area type, sufficient data for calibrating the distribution requires a set of ramp entrances (and adjacent freeway lanes) that have collectively experienced at least 200 crashes during a recent one- to three-year period (i.e., 200 crashes in the entire time period).

Crash Type for Ramp-Exit-Related Crashes
Table 18-12 presents the distribution of ramp-exit-related crashes by crash type for freeway ramp exits (and adjacent freeway lanes). The distribution is based on ramp-exit speed-change lane crashes. It does not include crashes associated with a ramp exit that drops a lane from the cross section. The distribution is categorized by two crash severity levels and two area types. If sufficient data are available, the values in Table 18-12 may be updated. This table represents a joint distribution of two variables for each area type. Therefore, for a given area type, sufficient data for calibrating the distribution requires a set of ramp exits (and adjacent freeway lanes) that have collectively experienced at least 200 crashes during a recent one- to three-year period (i.e., 200 crashes in the entire time period).

B.1.3.2. Replacement of Default Values for Ramps
Five default distributions for ramps may be updated with locally-derived replacement values. Procedures to develop each of these replacement values are described in the following subsections.

Crash Type for Multiple-Vehicle Crashes
Table 19-6 presents the distribution of multiple-vehicle crashes by crash type for ramp and C-D road segments. The distribution is categorized by two crash severity levels. If sufficient data are available, the values in Table 19-6 may be updated. Sufficient data for calibrating the distribution requires a set of ramp and C-D road segments that have collectively experienced at least 200 multiple-vehicle crashes during a recent one- to three-year period (i.e., 200 crashes in the entire time period).

Crash Type for Single-Vehicle Crashes
Table 19-9 presents the distribution of single-vehicle crashes by crash type for ramp and C-D road segments. The distribution is categorized by two crash severity levels and two area types. If sufficient data are available, the values in Table 19-9 may be updated. This table represents a joint distribution of two variables for each area type. Therefore, for a given area type, sufficient data for calibrating the distribution requires a set of ramp and C-D road segments that have collectively experienced at least 200 single-vehicle crashes during a recent one- to three-year period (i.e., 200 crashes in the entire time period).

Crash Type for Signal-Controlled Ramp Terminal Crashes
Table 19-16 presents the distribution of intersection-related crashes by crash type for signal-controlled crossroad ramp terminals. The distribution is categorized by two crash severity levels and two area types. If sufficient data are available, the values in Table 19-16 may be updated. This table represents a joint distribution of two variables for each area type. Therefore, for a given area type, sufficient data for calibrating the distribution requires a set of signal-controlled ramp terminals that have collectively experienced at least 200 intersection-related crashes during a recent one- to three-year period (i.e., 200 crashes in the entire time period).

Crash Type for One-Way Stop-Controlled Ramp Terminal Crashes
Table 19-21 presents the distribution of intersection-related crashes by crash type for one-way stop-controlled crossroad ramp terminals. The distribution is categorized by two crash severity levels and two area types. If sufficient data are available, the values in Table 19-21 may be updated. This table represents a joint distribution of two variables for each area type. Therefore, for a given area type, sufficient data for calibrating the distribution requires a set of one-way stop-controlled ramp terminals that have collectively experienced at least 200 intersection-related crashes during a recent one- to three-year period (i.e., 200 crashes in the entire time period).

Crash Type for All-Way Stop-Controlled Ramp Terminal Crashes
Table 19-45 presents the distribution of intersection-related crashes by crash type for all-way stop-controlled crossroad ramp terminals. The distribution is categorized by two crash severity levels and two area types. If sufficient data are available, the values in Table 19-45 may be updated. This table represents a joint distribution of two variables for each area type. Therefore, for a given area type, sufficient data for calibrating the distribution requires a set

of all-way stop-controlled ramp terminals that have collectively experienced at least 200 intersection-related crashes during a recent one- to three-year period (i.e., 200 crashes in the entire time period).

B.1.4. Calibration of Severity Distribution Functions

The SDFs used in the predictive methods of Chapters 18 and 19 were developed from the most complete and consistent databases available. If desired, these SDFs may be calibrated to local conditions. This calibration is optional, but it may yield more reliable estimates of expected average crash frequency by severity level.

The procedure described in this section is used to quantify the calibration factor for an SDF. The procedure consists of five steps. It requires data for a set of sites (i.e., freeway segments, speed-change lanes, ramp segments, or crossroad ramp terminals) that are located in the jurisdiction of interest.

The SDF calibration factors will have values greater than 1.0 for sites that, on average, experience more severe crashes than those used in the development of the SDFs. Similarly, the calibration factors for sites that experience fewer severe crashes on average than those used in the development of the SDFs will have values less than 1.0.

The procedures presented in this subsection should be used *after* the predictive models have been calibrated using the procedures described in Sections B.1.1 and B.1.3. The calibrated predictive models are used to determine the calibration factor for an SDF.

B.1.4.1. Step 1—Identify the site types for which the SDFs are to be calibrated.

Calibration is performed separately for each SDF provided in Chapters 18 and 19. Chapter 18 provides an SDF for freeway segments and speed-change lanes. Chapter 19 provides an SDF for ramp and C-D road segments. It also provides an SDF for one-way stop-controlled crossroad ramp terminals and an SDF for signal-controlled ramp terminals. The site types needed to calibrate a given SDF are identified in this step.

Also established in this step is the calibration period. Because crash severity is likely to change over time, a calibration period longer than three years is not recommended. The calibration period should have a duration that is a multiple of 12 months to avoid seasonal effects. It is recommended to use the same calibration period for all sites, but exceptions may be made where necessary.

B.1.4.2. Step 2—Select sites for calibration of the SDF.

Calibration sites are selected during this step. One set of calibration sites is assembled for each SDF identified in Step 1. It is desirable that these sites be reasonably representative of the range of site characteristics to which the predictive model will be applied. However, no formal stratification by traffic volume or other site characteristics is needed in selecting the calibration sites. As such, the sites can be selected in a manner to make the data collection needed for Step 3 as efficient as practical.

Each calibration site should be selected without regard to the number or severity of crashes reported during the calibration period. In other words, calibration sites should not be selected to intentionally limit the calibration database to include only sites with either high or low crash frequencies. Also, they should not be selected to intentionally limit the database to include sites with either more severe or less severe crashes. Where practical, this may be accomplished by selecting calibration sites randomly from a larger set of candidate sites.

The desirable minimum sample size for the calibration database for one site type is 30 to 50 sites. For segments, each site should be between 0.1 and 1.0 mi in length. Lengths in this range should be long enough to have statistical validity and short enough to be realistically homogeneous.

For large jurisdictions, such as entire states, with a variety of topographical and climate conditions, it may be desirable to assemble a separate set of calibration sites representing two or three different conditions. In this manner, separate calibration factors are developed for each specific terrain type or geographical region for a given site type.

For example, a state with distinct plains and mountain regions (or with distinct dry and wet regions), might choose to develop separate calibration factors for those regions. Where separate calibration factors are developed by terrain type or region, this needs to be done consistently for all site types applicable to those regions.

B.1.4.3. Step 3—Obtain data for each set of calibration sites for the calibration period.
This step is repeated for each SDF identified in Step 1 and its associated set of calibration sites assembled in Step 2. For this step, a calibration database is assembled for each set of calibration sites. The calibration data are assembled for a common calibration period for all sites. The calibration database should include the following information for each site represented in the database:

- all fatal or injury crashes that are reported during the calibration period, and
- site characteristics data needed to apply the predictive method for the same calibration period.

Only fatal or injury crashes should be included in the calibration database. Each observation in the database represents one site. It includes the site characteristics as well as the separate count of fatal, incapacitating injury, nonincapacitating injury, and possible injury crashes reported during the calibration period.

For a given site type, the calibration database should include at least 300 fatal or injury crashes per calibration period. If this minimum is not realized, then (a) additional sites should be added to the database following the guidelines in Step 2 or (b) the calibration period should be expanded to include additional years of crash data.

The crash data used for calibration should include all fatal or injury crashes related to each site selected for the calibration database. Crashes should be assigned to specific sites based on the guidelines presented in Section B.2.3.

Table B-2 identifies the site characteristics data that are needed to apply the predictive method. The table classifies each data element as either required or desirable for the calibration procedure. Data for each of the required elements are needed for calibration. For the desirable data elements, it is recommended that actual data be used if available. Assumptions are offered in the table when these data are not available.

B.1.4.4. Step 4—Apply the applicable predictive method to estimate the predicted average crash frequency by severity for each site during the calibration period.
This step is repeated for each SDF identified in Step 1 and its associated set of calibration sites assembled in Step 2. The site characteristics data assembled in Step 3 are used to apply the applicable predictive method to each site in the set of calibration sites. For this application, the predictive model should be applied without using the EB Method. The SDF calibration factor is set to 1.00. Through this process, the predicted average crash frequency for each severity level is obtained for each site in the set of calibration sites and for each year in the calibration period.

B.1.4.5. Step 5—Compute the calibration factors for use in the SDFs.
This step is repeated for each SDF identified in Step 1 and its associated set of calibration sites assembled in Step 2. It consists of three tasks.

During the first task, the observed crash data are used to calculate the observed probability of a severe crash (i.e., fatal K, incapacitating injury A, or nonincapacitating injury B), given that a fatal or injury crash has occurred. Equation B-2 is used for this purpose. In this manner, one overall average value is obtained for all sites represented in the database.

$$P_{o,aS,ac,at,KAB} = \frac{\sum_{i}^{n_{sites}} \sum_{j=1}^{n_c} \left(N_{o,w(i),x(i),at,K,j} + N_{o,w(i),x(i),at,A,j} + N_{o,w(i),x(i),at,B,j} \right)}{\sum_{i}^{n_{sites}} \sum_{j=1}^{n_c} \left(N_{o,w(i),x(i),at,K,j} + N_{o,w(i),x(i),at,A,j} + N_{o,w(i),x(i),at,B,j} + N_{o,w(i),x(i),at,C,j} \right)} \tag{B-2}$$

Where:

$P_{o,\,aS,\,ac,\,at,\,KAB}$ = observed probability of a severe crash (i.e., *K*, *A*, or *B*) for all crash types *at* at all sites *aS* and all cross sections or control types *ac*;

n_c = number of years in the crash period (yr);

$N_{o,\,w(i),\,x(i),\,at,\,m,\,j}$ = observed crash frequency for site *i* with site type *w(i)* and year *j* (includes cross section or control type *x(i)* for all crash types *at* and severity level *m*, with *m* = *K*, *A*, *B*, *C*) (crashes/yr); and

n_{sites} = number of sites.

In the second task, the predicted average crash frequency by severity from Step 4 is used to calculate the predicted probability of occurrence of a severe crash, given that a fatal or injury crash has occurred. Equation B-3 is used for this purpose. In this manner, one overall average value is obtained for all sites represented in the database.

$$P_{p,aS,ac,at,KAB} = \frac{\sum_{i}^{n_{sites}} \sum_{j=1}^{n_c} \left(N_{p,w(i),x(i),at,K,j} + N_{p,w(i),x(i),at,A,j} + N_{p,w(i),x(i),at,B,j} \right)}{\sum_{i}^{n_{sites}} \sum_{j=1}^{n_c} \left(N_{p,w(i),x(i),at,K,j} + N_{p,w(i),x(i),at,A,j} + N_{p,w(i),x(i),at,B,j} + N_{p,w(i),x(i),at,C,j} \right)} \tag{B-3}$$

Where:

$P_{p,\,aS,\,ac,\,at,\,KAB}$ = predicted probability a severe crash (i.e., *K*, *A*, or *B*) for all crash types *at* at all sites *aS* and all cross sections or control types *ac*;

n_{sites} = number of sites; and

$N_{p,\,w(i),\,x(i),\,at,\,m,\,j}$ = predicted crash frequency for site *i* with site type *w(i)* and year *j* (includes cross section or control type *x(i)* for all crash types *at* and severity level *m*, with *m* = *K*, *A*, *B*, *C*) (crashes/yr).

The final step is to compute the calibration factor using the following equation. The appropriate site-type subscript in this equation is uniquely defined for each SDF identified in Step 1.

$$C_{sdf,w} = \frac{P_{o,aS,ac,at,KAB}}{1.0 - P_{o,aS,ac,at,KAB}} \times \frac{1.0 - P_{p,aS,ac,at,KAB}}{P_{p,aS,ac,at,KAB}} \tag{B-4}$$

Where:

$C_{sdf,\,w}$ = calibration factor to adjust SDF for local conditions for site type *w*.

The computation is performed separately for each SDF identified in Step 1. The computed calibration factor is rounded to two decimal places for application in the appropriate SDF.

B.2. THE EMPIRICAL BAYES METHOD

The EB Method is used to combine the estimate from a predictive model with observed crash data to obtain a more reliable estimate of the expected average crash frequency. The development of the EB Method described in this appendix is documented by Hauer (1).

The EB Method improves the reliability of the estimate of expected average crash frequency by pooling the estimate from a predictive model with the subject site's observed crash data. The model estimate describes the safety of the typical site with attributes matching those of the subject site. However, it has some level of statistical uncertainty due to unexplained differences among the set of similar sites used to calibrate the predictive model. Similarly, an average crash frequency computed from crash data has uncertainty because of the random variability inherent to crash data. The EB Method produces an estimate of the expected average crash frequency that combines the model prediction and the site-specific crash data in proportion to the level of certainty that can be attached to each.

Each of Chapters 18 and 19 presents a four-step process for applying the EB Method. Before the EB Method can be applied, the appropriate predictive model must be used to determine the predicted average crash frequency for each site of interest. Each site's predicted average crash frequency is estimated for each year in a specified crash period. The steps in applying the EB Method are:

- Determine whether the EB Method is applicable.
- Determine whether observed crash data are available for the project or facility for a desired crash period. Acquire the crash data for this crash period.
- Apply the EB Method to estimate the expected average crash frequency by combining the predicted average crash frequency and observed crash data for the crash period.
- Adjust the estimated value of expected average crash frequency to a future time period, if appropriate.

B.2.1. Determine whether the EB Method is Applicable

The applicability of the EB Method to a particular project depends on the type of analysis being performed and the type of future project work that is anticipated. If the analysis is being performed to evaluate the safety of an existing project, then the EB Method should be applied.

If a future project is being planned, then the nature of that future project should be considered in deciding whether to apply the EB Method. Specifically, the EB Method should be applied for the analyses involving the following future project types:

- Sites at which the roadway geometrics and traffic control are not being changed (e.g., the "do-nothing" alternative).
- Projects in which the roadway cross section is modified but the basic number of through lanes remains the same. This could include projects for which lanes or shoulders were widened or the roadside was improved.
- Projects in which minor changes in alignment are made, such as flattening individual horizontal curves while leaving most of the alignment intact.
- Projects in which a weaving section is added to a freeway.
- Any combination of the above improvements.

The EB Method is not applicable to the following types of improvements.

- Projects in which a new alignment is developed for a substantial proportion of the project length.
- Crossroad ramp terminals at which the basic number of intersection legs or type of traffic control is changed as part of a project.

The reason that the EB Method is not used for the two improvement types in the previous list is that the observed crash data for a previous time period is not necessarily indicative of the crash experience that is likely to occur after a major geometric improvement.

If the EB Method is applied to individual sites and some sites within the project limits will not undergo a major geometric improvement, it is acceptable to apply the EB Method to these sites. In other words, the site-specific EB Method can be applied to some sites within the project limits and not applied to other sites.

If alternative improvements are being evaluated for a given project and the EB Method is being considered, then the EB Method will need to be consistently applied to all alternatives being evaluated. If the EB Method cannot be consistently applied to all alternatives, then it should not be used for any alternatives (i.e., the predictive method should be used without EB adjustment). This approach recognizes that there is typically a small difference in the results obtained from the predictive method when it is used with and without the EB Method. If the EB Method is not applied consistently, such differences will likely introduce a small bias in the comparison of expected crash frequency among alternatives.

If the EB Method is not applicable, do not proceed to the remaining steps. Instead, follow the predictive method in Chapter 18 or 19 but skip Steps 6, 13, and 15.

B.2.2. Determine whether Observed Crash Data are Available for the Project and, if so, Obtain those Data

If the EB Method is determined to be applicable to a given project, then it should be determined whether observed crash data are available directly from the jurisdiction's crash record system or indirectly from another source. At least two years of observed crash data are desirable to apply the EB Method.

Two variations of the EB Method are available. They are the site-specific EB Method and the project-level EB Method. The appropriate variation to use for a given project depends on the level of detail provided in the crash record system, the site types to which the method is applied, and the crash types associated with the predictive model that will be used. In general, the best results will be obtained if the site-specific EB Method is used. Figure B-1 provides a flow chart to assist in the determination of whether the site-specific or project-level variation of the EB Method is applicable for a given project.

B.2.2.1. Projects with One Site

Two considerations are discussed in this section. The first consideration relates to crash type. It is included in Figure B-1. The second consideration relates to crash severity. It is not included in the figure and may only apply in rare instances.

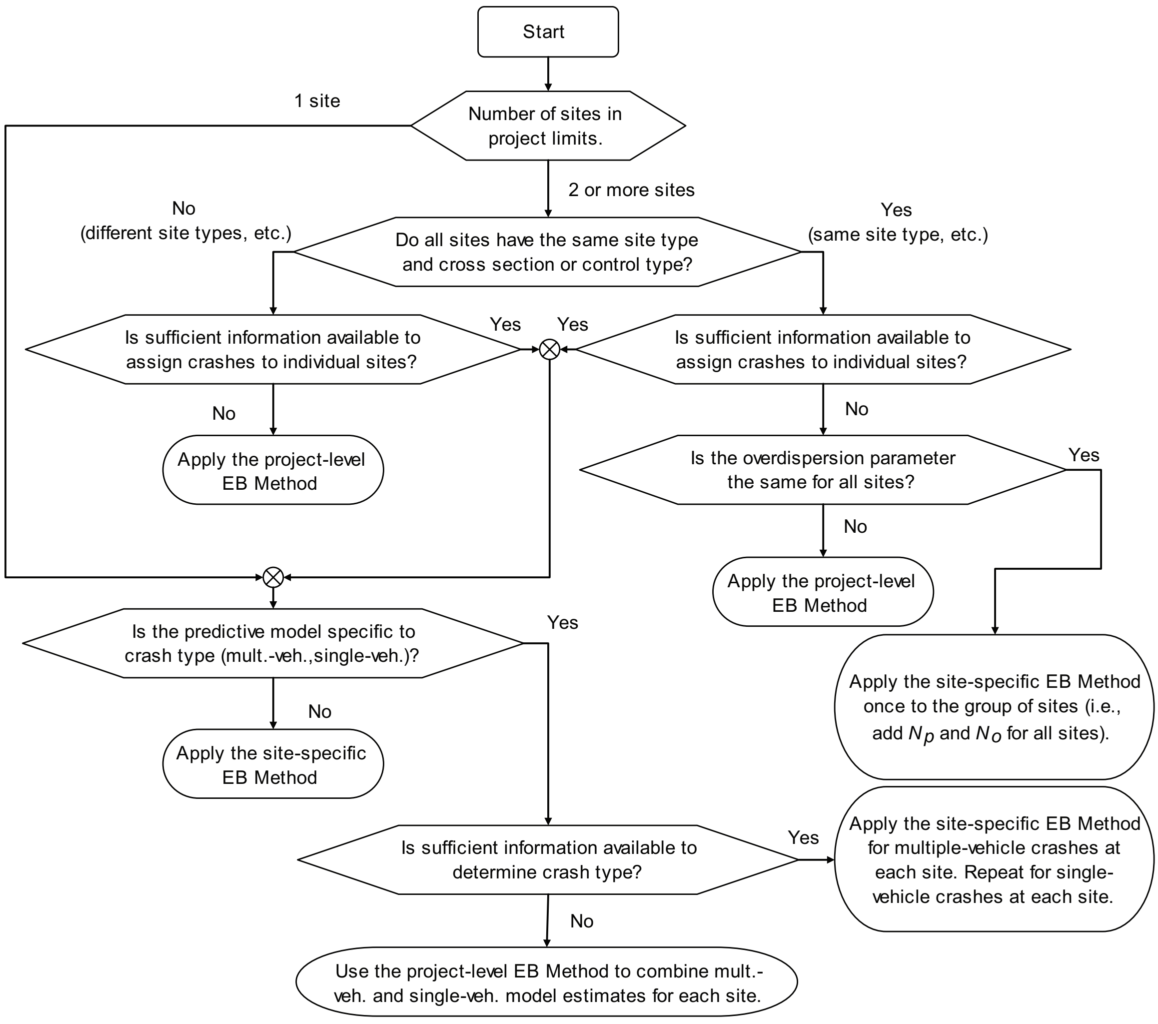


Figure B-1. Determination of the Appropriate Variation of the EB Method

Crash Type Considerations

For projects that consist of one site, Figure B-1 indicates that the first consideration is whether the applicable predictive model is specific to one crash type (i.e., multiple-vehicle crashes or single-vehicle crashes). The predictive methods in Chapters 18 and 19 include some models that are specific to crash type. These models have different overdispersion parameters for each crash type. They are identified in Table B-1.

If the crash record system provides sufficient information to determine crash type for each site and a model with this sensitivity from Chapter 18 or 19 is being used, then the site-specific EB Method is applicable. This method is described in Section B.2.4.

If the crash record system does *not* provide sufficient information to determine crash type and a model with this sensitivity from Chapter 18 or 19 is being used, then the project-level EB Method is applicable. This method is described in Section B.2.5.

Crash Severity Considerations

Although not shown in Figure B-1, another consideration is whether the applicable predictive model is specific to crash severity (i.e., fatal-and-injury crashes or property-damage-only crashes). The predictive methods in Chapters 18 and 19 include models that are specific to crash severity. These models have different overdispersion parameters for each severity.

If the crash record system provides sufficient information to determine crash severity for each site, then the site-specific EB Method is applicable. This method is described in Section B.2.4.

If the crash record system does *not* provide sufficient information to determine crash severity, then the project-level EB Method is applicable. This method is described in Section B.2.5.

Once the total expected average crash frequency is obtained, the estimate of expected average crash frequency for fatal-and-injury crashes is calculated by applying the proportion of predicted average crash frequency for fatal-and-injury crashes (i.e., $N_{p, w, x, y, fi} / N_{p, w, x, y, as}$) to the total expected average crash frequency. Similarly, the estimate of expected average crash frequency for property-damage-only crashes is calculated by applying the proportion of predicted average crash frequency for property-damage-only crashes (i.e., $N_{p, w, x, y, pdo} / N_{p, w, x, y, as}$) to the total expected average crash frequency.

B.2.2.2. Projects with Two or More Sites

For projects that consist of two or more sites, Figure B-1 indicates that there are several considerations when determining the appropriate EB Method variation. The first consideration relates to the site types and cross sections or control types represented within the project limits. In general, a project will consist of many sites that collectively represent different site types, cross sections, and control types. Occasionally, a project may consist of several sites that have the same site type and cross section or control type (e.g., a succession of segments along a specific freeway).

Projects with Different Types of Sites

If a project consists of several sites that collectively have different site types, cross sections, or control types, then the next consideration is whether the crash record system provides sufficient information to assign observed crashes to the individual sites. If the crashes can be assigned to individual sites, then the evaluation proceeds on a site-by-site basis. In this situation, the discussion in Section B.2.2.1 applies and the guidance therein is followed to determine the appropriate EB Method variation. Criteria for assigning crashes to individual sites are presented in Section B.2.3.

If the crashes cannot be assigned to individual sites, then the project-level EB Method is applicable. This method is described in Section B.2.5.

Projects with the Same Site Types

If a project consists of several sites that have the same site type and cross section or control type, then the next consideration is whether the crash record system provides sufficient information to assign observed crashes to the individual sites. If the crashes can be assigned to individual sites, then the evaluation proceeds on a site-by-site basis. In this situation, the discussion in Section B.2.2.1 applies and the guidance therein is followed to determine the appropriate EB Method variation. Criteria for assigning crashes to individual sites are presented in Section B.2.3.

If the crashes cannot be assigned to individual sites, then the next consideration is whether the overdispersion parameter is the same for all the sites. The overdispersion parameter is constant for some predictive models; for others, it is a function of segment length. For those models in which it is a function of segment length, the length of each segment would have to be the same to produce an overdispersion factor that is the same for all sites.

If the overdispersion parameter is the same for all sites, then the site-specific EB Method can be used. In this application, the predicted average crash frequency for each site is combined into a single estimate for the group of sites. Similarly, the observed crash count for each site is combined into a single estimate for the group of sites.

If the overdispersion parameter is not the same for all sites, then the project-level EB Method is applicable. This method is described in Section B.2.5.

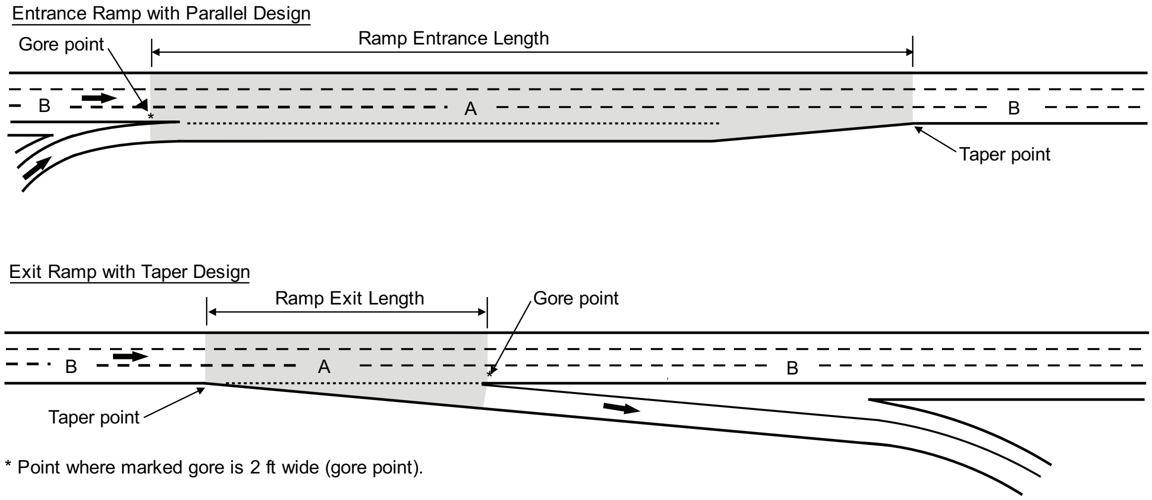
B.2.3. Assign Crashes to Individual Sites for Use in the EB Method

When using the site-specific EB Method, observed crashes for a site are combined with the predictive model estimate of crash frequency for that site to provide a more reliable estimate of its expected average crash frequency. To apply the site-specific EB Method, observed crashes are assigned to each individual site within the facility of interest. This assignment occurs during Step 6 of the predictive method. This section provides guidance for assigning crashes to sites associated with freeway facilities. Similar guidance for assigning crashes to sites associated with street and highway facilities is provided in Section A.2.3 of the *Highway Safety Manual*.

Guidance for Assigning Crashes to Freeway Segments and Speed-Change Lanes

Speed-change-lane-related crashes include all crashes that are located between the gore point and the taper point of a speed-change lane and that (a) involve vehicles in the speed-change lane, (b) involve vehicles in the freeway lanes on the same side of the freeway as the speed-change lane, or (c) occur in the median by a vehicle traveling in a lane on the same side of the freeway as the speed-change lane and in the same basic direction as vehicles in the speed-change lane. All freeway crashes that are not classified as speed-change-lane-related crashes are considered to be freeway segment crashes.

Figure B-2 illustrates the method used to assign crashes to freeway segments or speed-change lanes. All crashes that occur in Region A are assigned to the speed-change lane. Crashes that occur outside of Region A (i.e., in Region B) are assigned to the freeway segment.



A All crashes that occur within this region are classified as speed-change crashes.
B All crashes that occur within this region are classified as freeway segment crashes.

Figure B-2. Definition of Freeway Segments and Speed-Change Lanes

Guidance for Assigning Crashes to Ramp Segments and Crossroad Ramp Terminals

The guidance for assigning crashes to intersections described in Section A.2.3 of the *Highway Safety Manual* also applies to assigning crashes to crossroad ramp terminals. Exceptions to this guidance are described in the following paragraphs. Crashes that are not assigned to the crossroad ramp terminal are assigned to the crossroad or the intersecting ramp segments.

The predictive models for crossroad ramp terminals include consideration of crashes on the crossroad that are associated with an unsignalized driveway or public street approach located within 250 ft of the crossroad ramp terminal. The interaction between driveway traffic and ramp terminal traffic is complex. As a result, it is often difficult to determine whether crashes between the two traffic streams are related to the driveway or the ramp terminal geometry and traffic control features. Consideration of these crashes in the crossroad ramp terminal predictive models facilitates an examination of the safety implications of these interactions. Therefore, driveway- and public-street-related crashes on the crossroad within 250 ft of the crossroad ramp terminal should be assigned to the crossroad ramp terminal (they should not be assigned to the crossroad segment).

Rear-end crashes on exit ramps should be carefully scrutinized for their relationship to the downstream crossroad ramp terminal. Lengthy queues of stopped vehicles can exist on some ramps during peak traffic demand periods. If the crash is related to the presence of a queue created by the operation of the downstream ramp terminal, then the crash should be assigned to the ramp terminal regardless of the distance between the crash location and the ramp terminal.

In general, a ramp is defined to begin at a gore point and end at (a) another gore point (when ending at another ramp) or (b) the near edge of traveled way of the crossroad (when ending at a crossroad ramp terminal). Exit-ramp-related and entrance-ramp-related crashes represent crashes that occur on a ramp, between the near edge of traveled way of the crossroad and the freeway speed-change lane gore point (as shown in Figure B-2). Connector-ramp-related crashes represent all crashes that occur on a connector ramp, between the freeway speed-change lane gore point and the crossroad speed-change lane gore point.

Any crashes that occur in a ramp speed-change lane associated with a ramp-to-ramp junction are assigned to the originating ramp (i.e., they are not assigned to the merging or diverging ramp). The merging ramp ends at the gore point of the ramp speed-change lane. The diverging ramp begins at the gore point of the ramp speed-change lane.

C-D road crashes represent crashes that occur on a C-D road, between the freeway exit gore point and the freeway entrance gore point.

B.2.4. Apply the Site-Specific EB Method

This section describes the EB Method that is used when observed crash data are available for each site of interest. It is used to estimate the expected average crash frequency (in total, or by crash type or severity) for a specific site by combining the predictive model estimate with observed crash data.

The expected average crash frequency for reference year *r* at a site *i* with site type *w(i)* and cross section or control type *x(i)* for a specified crash type *y* and severity *z* is computed using the following equation:

$$N_{e,w(i),x(i),y,z,r} = w_{w(i),x(i),y,z} \times N_{p,w(i),x(i),y,z,r} + \left(1.0 - w_{w(i),x(i),y,z}\right) \times \frac{N^{*}_{o,w(i),x(i),y,z}}{C_{b,w(i),x(i),y,z,r}} \qquad \text{(B-5)}$$

with

$$w_{w(i),x(i),y,z} = \frac{1.0}{1.0 + \left(k_{w(i),x(i),y,z} \times \sum_{j=1}^{n_c} N_{p,w(i),x(i),y,z,j}\right)} \qquad \text{(B-6)}$$

$$C_{b,w(i),x(i),y,z,r} = \frac{1.0}{N_{p,w(i),x(i),y,z,r}} \times \sum_{j=1}^{n_c} N_{p,w(i),x(i),y,z,j} \qquad \text{(B-7)}$$

Where:

$N_{e,\,w(i),\,x(i),\,y,\,z,\,r}$ = expected average crash frequency for site *i* with site type *w(i)* and reference year *r* (includes cross section or control type *x(i)* for crash type *y* and severity *z*) (crashes/yr);

$w_{w(i), x(i), y, z}$ = weighted adjustment factor for site i with site type $w(i)$ and cross section or control type $x(i)$ for crash type y and severity z.

$N_{p, w(i), x(i), y, z, r}$ = predicted average crash frequency for site i with site type $w(i)$ and reference year r (includes cross section or control type $x(i)$ for crash type y and severity z) (crashes/yr);

$N^*_{o, w(i), x(i), y, z}$ = total observed number of crashes for site i with site type $w(i)$ and all years in the crash period (includes cross section or control type $x(i)$ for crash type y and severity z) (crashes);

$C_{b, w(i), x(i), y, z, r}$ = equivalent years in the crash period relative to reference year r at site i with site type $w(i)$ and cross section or control type $x(i)$ for crash type y and severity z (yr);

$k_{w(i),x(i), y, z}$ = overdispersion parameter for site i with site type $w(i)$, cross section or control type $x(i)$, crash type y and severity z;

n_c = number of years in the crash period (yr); and

$N_{p, w(i), x(i), y, z, j}$ = predicted average crash frequency for site i with site type $w(i)$ and year j (includes cross section or control type $x(i)$ for crash type y and severity z) (crashes/yr).

The following equation is an alternative form of Equation B-5 that is useful when the expected number of crashes for the crash period is desired.

$$N^*_{e,w(i),x(i),y,z} = w_{w(i),x(i),y,z} \times \left(\sum_{j=1}^{n_c} N_{p,w(i),x(i),y,z,j} \right) + \left(1.0 - w_{w(i),x(i),y,z}\right) \times N^*_{o,w(i),x(i),y,z} \tag{B-8}$$

Where:

$N^*_{e, w(i), x(i), y, z}$ = total expected number of crashes for site i with site type $w(i)$ and all years in the study period (includes cross section or control type $x(i)$ for crash type y and severity z) (crashes).

The expected average crash frequency is specific to a given site type w, cross section or control type x, crash type y, and severity z. As a result, the variables used in these equations must all be consistent in their representation of site type, cross section or control type, crash type, and severity. Also, the reference year r must be one of the years represented in the crash period. The crash period is defined as the consecutive years for which observed crash data are available.

The predicted average crash frequencies used in Equation B-5 to Equation B-7 are obtained from the appropriate predictive model described in Chapter 18 or 19. Similarly, the overdispersion parameter used in Equation B-6 is obtained from the same predictive model as used to estimate the predicted average crash frequencies.

The overdispersion parameter is shown to be specific to site i. This situation will apply whenever it is computed as a function of segment length, which is the case for Chapters 18 and 19. When it is not a function of segment length, the subscript components referencing site i are removed.

Equation B-6 shows an inverse relationship between the overdispersion parameter and the weight w. This implies that when a model with little overdispersion is available, more reliance will be placed on the predictive model estimate N_p and less reliance on the observed crash count N^*_o. The opposite is also the case; when a model with substantial overdispersion is available, less reliance will be placed on the predictive model estimate and more reliance on the observed crash count.

It is important to note in Equation B-5 that, as N_p increases, there is less weight placed on it and more on N^*_o. This might seem counterintuitive at first. However, this implies that for longer sites and for longer study periods, there are more opportunities for crashes to occur. Thus, the observed crash history is likely to be more meaningful and the model prediction less important. So, as N_p increases, the EB Method places more weight on the number of crashes that actually occur. When few crashes are predicted, the observed crash count is not likely to be meaningful, in statistical terms, so greater reliance is placed on the predicted crash frequency.

Chapters 18 and 19 present worksheets that can be used to apply the site-specific EB Method as presented in this section.

Section B.2.6 explains how to us Equation B-5 to estimate the expected average crash frequency for a time period other than the crash period, such as the time period when a proposed future project will be implemented.

B.2.5. Apply the Project-Level EB Method

This section describes an alternative EB Method that is used when observed crash data are aggregated across several sites (e.g., for an entire facility or project). The development of this variation of the EB Method is documented by Hauer et al. (2).

In general, the EB Method described in this section is used when the predictive model and its overdispersion parameter are not uniquely defined for the combined set of sites being evaluated. It is also needed when the predictive model is specific to crash type (or severity) but the information in the crash database is insufficient to make crash type determinations.

When the crash data cannot be disaggregated to the level of the predictive model, the estimates from each of the predictive models for the various sites have different weights. These estimates cannot be directly combined to compute an overall weighted adjustment factor *w* because they are likely correlated to some degree (e.g., all sites in a given project may be consistently safer [or less safe] than the similar sites used to calibrate the predictive model). Because the degree of correlation is unknown, an approximate method is used to estimate the expected average crash frequency for each of two extreme conditions of correlation. The first condition assumes that the estimates among sites are independent. The second condition assumes that the estimates among sites are perfectly correlated. The best estimate of expected average crash frequency is rationalized to be the average of these two extreme conditions.

The following procedure describes the sequence of calculations necessary to implement the project-level EB Method. To facilitate the presentation of this procedure, the equations shown in this section have subscripts denoting fatal-and-injury *fi* crashes of all crash types (i.e., multiple-vehicle and single-vehicle fatal-and-injury crashes combined). The conversion of these equations so that they are applicable to property-damage-only crashes (or crashes of a specific crash type) requires only the substitution of the appropriate subscripts.

Step 1—Sum the predicted average crash frequency and observed crash counts.
The desired crash type *w* and crash severity *z* are specified during this step. The crash type chosen must have an associated predictive model. For example, if an estimate of the expected average multiple-vehicle crash frequency is desired, then an SPF that predicts multiple-vehicle crash frequency must be available in the predictive model. Similarly, if an estimate of the expected average fatal-and-injury crash frequency is desired, then an SPF that predicts fatal-and-injury crash frequency is required.

The predicted average crash frequency is summed for each site and year represented in the crash period to obtain the predicted number of crashes for all sites and all years in the crash period. Each site *i* will have a specific site type *w(i)* and cross section or control type *x(i)*. Similarly, the observed crash counts are summed for each site and year represented in the crash period to obtain the observed number of crashes for all sites and all years in the crash period. The following equations are used to compute the desired sums for fatal-and-injury crashes of all crash types.

$$N^{*}_{p,aS,ac,at,fi} = \sum_{i}^{n_{sites}} \sum_{k}^{n_{\substack{crash\\types}}} \sum_{j=1}^{n_c} N_{p,\,w(i),x(i),y(k),fi,j} \tag{B-9}$$

$$N^*_{o,aS,ac,at,fi} = \sum_{i}^{n_{sites}} \sum_{k}^{n_{crash\ types}} \sum_{j=1}^{n_c} N_{o,\,w(i),x(i),y(k),\,fi,\,j} \qquad \text{(B-10)}$$

Where:

$N^*_{p,\,aS,\,ac,\,at,\,fi}$ = total predicted number of crashes for all sites *aS* and all years in the crash period (includes all cross sections *ac* and fatal-and-injury crashes *fi* of all crash types *at*) (crashes);

n_{sites} = number of sites;

$n_{crash\ types}$ = number of crash types;

$N_{p,\,w(i),\,x(i),\,y(k),\,fi,\,j}$ = predicted average crash frequency for site *i* with site type *w(i)* and year *j* (includes cross section or control type *x(i)* for fatal-and-injury crashes *fi* of crash type *y(k)*) (crashes/yr);

$N^*_{o,\,aS,\,ac,\,at,\,fi}$ = total observed number of crashes for all sites *aS* and all years in the crash period (includes all cross sections *ac* and fatal-and-injury crashes *fi* of all crash types *at*) (crashes); and

$N_{o,\,w(i),\,x(i),\,y(k),\,fi,\,j}$ = observed crash frequency for site *i* with site type *w(i)* and year *j* (includes cross section or control type *x(i)* for fatal-and-injury crashes *fi* of crash type *y(k)*) (crashes/yr).

The predicted average crash frequencies used in Equation B-9 are obtained from the appropriate predictive model described in Chapter 18 or 19. Because the EB Method is applied at the project level, it is likely that observed crashes cannot be associated with specific sites and Equation B-10 cannot be directly used. In this situation, the analyst should use the equation as guidance when consulting crash records to identify crashes of all types that are associated with one of the sites in the project limits and that occur during the crash period.

Step 2—Compute the variance of the predicted average crash frequency.
Two variance estimates are computed in this step. One estimate is based on the assumption that the sites are independent, and the other estimate is based on the assumption that the sites are perfectly correlated. The following equations are used for these computations.

$$V_{0,aS,ac,at,fi} = \sum_{i}^{n_{sites}} \sum_{k}^{n_{crash\ types}} k_{w(i),x(i),y(k),fi} \times \left[\sum_{j=1}^{n_c} N_{p,\,w(i),x(i),y(k),\,fi,\,j}\right]^2 \qquad \text{(B-11)}$$

$$V_{1,aS,ac,at,fi} = \left(\sum_{i}^{n_{sites}} \sum_{k}^{n_{crash\ types}} \sqrt{k_{w(i),x(i),y(k),fi} \times \left[\sum_{j=1}^{n_c} N_{p,\,w(i),x(i),y(k),\,fi,\,j}\right]^2}\right)^2 \qquad \text{(B-12)}$$

Where:

$V_{0,\,aS,\,ac,\,at,\,fi}$ = variance of the predicted average crash frequency assuming independence for all sites *aS* and all years in the crash period (includes all cross sections *ac* and fatal-and-injury crashes *fi* of all crash types *at*) (crashes²/yr²);

n_{sites} = number of sites;

$n_{crash\ types}$ = number of crash types;

$k_{w(i),\,x(i),\,y(k),\,fi}$ = overdispersion parameter for site i with site type $w(i)$ and cross section or control type $x(i)$ for fatal-and-injury crashes of crash type $y(k)$; and

$V_{1,aS,\,ac,\,at,\,fi}$ = variance of the predicted average crash frequency assuming perfect correlation for all sites aS and all years in the crash period (includes all cross sections ac and fatal-and-injury crashes fi of all crash types at) (crashes^2/yr^2).

The overdispersion parameters used in Equation B-11 and Equation B-12 are obtained from the same predictive model that was used to estimate the predicted average crash frequencies.

Step 3—Compute the weighted adjustment factor.
Two weighted adjustment factors are computed in this step. One factor is based on the assumption that the sites are independent, and the other factor is based on the assumption that the sites are perfectly correlated. The following equations are used for these computations.

$$w_{0,aS,ac,at,fi} = \frac{1.0}{1.0 + \dfrac{V_{0,aS,ac,at,fi}}{N^*_{p,aS,ac,at,fi}}} \tag{B-13}$$

$$w_{1,aS,ac,at,fi} = \frac{1.0}{1.0 + \dfrac{V_{1,aS,ac,at,fi}}{N^*_{p,aS,ac,at,fi}}} \tag{B-14}$$

Where:

$w_{0,\,aS,\,ac,\,at,\,fi}$ = weighted adjustment factor assuming independence for all sites aS and all years in the crash period (includes all cross sections ac and fatal-and-injury crashes fi of all crash types at); and

$w_{1,\,aS,\,ac,\,at,\,fi}$ = weighted adjustment factor assuming perfect correlation for all sites aS and all years in the crash period (includes all cross sections ac and fatal-and-injury crashes fi of all crash types at).

Step 4—Compute the equivalent years in the crash period.
The equivalent number of years in the crash period reflects changes in traffic volume and other factors during the crash period. The changes are relative to a specified reference year r. Any year in the crash period can be designated as the reference year. It is the year for which the expected average crash frequency will be estimated in Step 5. The equivalent number of years is computed using the following equation.

$$C_{b,aS,ac,at,fi,r} = \frac{N^*_{p,aS,ac,at,fi}}{N_{p,aS,ac,at,fi,r}} \tag{B-15}$$

with

$$N_{p,aS,ac,at,fi,r} = \sum_{i}^{n_{sites}} \sum_{k}^{n_{crash\ types}} N_{p,\,w(i),x(i),y(k),fi,r} \tag{B-16}$$

Where:

$C_{b,\,aS,\,ac,\,at,\,fi,\,r}$ = equivalent years in the crash period relative to reference year r for all sites aS, all cross sections ac, and fatal-and-injury crashes fi of all crash types at (yr);

$N_{p,\, aS,\, ac,\, at,\, fi,\, r}$ = predicted average crash frequency for all sites *aS* and reference year *r* (includes all cross sections *ac* and fatal-and-injury crashes *fi* of all crash types *at*) (crashes/yr);

n_{sites} = number of sites; and

$n_{crash\ types}$ = number of crash types.

Step 5—Compute the expected average crash frequency.
The expected average crash frequency for the reference year *r* is computed in this step. Steps 4 and 5 are repeated for other reference years of interest. The expected average crash frequency is computed as the average of the expected values for the two assumed conditions (i.e., sites are independent and sites are perfectly correlated). The following equation is used for this calculation.

$$N_{e,aS,ac,at,fi,r} = \frac{N_{0,aS,ac,at,fi,r} + N_{1,aS,ac,at,fi,r}}{2} \tag{B-17}$$

with

$$N_{0,aS,ac,at,fi,r} = w_{0,aS,ac,at,fi} \times N_{p,aS,ac,at,fi,r} + \left(1.0 - w_{0,aS,ac,at,fi}\right) \times \frac{N^*_{o,aS,ac,at,fi}}{C_{b,aS,ac,at,fi,r}} \tag{B-18}$$

$$N_{1,aS,ac,at,fi,r} = w_{1,aS,ac,at,fi} \times N_{p,aS,ac,at,fi,r} + \left(1.0 - w_{1,aS,ac,at,fi}\right) \times \frac{N^*_{o,aS,ac,at,fi}}{C_{b,aS,ac,at,fi,r}} \tag{B-19}$$

Where:

$N_{e,\, aS,\, ac,\, at,\, fi,\, r}$ = expected average crash frequency for all sites *aS* and reference year *r* (includes all cross sections *ac* and fatal-and-injury crashes *fi* of all crash types *at*) (crashes/yr);

$N_{0,\, aS,\, ac,\, at,\, fi,\, r}$ = expected average crash frequency for all sites *aS* and reference year *r* assuming independence among sites (includes all cross sections *ac* and fatal-and-injury *fi* crashes of all crash types *at*) (crashes/yr); and

$N_{1,\, aS,\, ac,\, at,\, fi,\, r}$ = expected average crash frequency for all sites *aS* and reference year *r* assuming perfect correlation among sites (includes all cross sections *ac* and fatal-and-injury *fi* crashes of all crash types *at*) (crashes/yr).

The following equation is an alternative form of Equation B-17 that is useful when the expected number of crashes for the crash period is desired.

$$N^*_{e,aS,ac,at,fi} = \frac{N^*_{0,aS,ac,at,fi} + N^*_{1,aS,ac,at,fi}}{2} \tag{B-20}$$

with

$$N^*_{0,aS,ac,at,fi} = w_{0,aS,ac,at,fi} \times N^*_{p,aS,ac,at,fi} + \left(1.0 - w_{0,aS,ac,at,fi}\right) \times N^*_{o,aS,ac,at,fi} \tag{B-21}$$

$$N^*_{1,aS,ac,at,fi} = w_{1,aS,ac,at,fi} \times N^*_{p,aS,ac,at,fi} + \left(1.0 - w_{1,aS,ac,at,fi}\right) \times N^*_{o,aS,ac,at,fi} \tag{B-22}$$

Where:

$N^*_{e,\,aS,\,ac,\,at,\,fi}$ = total expected number of crashes for all sites *aS* and all years in the study period (includes all cross sections *ac* and fatal-and-injury crashes *fi* of all crash types *at*) (crashes);

$N^*_{0,\,aS,\,ac,\,at,\,fi}$ = total expected number of crashes for all sites *aS* and all years assuming independence among sites (includes all cross sections *ac* and fatal-and-injury crashes *fi* of all crash types *at*) (crashes); and

$N^*_{1,\,aS,\,ac,\,at,\,fi}$ = total expected number of crashes for all sites *aS* and all years assuming perfect correlation among sites (includes all cross sections *ac* and fatal-and-injury crashes *fi* of all crash types *at*) (crashes).

Chapters 18 and 19 present worksheets that can be used to apply the project-level EB Method as presented in this section.

Section B.2.6 explains how to use Equation B-17 to estimate the expected average crash frequency for a time period other than the crash period, such as the time period when a proposed future project will be implemented.

B.2.6. Estimate the Expected Average Crash Frequency for a Future Time Period

The estimate obtained from Equation B-5 or Equation B-17 represents the expected average crash frequency for a given site or project, respectively, during the crash period.

This section describes a procedure that is used to obtain an estimate of the expected average crash frequency during the study period. The study period is defined as the consecutive years for which an estimate of the expected average crash frequency is desired. This procedure is used when the study period includes years that are not represented in the crash period. Typically, the study period includes one or more future years that are coincident with a proposed or anticipated change in some feature or characteristic of the project.

The procedure yields an estimate of the expected average crash frequency for a specified study year *j*. This estimate is corrected for (a) any growth or decline in AADTs between the crash period and the study period and (b) any change in geometric design or traffic control features between the crash period and the study period (as represented by the values of the associated CMFs). The estimates for each study year *j* are added to obtain the expected number of crashes for the study period.

Site-Specific EB Method

The expected average crash frequency for a site for year *j* can be estimated using the following equation. In this application, the year of interest is year *j* and the reference year *r* is any one year in the crash period (by convention, the reference year is typically selected to be the first year in the crash period).

$$N_{e,w(i),x(i),y,z,j} = N_{e,w(i),x(i),y,z,r} \times \frac{N_{p,w(i),x(i),y,z,j}}{N_{p,w(i),x(i),y,z,r}} \qquad \text{(B-23)}$$

Where:

$N_{e,\,w(i),\,x(i),\,y,\,z,\,j}$ = expected average crash frequency for site *i* with site type *w(i)* and year *j* (includes cross section or control type *x(i)* for crash type *y* and severity *z*) (crashes/yr);

$N_{e,\,w(i),\,x(i),\,y,\,z,\,r}$ = expected average crash frequency for site *i* with site type *w(i)* and reference year *r* (includes cross section or control type *x(i)* for crash type *y* and severity *z*) (crashes/yr);

$N_{p,\,w(i),\,x(i),\,y,\,z,\,j}$ = predicted average crash frequency for site *i* with site type *w(i)* and year *j* (includes cross section or control type *x(i)* for crash type *y* and severity *z*) (crashes/yr); and

$N_{p,\, w(i),\, x(i),\, y,\, z,\, r}$ = predicted average crash frequency for site *i* with site type *w(i)* and reference year *r* (includes cross section or control type *x(i)* for crash type *y* and severity *z*) (crashes/yr).

The expected average crash frequency is specific to a given site type *w*, cross section or control type *x*, crash type *y*, and severity *z*. As a result, the variables used in this equation must all be consistent in their representation of site type, cross section or control type, crash type, and severity.

The predicted average crash frequencies used in Equation B-23 are obtained from the appropriate predictive model described in Chapter 18 or 19. The expected average crash frequency is obtained from Equation B-5.

The expected number of crashes for a site for a specified study period is computed using the following equation.

$$N^*_{e,w(i),x(i),y,z} = \sum_{j=1}^{n_s} N_{e,w(i),x(i),y,z,j} \tag{B-24}$$

Where:

$N^*_{e,\, w(i),\, x(i),\, y,\, z}$ = total expected number of crashes for site *i* with site type *w(i)* and all years in the study period (includes cross section or control type *x(i)* for crash type *y* and severity *z*) (crashes); and

n_s = number of years in the study period (yr).

The expected number of crashes for all sites for a specified study period is computed using the following equation.

$$N^*_{e,aS,ac,at,as} = \sum_{i}^{n_{sites}} \sum_{k}^{n_{crash\ types}} \sum_{l}^{n_{severities}} N^*_{e,w(i),x(i),y(k),z(l)} \tag{B-25}$$

Where:

$N^*_{e,\, aS,\, ac,\, at,\, as}$ = total expected number of crashes for all sites *aS* and all years in the study period (includes all cross sections *ac*, all crash types *at*, and all severities *as*) (crashes/yr);

n_{sites} = number of sites;

$n_{crash\ types}$ = number of crash types; and

$n_{severities}$ = number of severities.

Project-Level EB Method

The following procedure describes the sequence of calculations necessary to adjust the estimate of expected average crash frequency to a future year (or years). To facilitate the presentation of this procedure, the equations shown in this subsection have subscripts denoting fatal-and-injury *fi* crashes of all crash types (i.e., multiple-vehicle and single-vehicle fatal-and-injury crashes combined). The conversion of these equations so that they are applicable to property-damage-only crashes (or crashes of a specific crash type) requires only the substitution of the appropriate subscripts.

The expected average crash frequency for all sites for year *j* can be estimated using the following equation. In this application, the year of interest is year *j* and the reference year *r* is any one year in the crash period (by convention, the reference year is typically selected to be the first year in the crash period).

$$N_{e,aS,ac,at,fi,j} = N_{e,aS,ac,at,fi,r} \times \frac{N_{p,aS,ac,at,fi,j}}{N_{p,aS,ac,at,fi,r}} \tag{B-26}$$

Where:

$N_{e,\,aS,\,ac,\,at,\,fi,\,j}$ = expected average crash frequency for all sites *aS* and year *j* (includes all cross sections *ac* and fatal-and-injury crashes *fi* of all crash types *at*) (crashes/yr);

$N_{e,\,aS,\,ac,\,at,\,fi,\,r}$ = expected average crash frequency for all sites *aS* and reference year *r* (includes all cross sections *ac* and fatal-and-injury crashes *fi* of all crash types *at*) (crashes/yr);

$N_{p,\,aS,\,ac,\,at,\,fi,\,j}$ = predicted average crash frequency for all sites *aS* and year *j* (includes all cross sections *ac* and fatal-and-injury crashes *fi* of all crash types *at*) (crashes/yr); and

$N_{p,\,aS,\,ac,\,at,\,fi,\,r}$ = predicted average crash frequency for all sites *aS* and reference year *r* (includes all cross sections *ac* and fatal-and-injury crashes *fi* of all crash types *at*) (crashes/yr).

The predicted average crash frequencies used in Equation B-26 are computed using Equation B-16. The expected average crash frequency is obtained from Equation B-17.

The expected number of crashes for all sites for a specified study period is computed using the following equation.

$$N^{*}_{e,aS,ac,at,fi} = \sum_{j=1}^{n_s} N_{e,aS,ac,at,fi,j} \tag{B-27}$$

Where:

$N^{*}_{e,\,aS,\,ac,\,at,\,fi}$ = total expected number of crashes for all sites *aS* and all years in the study period (includes all cross sections *ac* and fatal-and-injury crashes *fi* of all crash types *at*) (crashes.

B.2.7. EB Method for Segments with an Odd Number of Lanes

Most roadway cross sections have an even number of through traffic lanes. As a result, researchers can typically acquire data for segments with even numbers of lanes in sufficient number to permit the development of statistically valid predictive models. On the other hand, some roadways do exist with an odd number of through lanes. The development of statistically valid models for these cross sections is sometimes not possible due to inadequate sample size.

This section describes a procedure for evaluating a segment of interest that has a cross section with an odd number of through lanes. This procedure can be used when a predictive model is not available for the specified cross section. It is described in the form of supplemental equations that are used in the steps of the predictive method. The step numbers of the procedure match those of the predictive method to which they apply. The procedure is viable if the following checks are satisfied.

- The segment has X total lanes that represent Y lanes in one direction and Z lanes in the opposite direction (i.e., $X = Y + Z$) and Y is not equal to Z.
- The predictive model for segments includes an SPF for $2{\times}Y$ lanes.
- The predictive model for segments includes an SPF for $2{\times}Z$ lanes.

If these checks are satisfied, then the procedure can be applied.

Step 9—For the selected site, determine and apply the appropriate SPF.
The applicable predictive model is identified from the appropriate chapter. The site of interest is determined to have a site type *w* with an *X*-lane cross section, and the analysis is focused on crash type *y* and severity *z*.

Select an SPF for the subject site based on it being applicable to a cross section of 2×*Y* lanes. Select a second SPF for the subject site based on it being applicable to a cross section of 2×*Z* lanes. The best estimate of the predicted average crash frequency for base conditions is computed as the average of the estimates from the two SPFs. This calculation is shown using the following equation.

$$N_{spf,w,X,y,z,j} = \frac{N_{spf,w,2Y,y,z,j} + N_{spf,w,2Z,y,z,j}}{2} \tag{B-28}$$

Where:

$N_{spf,w,n,y,z,j}$ = predicted average crash frequency for year *j* determined for base conditions of the SPF developed for site type *w*, *n*-lane cross section ($n = X$, $2Y$, $2Z$), crash type *y*, and severity *z* (crashes/yr).

Step 10—Multiply the result obtained in Step 9 by the appropriate CMFs.
The predictive model is used to compute the predicted average crash frequency for the subject site. The general form of this model is shown in the equation below. The specific CMFs and calibration factor are obtained from the appropriate chapter.

$$N_{p,w,X,y,z,j} = N_{spf,w,X,y,z,j} \times \left(CMF_{1,w,X,y,z} \times CMF_{2,w,X,y,z} \times \ldots \times CMF_{m,w,X,y,z}\right) \times C_{w,X,y,z} \tag{B-29}$$

Where:

$N_{p,w,X,y,z,j}$ = predicted average crash frequency for year *j* for site type *w*, *X*-lane cross section, crash type *y,* and severity *z* (crashes/yr);

$CMF_{m,w,X,y,z}$ = crash modification factors specific to site type *w*, *X*-lane cross section, crash type *y,* and severity *z* for specific geometric design and traffic control features *m*; and

$C_{w,X,y,z}$ = calibration factor to adjust SPF for local conditions for site type *w*, *X*-lane cross section, crash type *y,* and severity *z*.

Step 13—Apply site-specific EB Method (if applicable) and apply SDFs.
If the EB Method is used in the predictive method, then the variance of the predicted average crash frequency is computed in this step using the following equation.

$$V_{p,w,X,y,z} = \left(\sqrt{k_{w,2Y,y,z} \times \left[\sum_{j=1}^{n_c}\left(0.5 \times N_{p,w,2Y,y,z,j}\right)\right]^2} + \sqrt{k_{w,2Z,y,z} \times \left[\sum_{j=1}^{n_c}\left(0.5 \times N_{p,w,2Z,y,z,j}\right)\right]^2}\right)^2 \tag{B-30}$$

with

$$N_{p,w,2Y,y,z,j} = N_{spf,w,2Y,y,z,j} \times \left(CMF_{1,w,X,y,z} \times CMF_{2,w,X,y,z} \times \ldots \times CMF_{m,w,X,y,z}\right) \times C_{w,X,y,z} \tag{B-31}$$

$$N_{p,w,2Z,y,z,j} = N_{spf,w,2Z,y,z,j} \times \left(CMF_{1,w,X,y,z} \times CMF_{2,w,X,y,z} \times \ldots \times CMF_{m,w,X,y,z}\right) \times C_{w,X,y,z} \tag{B-32}$$

Where:

$V_{p, w, X, y, z}$ = variance of the predicted average crash frequency for site type *w*, *X*-lane cross section, crash type *y*, and severity *z* (crashes^2/yr^2);

$k_{w, n, y, z}$ = overdispersion parameter for site type *w*, *n*-lane cross section ($n = X$, $2Y$, $2Z$), crash type *y*, and severity *z*;

n_c = number of years in the crash period (yr); and

$N_{p, w, n, y, z, j}$ = predicted average crash frequency for a year *j* for site type *w*, *n*-lane cross section ($n = X$, $2Y$, $2Z$), crash type *y*, and severity *z* (crashes/yr).

The overdispersion parameters used in Equation B-30 are obtained from the same predictive model as used to estimate the predicted average crash frequencies.

An overdispersion parameter is needed to apply the EB Method. An equivalent overdispersion parameter that is associated with the predicted average crash frequency from Equation B-29 is computed using the following equation.

$$k^*_{w,X,y,z} = \frac{V_{p,w,X,y,z}}{\left[\sum_{j=1}^{n_c} N_{p,w,X,y,z,j}\right]^2} \tag{B-33}$$

Where:

$k^*_{w, X, y, z}$ = effective overdispersion parameter for site type *w*, *X*-lane cross section, crash type *y*, and severity *z*.

The effective overdispersion parameter computed using Equation B-33 is used in Equation B-5 of the site-specific EB Method described in Section B.2.4.

Step 15—Apply the project-level EB Method (if applicable).
The effective overdispersion parameter computed using Equation B-33 is used in Step 2 of the project-level EB Method described in Section B.2.5.

B.3. REFERENCES

(1) Hauer, E. *Observational Before-After Studies in Road Safety*. Pergamon Press, Elsevier Ltd., Oxford, United Kingdom, 1997.

(2) Hauer, E., D. Harwood, F. Council, and M. Griffith. “Estimating Safety by the Empirical Bayes Method.” In *Transportation Research Record 1784*. TRB, National Research Council, Washington, DC, 2002, pp. 126–131.