U.S. Department of Transportation
National Highway Traffic Safety Administration

TRAFFIC SAFETY FACTS
Research Note

NHTSA
www.nhtsa.gov

DOT HS 812 099 | Summary of Statistical Findings | January 2015

Validation of the National Estimates Produced From NASS GES

The National Automotive Sampling System (NASS) General Estimates System, operated by the National Highway Traffic Safety Administration, is responsible for producing national estimates for all types of motor vehicle traffic crashes that occur throughout the United States each year.

The objective of this research note is to validate the GES estimate on the total number of motor vehicle traffic crashes occurring annually. The method used was to compare an annual GES estimate with numbers of motor vehicle traffic crashes reported by the 50 States and the District of Columbia in the same year.

Background

NHTSA is responsible for reducing the number of fatal and injury crashes across the country. To fulfill its mission, NHTSA obtains data on motor vehicle traffic crashes from sources such as the Fatality Analysis Reporting System (FARS), NASS, and the State Data System (SDS).

The FARS is a census of all fatal crashes that occur in the 50 States and District of Columbia. The information comes mainly from police accident reports (PARs), State driver records, State vehicle registration files, death certificates, and State roadway files.

Since it is not possible to obtain all PARs for all crashes that occur in the United States, the NASS was designed to obtain a sample of crashes to produce national estimates on the number of motor vehicle traffic crashes in a given year. The NASS is comprised of two systems, the Crashworthiness Data System (CDS) and the General Estimates System (GES), and began operation in 1988. The CDS provides national estimates on detailed vehicle and occupant characteristics of passenger vehicle traffic crashes obtained mainly through vehicle and scene inspections, medical records, and interviews on about 4,000 crashes per year.

The GES is a multi-stage probability sample that provides national estimates on general characteristics of motor vehicle traffic crashes including the vehicles and people involved. The first stage of the sample is the selection of the sites that represent a central city, a large county, or a group of counties. The second stage of the sample is the selection of the police jurisdictions within each site. The third stage is the selection of a sample of PARs from each of the selected police jurisdictions, about 50,000 per year.

The GES obtains its information from PARs only. Examples of general characteristics include estimated number of injury crashes that occur across the county and by region of the county, estimated numbers of crashes involving pedestrians, large truck, or motorcycles, the estimated number of crashes involving rollovers, etc.

The original 1988 GES sample design projected an estimated 6.2 million police-reported motor vehicle traffic crashes. The motor vehicle crash counts used to provide the framework for the first stage sample design (selection of sites to collect the data) were obtained from 1983 and 1984 State crash files. Sixty sites in 26 States were selected across the United States.

Since the early 1980s, the NHTSA has been obtaining, from various States, computer data files coded from PARs; these are referred to as the SDS. The SDS contains crash data from 34 States that provide data files those States use to produce their annual crash numbers. Participation in the SDS program is voluntary. The data is received annually in various formats and converted to Statistical Analysis System (SAS, the name of the software system) data files. Although NHTSA does not change the coding used by each State, it does perform some basic standardization procedures used to help researchers, such as adopting a common file structure and names of variables, eliminating variables that contain personal identifiers, and making SAS character variables for all categorical numeric data. The data files received from the States are usually one to two years after the crash year.

Methodology

To verify the NASS GES overall national estimate on the number of motor vehicle crashes, the SDS data, supplemented by State Web sites, was used to obtain the number of crashes reported by all of the States and DC for 2009.

The data from two-thirds of the States was obtained from NHTSA's SDS. The motor vehicle crash data received by the SDS is the same as each State uses to produce its annual statistics. Of the 34 States in the SDS, 31 States had 2009 data available. For the 3 SDS States that did not have 2009 data, 2 used FARS fatal crash counts and factored in the history of the State crash counts, and the third used 2008 crash counts after comparing the crash counts across several years. For the 16 States plus DC that are not in the SDS, the data was obtained from the States' Web sites.

For a crash to qualify for the NASS GES, it must be reported on the State PAR and meet the criteria described in the National Safety Council's ANSI D16.1 Manual on Classification of Motor Vehicle Traffic Accidents, 7th Edition. These guidelines describe the differences in what constitutes a motor vehicle traffic crash and an incident. Note: Motor vehicle traffic crashes occur within the trafficway and incidents occur outside of the trafficway, where *trafficway* is defined as *any land way open to the public as a matter of right or custom for moving persons or property from one place to another.* Most, if not all, of the States follow these guidelines to determine if a case qualifies for its system as described in the State's PAR manual.

While the State PAR manuals describe the crashes that qualify for annual reporting, the local jurisdictions may not completely follow these guidelines. For example, a PAR manual might list the threshold for reporting crashes at $1,500 for vehicle/property damage, but the local jurisdictions may respond to crashes only when someone is injured.

Five States in the SDS include incident reports with their count of motor vehicle traffic crashes. These incidents can be identified by specific variables at the crash level. The crash counts shown in the table below exclude the incidents in these 5 SDS States. However, when a State publishes its crash counts, these incidents may be included in the total counts. The remaining States may include some of these incidents in their crash count totals, but the numbers they include are assumed to be so few that it would have minimal or no effect on the total number of crashes.

Findings

Table 1 shows the number of crashes in each State plus the District of Columbia. The column "Source" describes the source of the crash counts: SDS files or the States' individual Web sites. The comment column describes any issue that arose in obtaining the crash counts for each of the States. Summing the crash counts from all of the States plus DC, the total number of crashes that occurred in 2009 came to 6,085,916. Due to the statistical design of the GES, State level estimates cannot be produced, so direct comparison of GES versus States' data on a State-by-State basis is not possible.

In 2009, GES estimated 5,498,000 motor vehicle traffic crashes, about a 10-percent difference compared to the State total. The 95-percent confidence interval for this GES estimate is (4,985,583 – 6,009,429). In this same year, the number of crashes reported by the 50 States and DC was slightly more than 6 million motor vehicle traffic crashes. The States-plus-DC crash count is slightly higher than the upper bound of the confidence interval.

The differences between the GES estimates and the State totals may be attributable to a number of reasons.

1. The GES is a survey operating in a central city, large county, or group of counties, and these sites were not selected to represent the State in which it is located. Each State includes all locations within it and thus represents itself.

2. The GES uses specific guidelines to define what constitutes a motor vehicle traffic crash, i.e., crashes that occur within the trafficway. Most, if not all, of the States use the same guidelines as the GES but may also include additional crashes based on the State's needs. For example, some States include ALL fatal crashes regardless of where the crash occurred.

3. The GES includes only those crashes where the police jurisdiction from which the PAR is obtained indicates that it sends its PARs to the State for inclusion in the State's annual statistics. The States may include any crash that involves a fatality or of specific relevance to the State.

4. The changes in the threshold—dollar values used by the States to determine the minimum criteria for the State to report the crashes and include in their annual statistics—reduced the number of crashes by about 8 percent as shown by the comparisons from previous years' State data.

5. The GES may not have access to or obtained all of the qualifying PARs from the participating police jurisdictions.
6. The GES uses standardized coding procedures and definitions to code the data obtained from PARs from the 26 States to produce the GES analytical file.
7. Coding methodologies and definitions are unique to each State.

Conclusions

Comparing the national estimates from GES to the crash counts obtained from the States and DC show that the GES does a fairly decent job in producing these estimates given that the original data for GES was obtained in 1983 and 1984. The results from this study help to support the need for a new and updated record-based sample system that NHTSA is currently designing.

U.S. Department of Transportation
National Highway Traffic Safety Administration

This research note and other general information on highway traffic safety may be accessed by Internet users at: www-nrd.nhtsa.dot.gov/CATS/index.aspx

Table 1
Number of Motor Vehicle Traffic Crashes in the States and GES

State	# Crashes	Year	Source	Comments
Alabama	123,503	2009	SDS	
Alaska	12,890	2009	State Web Site	
Arizona	106,767	2009	State Web Site	
Arkansas	62,808	2009	SDS	
California	426,228	2009	SDS	
Colorado	105,000	2009	SDS	Estimated
Connecticut	103,719	2009	SDS	
Delaware	16,723	2009	SDS	
District of Columbia	16,841	2009	State Web Site	
Florida	235,803	2009	SDS	
Georgia	318,531	2009	SDS	
Hawaii	10,000	2009	State Web Site	Estimated
Idaho	22,992	2009	State Web Site	
Illinois	292,437	2009	SDS	
Indiana	189,983	2009	SDS	
Iowa	55,488	2009	SDS	
Kansas	61,119	2009	SDS	
Kentucky	126,237	2009	SDS	
Louisiana	155,857	2009	SDS	
Maine	33,118	2009	State Web Site	
Maryland	96,391	2009	SDS	
Massachusetts	136,384	2009	State Web Site	
Michigan	293,403	2009	SDS	
Minnesota	73,498	2009	SDS	
Mississippi	74,122	2009	State Web Site	
Missouri	153,015	2009	SDS	
Montana	21,971	2008	State Web Site	2009 data is not available to SDS
Nebraska	34,664	2009	SDS	
Nevada	53,151	2009	State Web Site	
New Hampshire	33,265	2009	State Web Site	
New Jersey	301,233	2009	SDS	
New Mexico	46,213	2009	SDS	
New York	314,974	2009	SDS	
North Carolina	209,695	2009	SDS	
North Dakota	17,686	2009	SDS	
Ohio	299,040	2009	SDS	
Oklahoma	71,218	2009	State Web Site	
Oregon	41,271	2009	State Web Site	
Pennsylvania	121,298	2009	SDS	
Rhode Island	41,788	2009	State Web Site	
South Carolina	106,864	2009	SDS	
South Dakota	16,994	2009	State Web Site	
Tennessee	155,099	2009	State Web Site	
Texas	428,667	2009	SDS	
Utah	51,367	2009	State Web Site	
Vermont	12,640	2009	State Web Site	
Virginia	116,742	2009	SDS	
Washington	110,070	2009	SDS	
West Virginia	39,906	2009	State Web Site	
Wisconsin	121,736	2009	SDS	
Wyoming	15,507	2009	SDS	
Total	**6,085,916**	**2009**	**States**	
Total	**5,498,000**	**2009**	**GES**	