CHAPTER 4
EXTRACTION OF DATA FROM THE GES DATABASE

4.1 Introduction

In order to develop the accident rates to be used as input for the safety model, the crash data needed to be extracted from the GES database. This chapter presents the procedure for extracting the data, the GES estimation methods and the reliability of the estimates. The different accident categories and configurations will be presented since the data presented in this chapter comprise the input for the safety model that will be described in chapter 5. Finally, the raw accident rate data will be presented along with a special focus on rear-end crashes.

4.2 The GES Database as a useful tool

GES serves as a national data source for the National Traffic Safety Administration (NHTSA) in order to support highway safety research and develop a variety of programs. GES addresses several crucial questions related to traffic safety. Regarding traffic crashes, information is provided as of when and how often do crashes occur. Also where the crashes occur and what are the events during the crash.

Additional information is provided for the vehicles involved in the crashes. These include the types of vehicles and their action prior to the crash, as well as what area of the vehicle was damaged. Finally, for the people involved in the crashes, the GES provides information such as how many drivers, pedestrians and pedal-cyclists were involved, how severely they were injured, what was their age and sex etc.

GES provides a means to track trends in these national level estimates. Some background information will be provided regarding how the data are derived and how reliable they are. GES includes crashes from all severities, from property damage only, to fatal. Also, all motor vehicles are included. As mentioned earlier, a total number of 6.8 million crashes were included in the 1996 GES file that was used for developing the accident trends and rates for the safety model.

4.2.1 GES Estimation Methods

Before presenting the actual crash frequencies as they were extracted from the GES database, it is important to describe how these data are extracted and how they are grouped in order to produce the crash rates. Before each file is published every year, NHTSA groups the crash data and constructs the file in different stages. The sample design of the GES consists of the following three stages:

Stage 1: Primary Sampling Unit's (PSU's) are selected and they can be a central city, a county or a group of counties. The PSU's are assigned to different geographic and urbanization classes such as the geographic region, i.e. Northeast, South, Central and West. The selection of these PSU's is reviewed every year. In 1996, the U.S was divided into 1,195 PSUs that were grouped into 12 categories.

Stage 2: in this stage, the Police Jurisdictions (PJ) are selected within the Primary Sampling Units. A probability sample of PJ's within each PSU was selected with probability proportional to the number of crashes investigated by that PJ. The more crashes investigated by the PJ, the greater the likelihood that jurisdiction would be chosen.

Stage 3: In this stage the selection of the Police Accident Reports within the Police Jurisdictions will be done. This third and final stage of selection is the only stage where new sampling units are introduced each year and year-to-year variation will occur. In the GES 1996 file there were approximately 56,000 Police Accident Reports.

A list of PAR's is made with each PAR being assigned to different groups or strata based on the review criteria. In this third and final stage the Police Accident Reports are grouped into one of four groups by the data collector. In Group 1, all crashes involving a towed passenger car are grouped. In Group 2, crashes involving a medium or heavy truck are grouped. In Group 3, all crashes not involving a towed passenger vehicle or medium or heavy truck are grouped together and finally in Group 4, all other crashes are considered.

The GES follows a procedure in order to calculate estimates of national crash characteristics. The data from each Police Accident Report have to be weighted to reflect the probability of selection. As explained above, the GES sampling process consists of three stages therefore the final weight is the product of the probability of selection at each of the stages. Because of this fact, the extracted crash frequencies from the database are non-integer numbers and had to be rounded off to reflect the number of crashes.

P_i is the probability of including the ith PAR in the sample and is the product of the probabilities of selection at each stage of sampling and can be expressed as in equation 1 below:

$$P_i = P_{psu\ selection} * P_{pj\ selection} * P_{par\ selection} \qquad [1]$$

4.2.2 Reliability of Estimates

Due to the fact that the GES data are derived from an annual probability sample of more than 55,000 Police Accident Reports (PAR's) and not from a census of all 6.8 police reported million crashes in the U.S., the estimates are subject to sampling errors, as well as non-sampling errors. Table A-3 in Appendix A illustrates the Standard Error (SE) of the GES estimates for the year 1996. The estimates have a different SE for crash estimates and vehicle estimates. These come from a different file as well, i.e. crash estimates from the accident file and vehicle estimates from the vehicle file. The formulae used in each case are also provided and the derivation of them belongs to the research team of NHTSA.

It is important to note that the GES variables are revised every two years in order for the database to be more accurate and complete to meet challenging transportation safety issues. The above-mentioned sampling errors are revised every year in order for them to be adjusted to the new variables and procedures within the database.

4.2.3 Extraction Procedures and SAS Data Sets

The Statistical Analysis System (SAS) software package was used to extract the data from the GES database. SAS file *main1.sas* is presented in Appendix A and served, as the master program that aided in identifying the variables needed from each file. These variables will be described in more detail in the next section.

The main objective of this program was to create a library to store the extracted crash data and create a working data set from each of the three files. As mentioned in the literature review, the GES database includes three files, the accident, the vehicle and the person file. At each data set, the variables from each file were chosen and finally the files were merged by case number (casenum in the database) and primary sampling units. This was done due to the fact that the accident file refers to crashes and the vehicle to vehicles. *main1.sas* was run every time data needed to be extracted from the database.

In order to extract a specific set of data in a tabular format, smaller SAS programs were created. One sample program (*spd.sas*) is presented in Appendix A. This program extracts the number of crashes for all speed limits in the database by time-of-day. Figure 4-1 below illustrates the complete procedure used in extracting all data in tables. Ellipses describe the processing steps and squares the data sets and tabulated results. The purpose of the PROC step is to process the data for extraction. The GES variable WEIGHT used in all programs in order to produce the national estimates from the 56,000 Police Accident Reports. Finally, the results of analysis were processed using Excel spreadsheets.

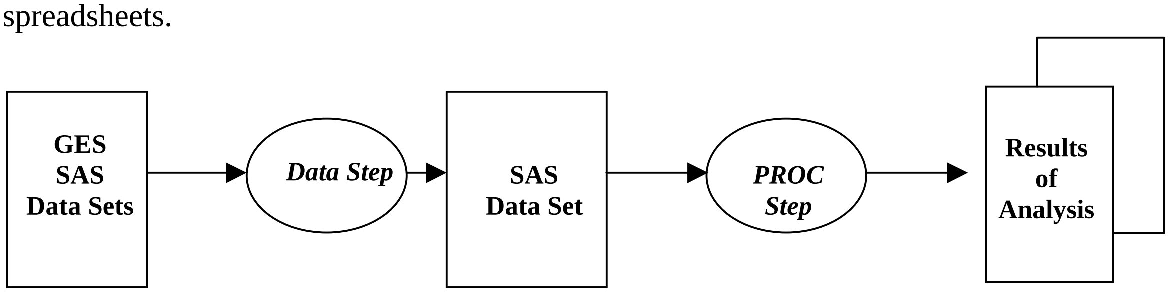

Figure 4-1 GES/SAS Extraction Procedure

4.3 GES Variables

In chapter 3 some of the important variables for the safety model were discussed together with a presentation of some of the most recent trends in transportation safety in the United States. The variables drawn directly from the GES database 1996 files are presented in a tabular form in Table A-1 in Appendix A.

These variables are:

- Hour the crash occurred (time-of-day)

- Speed Limit: the actual posted speed limit in miles per hour
- Maximum Injury Severity: most severe injury of all persons in the crash
- Vehicle Severity: the severity of the vehicle damage
- Manner of Collision: the orientation of the vehicles in a collision
- Accident Type: the pre-crash situation in 6 categories and 14 configurations

4.3.1 First Harmful Event and Manner of Collision

Before proceeding in the description of the accident type variable that was used for the development of the crash rates, it is important to present a description for variables such as the first harmful event and the manner of collision. These two variables are located in the accident and vehicle files of the database and belong to the variables in the GES database that attempt to characterize the manner of collision and the actual crash.

The first harmful event is one of the basic variables that characterize an accident in the GES database. Every motor vehicle traffic accident consists of a series of events. In classification by type, one of the events must be selected before further classification can be made. For uniformity in classification, the "First Harmful Event" is the "first" property damage or injury-producing event that can be determined to have happened in the accident. It is important to mention the “first harmful event” at this point because of the fact that for a crash to occur there is a pre-crash state, the first harmful event and the actual crash at the end.

The manner of collision variable appears in the accident file. This variable indicates the orientation of the vehicles in a collision. The GES database includes the following manners of collision:

1. *Rear-End:* when a collision occurs between the rear of one vehicle and the front of another vehicle.
2. *Head-On:* when a collision occurs between the front end of one vehicle and the front end of another vehicle.
3. *Rear-To-Rear:* when a collision occurs between the rear of one vehicle and the rear of another vehicle.

4. *Angle:* when the impact configuration is known but cannot be classified with any other element. Included here are also end-swipes. Can be characterized by perpendicular movements on intersections.
5. *Sideswipe, Same Direction:* when the Police Accident Report, reports that a sideswipe occurred while the two vehicles were traveling in the same directions. Such conditions might include lane changing.
6. *Sideswipe, Opposite Direction:* when the Police Accident Report reports that a sideswipe occurred while the two vehicles were travelling in opposite directions. A left turn movement can serve as an example for this type.
7. *Other*: when the First Harmful Event involves a vehicle that is parked off the roadway and an open door over the roadway is struck by a motor vehicle in transport.
8. *Unknown*

The above-mentioned eight manners of collision are the official crash types that the GES identifies. As it will be described later in this chapter, the pre-crash state variables will be used. Each of the 80 pre-crash states is furthermore grouped in a more general type, for example, the angle crashes can include sideswipes and even some forward impact crashes.

4.3.2 Speed Limit and Time-of-Day

The speed limit values extracted from the database aided in identifying the facility type since the database did not provide such a variable to indicate the type of facility that the crash occurred on. However, it must be noted that there is a variable in the GES, providing frequencies for crashes that occurred on the National Highway System (NHS) and Interstates. The Interstate facilities are identified as facilities with speed limits of 40 mph and higher.

As shown in Table A-1 in Appendix A, the actual posted speed limit in miles per hour was reported in increments of 5mph in the range of 0-75mph. For zero speed limits the crashes were categorized as ones that occurred in a parking lot, alley etc.

In order to identify the type of facility on which a crash occurred, the speed limit variable from the GES database was used with the assumptions presented in Table 4-1 below:

Table 4-1 Facility types and corresponding speed limits

Facility Type	Speed Limit (mph)
Locals	5-25
Collectors	30-35
Minor Arterials	40-45
Principal Arterials & Interstates	50-75

In addition to the speed limit, another variable used was the hour the crashes occurred. The variable HOUR was extracted from the accident file of the database and military time was used (0-24 hrs). A description and all time intervals covered for this variable can be found in Table A-1 in Appendix A.

4.3.3 Accident Severity

There are three main categories of accident severity; property damage only, injury severity and fatality. All these categories are described in the GES files and were used in grouping the accident rates by their injury level. Property Damage Only (PDO) crashes were extracted using the variable for Vehicle Severity from the vehicle file. The four levels of vehicle damage used were:

i. No Damage
ii. Minor
iii. Moderate
iv. Severe (major)

Injury severity crashes were extracted using the maximum severity variable from the accident file of the GES. The five levels considered were:

i. No Injury
ii. Possible Injury
iii. Non-incapacitating Injury
iv. Incapacitating Injury
v. Fatal Injury

4.3.4 Accident Type

The vehicle file of the GES database provides a more detailed description for the pre-crash situation. Data obtained from the first harmful event and the manner of collision variables provide information for 80 different accident types. A description of the first harmful event and the manner of collision variables was presented earlier. Variable 23 of the vehicle file includes six categories and fourteen major accident configurations, as they will be described below. This variable is used for categorizing the collisions of drivers involved in accidents. A collision is defined here as the first harmful event in an accident between a vehicle and some object, accompanied by property damage or human injury. The object maybe another vehicle, a person, an animal, a fixed object, the road surface, or the ground. If the first collision is a rollover, the impact is with the ground or road surface. The collision may also involve plowing into soft ground, if severe vehicle deceleration results in damage or injury. A road departure without damage or injury is not defined as a collision.

To determine the proper Accident Type (AT), the following three-step decision process is followed:

Step 1: the appropriate category is determined.

Step 2: the appropriate configuration is determined.

Step 3: the specific accident type is determined.

Each category is defined by an accident configuration(s). The categories are divided into six sections and are described as follows:

Category I. Single Driver

The first harmful event involves a collision between an in-transport vehicle and an object. A harmful event involving two in-transport vehicles is excluded from this category.

Category II. Same Traffic Way, Same Direction

The first harmful event occurred while both vehicles were traveling in the same direction on the same traffic way.

Category III. Same Traffic Way, Opposite Direction

The first harmful event occurred while both vehicles were traveling in the opposite directions on the same traffic way.

Category IV. Change Traffic Way, Vehicle Turning

The first harmful event occurred when the vehicle is either turning or merging while attempting to change from one traffic way to another traffic way. Traffic way for this variable is loosely defined to include driveways, alleys and parking lots when a vehicle is either entering or exiting a traffic way.

Category V. Intersecting Paths

The first harmful event involves situations where vehicle trajectories intersect.

Category VI. Miscellaneous

The first harmful event involves an accident type, which cannot be described in categories I-V and thus is included in this category. All the above categories are furthermore described in Table B-1 in Appendix B. Figures illustrating each accident type are also presented in Figure B-1 in Appendix B. Figure B-1 illustrates how each variable is used for a specific accident type. A rear-end sample case is presented in the next section.

4.3.5 Rear-End Accident Types

One of the most common accident types is the rear-end type. Almost 28% of the total number of police reported crashes are rear-ends (U.S. DOT, 1997). Rear-end crashes is potentially the dominant crash type to be impacted by signal coordination. In Figure 4-2 below, two types of rear end crash configurations are illustrated.

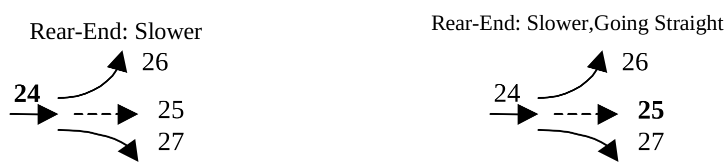

Figure 4-2: Rear-End Accident Types

For each particular variable different crash frequencies can be extracted. Rear-end types for example can be of many different configurations. In the figure above, the rear-end slower pre-crash state is illustrated. There are four different variables within each of the rear-end types, if the vehicle in the front was going straight and moving slower than the vehicle behind, then the variable number 25 is used. If one chooses the variable with number 24, this means the all the crash frequencies of rear-end (slower) type will be extracted regardless if the vehicle in the front was turning left, right or was going straight. By extracting the data in this manner, one can easily group all these types and produce statistics for all the pre-crash state configurations in the database. Grouping the data from this variable and comparing them with the manner of collision variable (e.g. rear-end as a manner of collision), the results are very similar taking into account the standard error produced during the extraction process. Some accident rates regarding rear-end crashes are presented in the last section of this chapter.

4.4 Development of the Accident Rates

The previous sections provided a background on how the crash data were extracted from the database and the variables used. The rates hereby presented were derived using the extracted data and the corresponding vehicle miles traveled from Table 3.5 in chapter 3. Equation 1 below was used to develop the accident rates to be used as input to the model.

Accident Rate = GES Crash Frequency / Facility based VMT [1]

The total crash rate was computed by speed limit and time-of-day. In addition, rates were computed for each of the 14 configurations (or crash types) as shown in Table 4-2 below. Several crosschecks were performed in order to ensure consistency across the rates and produce charts illustrating the several trends.

In order for the model to be efficient in computing the accident risk, the total number of 80 pre-crash states were grouped in 14 accident types as shown in Table 4-2 below.

At this point it is important to distinguish the differences between configurations and accident types. As mentioned earlier, there are 80 pre-crash states in Variable 23 of the GES vehicle file. These types are grouped in 14 different configurations (or categories) where a configuration means for example, a right roadside departure, rear-end crashes, angle crashes or side swipes. More detailed information and description of all the pre-crash states and configurations, is included in Appendix B (Table B-1).

Table 4-2 Crash Types as an input to the model

Crash Type (Label)	Description
1	**Single Driver - Right Roadside Departure**
2	**Single Driver - Left Roadside Departure**
3	**Single Driver - Forward Impact**
4	**Same Traffic Way and Same Direction - Rear-end**
5	**Same Traffic Way and Same Direction - Forward Impact**
6	**Same Traffic Way and Same Direction - Sideswipe/Angle**
7	**Same Traffic Way and Opposite Direction - Head-on**
8	**Same Traffic Way and Opposite Direction - Forward Impact**
9	**Same Traffic Way and Opposite Direction - Sideswipe/Angle**
10	**Change Traffic Way and Vehicle Turning – Turn Across Path**
11	**Change Traffic Way and Vehicle Turning – Turn Input Path**
12	**Intersecting Paths – Perpendicular Crash**
13	**Backing Vehicle**
14	**Other or Unknown**
15	**Total Accident Rate**

The fourteen crash types correspond to the configurations set by the GES database as they are divided in six categories and fourteen configurations. These types will be used in greater detail for the calculation of the accident risk in chapters 5 and 6.

4.4.1 Crash rates by time-of-day

The distribution of accidents in terms of an accident rate per million vehicle miles was examined using the database results before creating the input rates for the safety model. Figure 4-3 below illustrates this trend of accident rates per time-of-day. The rates are in the range of 50-100 accidents per million VMT for the interval of 6 am to midnight. However, the rates are higher during the early morning hours due to the lower volume on the roads during these times.

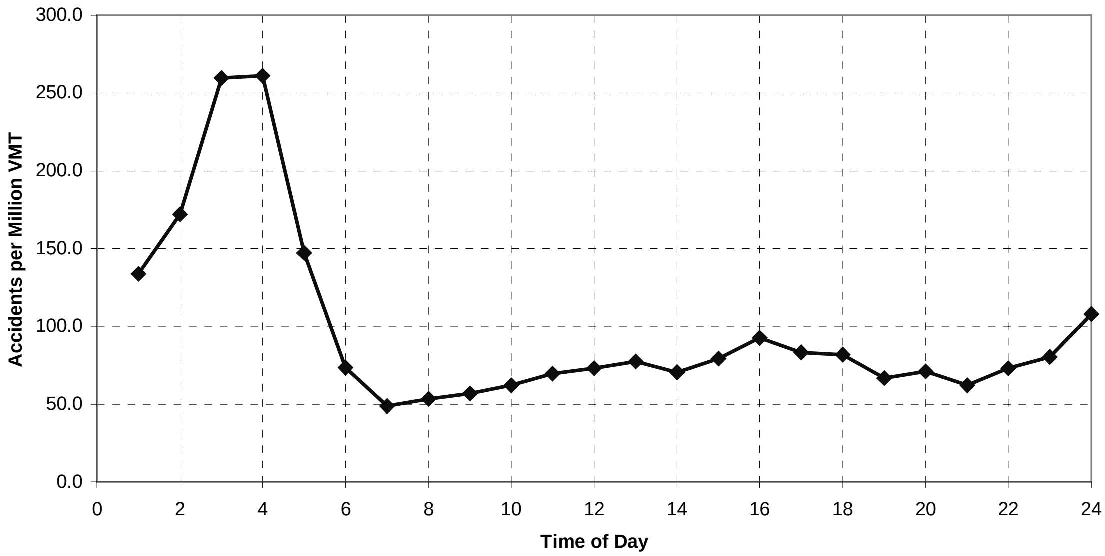

Figure 4-3 Distribution of Accident Rates per time-of-day

4.4.2 Crash rates by time-of-day and accident type

By grouping all the accidents in fourteen types, rates can be easily produced for each of these types. Grouped rates are illustrated in Table 4-3 below. The speed limit is in miles per hour and the rates in crashes per million vehicle miles traveled according to the GES database and the highway statistics that were used for the development of these rates. It can be observed that for crash types 1 and 2 (left and right roadside departures) the rates were higher. The same trend occurred for most of the same traffic way opposite direction crash types and more specifically, cases such as head-on, sideswipe/angle and forward impact collisions. Speed limits are in miles per hour and rates in crashes per million-VMT.

Table 4-3 Crash rates by accident type and speed limit (in crashes per million VMT)

Speed	Accident Type														
Limit	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
0	0.40	0.61	1.43	1.32	0.00	1.24	0.12	0.60	0.31	1.00	0.87	0.48	0.50	2.30	11.1
5	0.00	0.00	1.38	3.72	0.00	0.00	0.00	0.00	0.69	0.00	2.09	2.27	0.00	0.00	10.1
10	0.65	0.17	1.72	0.86	0.00	0.17	0.15	0.00	0.55	1.04	2.16	0.56	0.00	1.21	9.24
15	0.73	0.47	1.27	1.59	0.00	0.42	0.18	0.23	0.51	0.86	0.97	0.41	0.09	0.68	8.40
20	0.57	0.57	1.07	1.08	0.06	0.40	0.12	0.36	0.46	0.97	0.75	0.57	0.05	0.61	7.64
25	0.55	0.45	0.92	1.15	0.06	0.30	0.08	0.18	0.45	0.65	0.86	0.46	0.10	0.71	6.94
30	0.49	0.38	0.92	1.05	0.05	0.32	0.06	0.14	0.44	0.57	0.73	0.45	0.06	0.66	6.31
35	0.37	0.32	0.65	0.76	0.03	0.22	0.05	0.12	0.31	0.43	0.52	0.32	0.06	0.45	4.60
40	0.37	0.30	0.67	0.70	0.04	0.22	0.04	0.10	0.31	0.45	0.58	0.28	0.06	0.46	4.56
45	0.37	0.26	0.58	0.69	0.02	0.21	0.05	0.08	0.28	0.35	0.50	0.29	0.05	0.40	4.14
50	0.25	0.19	0.36	0.36	0.02	0.13	0.03	0.08	0.16	0.23	0.31	0.21	0.04	0.24	2.60
55	0.19	0.15	0.33	0.37	0.01	0.11	0.03	0.06	0.16	0.23	0.29	0.17	0.03	0.23	2.37
60	0.07	0.07	0.19	0.23	0.01	0.06	0.01	0.02	0.07	0.11	0.13	0.10	0.02	0.14	1.21
65	0.07	0.06	0.18	0.20	0.01	0.05	0.01	0.02	0.07	0.10	0.14	0.09	0.01	0.10	1.10
70	0.09	0.06	0.17	0.14	0.00	0.04	0.02	0.02	0.07	0.12	0.10	0.06	0.01	0.11	1.00
75	0.08	0.05	0.11	0.15	0.00	0.07	0.00	0.01	0.04	0.07	0.16	0.09	0.00	0.09	0.91

4.4.3 Crash rates for injury severity accidents by speed limit and accident type

Five injury severity levels were considered for the development of rates in this category. The rates are shown in Table B-2 in Appendix B and they are tabulated in terms of accident type and speed limit (due to its large size this table is included in its whole in the Appendix). As expected, the fatal injury crashes have the lowest rates due to the relatively small number of such cases as explained in chapter 3. Higher rates were detected for no injury crashes with those rates being almost stable for all accident types. Accident type 4 (rear-end crashes) are shown in italics in the table. The lower crash rates for rear-end injury severity crashes are found on lower speed limits, i.e. facility types such as minor arterials, collectors and locals. However, the rates on higher speed limits cannot be neglected, as this crash type is very common for facilities like principal arterials and interstates.

4.4.4 Crash rates for damage severity accidents by speed limit and accident type

A similar approach like the injury severity crashes was followed for the damage severity cases. Damage severity crashes of four levels were analyzed and rates were produced. Vehicle damage crashes account for almost 60% of the total number of police reported accidents every year, hence such an analysis is very important. Table B-3 in Appendix B illustrates the rates produced using the GES database and the highway statistics. The analysis yielded some interesting results in terms of vehicle damage level.

Minor and moderate (functional) damage crash rates yielded more observations, i.e. all accident types and speed limits had a significant amount of crashes in order to produce the rates. As expected, higher rates were observed for higher speed limits with only a few exceptions. Rear-end crash rates are also shown in italics and it can be observed that the rates are very low for no damage whereas for higher damage levels the rates increase significantly with the result to have high accident rates for rear-end crashes with severe damage. At this point it is important to emphasize the number of unknown crash types found in the database in both severity levels (injury and property damage). Despite the efforts by the creators of the GES database to minimize this fact, a significant amount of crashes with unknown type exist in the database.

4.5 Rear-End Crashes - Case Study

Using the extracted frequencies from the database and traffic volume data from the Scottsdale/Rural road corridor in Phoenix, the rear-end crashes (see Category II, configuration D in Table B-1 in Appendix B) were further examined. The crash rates were produced based on the actual frequencies from the GES database and tube counts from Scottsdale/Rural road. The speed limit on this facility is 45 mph and therefore according to the definitions set in chapter 3 this is a principal arterial. It is particularly important to focus on rear-end crashes, especially after closely examining the different configurations set within the database. As shown in section 4.3.5 rear-ends can be of several types and fall into the same traffic way same direction category of accidents.

Traffic signal coordination is hypothesized to smoothen traffic flow, thus reducing the risk for having a rear-end crash especially when approaching an intersection.

Figure 4-4 below illustrates the tube counts obtained for the Scottsdale/Rural road corridor. The hourly volume in vehicles per hour is presented versus the time-of-day. The three periods with higher volumes are the morning and afternoon peak, together with the midday peak. The low volumes of the early morning hours are also observed, something that will affect the crash rates.

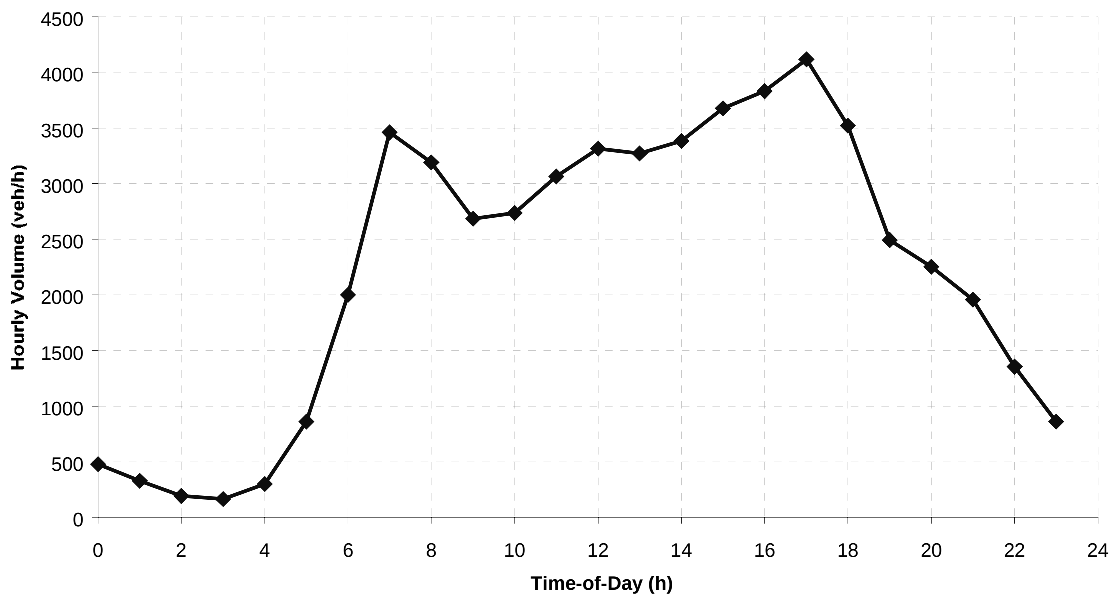

Figure 4-4 Scottsdale/Rural Tube Counts

Using the crash frequencies obtained from the GES database in general terms, the chart of rear-end crashes for speed limit of 45 mph was plotted. Figure 4-5 below presents the hourly rear-end crash types grouped together as a function of the time-of-day they occurred. The afternoon peak period tends to experience most of these crashes. This is somehow expected because a speed limit of 45mph identifies a principal arterial, a facility type that experiences heavy traffic during the afternoon rush hours.

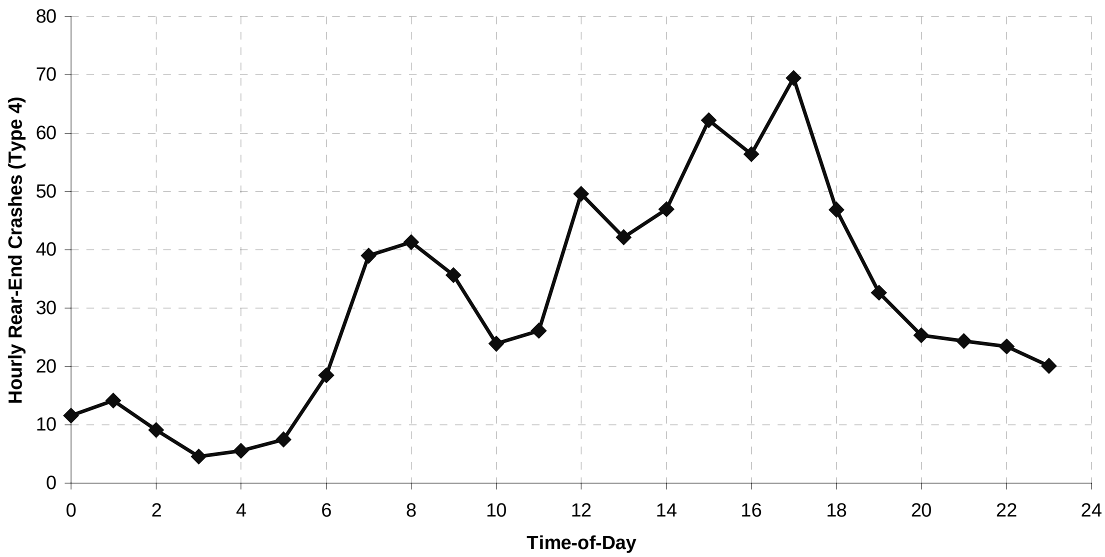

Figure 4-5 Rear-End Crashes –45 mph

The next step was to produce some crash rates for the rear-end crashes that occurred on principal arterials (speed limit of 45 mph). The rates are hereby presented below in Figures 4-6 and 4-7. Figure 4-6 illustrates the crash rates as a function of the time-of-day, provided by the GES database. The crash rate appears to be higher during the early morning hours mainly due to the limited traffic volume on the road. For the rest of the day starting from 5 am, the rear-end crash rates follow a steady pace with some minor increase experienced during the midday and afternoon peak hours. Figure 4-7 presents the rear-end crash rate as a function of volume, by using the actual tube counts from Scottsdale/Rural road and the highway statistics data (from chapter 3, VMT for principal arterials). The data can be characterized by a negative exponential distribution, thus yielding lower crash rates for higher volumes and higher crash rates for very low volumes. However, the trend appears to be steady for most of the volumes that are higher than 400 vehicles per hour. The shape of the chart in Figure 4-7 is consistent with the U-shape presented in the literature review (developed by Zhou and Sissiopiku). The crash rates are observed to be steady during the day, from 6 am to 10 p.m. An increase is also observed during the early morning hours mainly due to the differences in the traffic volumes during these hours.

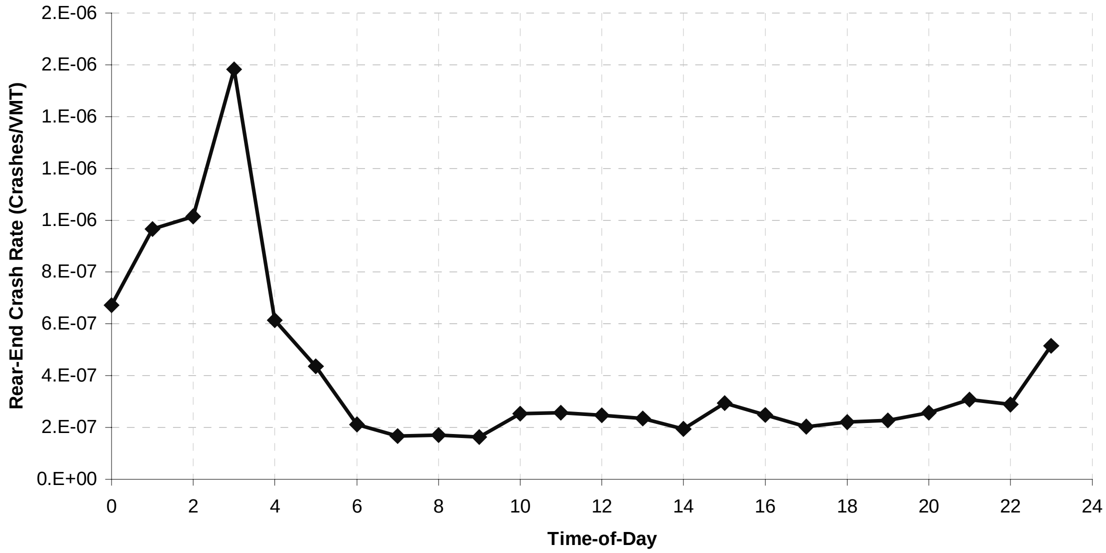

Figure 4-6 Rear-End Crash Rates – 45 mph

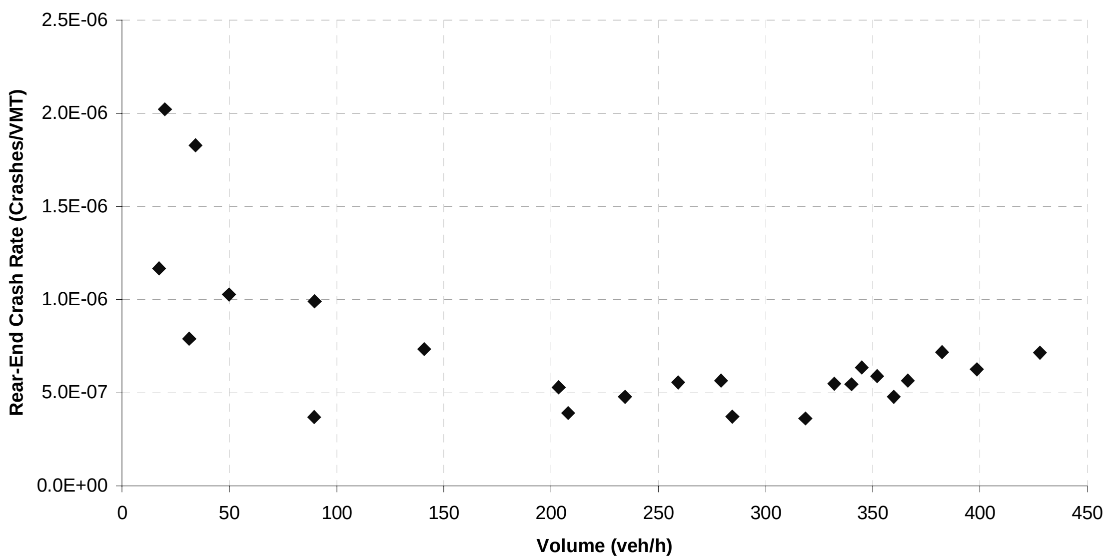

Figure 4-7 Rear-End Crash Rate As a Function of Volume

4.6 Summary

This chapter focused on the General Estimates System database, its extracting and estimation methods as well as the reliability of the estimates. This was particularly important since the rates produced included some assumptions. Another factor that stressed the need for such an analysis was the fact that highway statistics were used for the development of the rates. Annual vehicle miles traveled used as the exposure in calculating the number of crashes per vehicle miles traveled were grouped in terms of facility type but not in an urban or rural categorization.

Another important aspect of the study was also presented in this chapter; the accident type was one of the most important variables used for developing the rates. The difference in accident type and manner of collision is explained and all the 80 pre-crash states are presented in Appendix B along with explanatory figures. Rear-end cases were emphasized as well as the different injury levels were presented. The rear-end case study in Scottsdale/Rural road in Phoenix was presented along with some important results for the outcome.

The development of the rates in this chapter prepared the more detailed study that will follow in chapter 5 where the safety model is explained in more detail. This initial analysis aided in understanding the various trends within the database and the results yielded after implementing the highway statistics variables.