Information Collection Request Supporting Statements: Part B
Crash Report Sampling System
OMB Control No. 2127-0714

Abstract:[1]

The National Highway Traffic Safety Administration (NHTSA) is seeking approval from OMB of this information collection request (ICR) for extension with modification of its currently approved information collection for Crash Report Sampling System (CRSS). NHTSA is authorized by 49 U.S.C. § 30182 and 23 U.S.C. § 403 to collect data on motor vehicle traffic crashes to aid in the identification of issues and the development, implementation, and evaluation of motor vehicle and highway safety countermeasures. The information collected serves to identify and develop safety countermeasures that will reduce the severity of injury and property damage caused by motor vehicle crashes.

The Crash Report Sampling System (CRSS) voluntary collection of data from police-reported crashes involving all types of motor vehicles, pedestrians, and cyclists; this includes property damage only crashes as well as those resulting in injuries and fatalities. CRSS obtains its data from a nationally representative probability sample selected from the estimated six million police-reported crashes that occur annually in the United States. By focusing attention on police-reported crashes, CRSS concentrates on the crashes of greatest concern to the highway safety community and the general public.

CRSS depends on the voluntary participation and cooperation of State and law enforcement agencies. This allows the National Highway Traffic Safety Administration (NHTSA) and its contractors to access the crash reports to review, list, and categorize the crashes. CRSS data is solely based on crash reports. The crash reports provide essential data: detailed information regarding the location of the crash, the vehicles, and the people involved. The crash reports are official local and State government forms that include the location of the crash and the pre-crash environment, explains the number and types of vehicles involved as well as describing the persons, injuries and other variables to express how the person was involved in the crash. CRSS respondents are local law enforcement agencies and State agencies that provide access to repository/website or database of motor vehicle crashes.

For the crash report data acquisition process, NHTSA's technicians regularly obtain PARs from the sampled police agencies and select a sample of the in-scope PARs. Once a PAR has been selected for data collection, NHTSA's data coders review and retrieve the general crash information from the sampled PAR. PAR is the sole source for the CRSS data collection.

These information collections support NHTSA's mission to save lives and prevent injuries due to traffic crashes. The data collected from the CRSS provide annual, nationally representative estimates of the number, types, and characteristics of police-reported motor vehicle crashes. These data are used by NHTSA to support its highway safety research, policy making, and regulation program development.

The previous request for this information collection (OMB No. 2127-0714) indicated 35,680 burden hours, this request increases the burden to 42,680. The request for the collection of information is adjusted due to a) reducing the burden hour estimates for CRSS information collection to be more accurate and reflect current efficiencies, b) adding the non-sampled PJ Crash Count Special Study and c) adding the PJ Frame Evaluation Special Study into this package. The CRSS special studies are the biggest change to the previous clearance. The special studies are critical to assessing the quality of the PJ frame of the CRSS PSUs to determine PJ weights and measure of size for the CRSS PJ sample selection. Without the special studies, NHTSA may fail to accurately assess the national crash picture by missing pertinent crash data. Thus, the importance of including the special studies and the estimated additional burden for the

[1] The Abstract must include the following information: (1) whether responding to the collection is mandatory, voluntary, or required to obtain or retain a benefit; (2) a description of the entities who must respond; (3) whether the collection is reporting (indicate if a survey), recordkeeping, and/or disclosure; (4) the frequency of the collection (e.g., bi-annual, annual, monthly, weekly, as needed); (5) a description of the information that would be reported, maintained in records, or disclosed; (6) a description of who would receive the information; (7) the purpose of the collection; and (8) if a revision, a description of the revision and the change in burden.

CRSS.

The combined impact is an increase of 7,000 burden hours to NHTSA's overall total.

B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

1. Describe the potential respondent universe and any sampling or other respondent selection methods to be used.

CRSS is a major record-based crash data collection system.[2] CRSS is a multi-stage complex survey of police crash reports. CRSS sample is comprised of PSU, PJ, and PAR samples.

In the following, we describe in detail the CRSS population and how the CRSS PSU, PJ, and PAR samples were selected.

The purpose of the CRSS is to provide annual, nationally representative estimates of police-reported motor vehicle crashes as well as characteristics of these motor vehicle crashes. PAR is the sole source of data for the CRSS. CRSS population is the set of police-reported motor vehicle crashes on a traffic-way (strata 2 – 10 of Table 2).

Table 2. CRSS PAR Strata, Target Sample Allocation, and Population Sizes

CRSS PAR Stratum	Description	Target Percent of Sample Allocation	Estimated Population (GES 2011) ****	Population Percent
1	An in-scope Not-in-Traffic Surveillance (NTS) crash (take all)*			
2	Crashes not in Stratum 1 in which: Involves a killed or injured (includes injury severity unknown) non-motorist	9%	119,579	2.2%
3	Crashes not in Stratum 1 or 2 in which: Involves a killed or injured (includes injury severity unknown) motorcycle or moped rider	6%	76,513	1.4%
4	Crashes not in Stratum 1-3 in which: At least one occupant of a late model year passenger vehicle** is killed or incapacitated	4%	22,272	0.42%
5	Crashes not in Stratum 1-4 in which: At least one occupant of an older passenger vehicle*** is killed or incapacitated	7%	84,659	1.6%
6	Crashes not in Stratum 1-5 in which: at least one occupant of a late model year passenger vehicle** is injured (including injury severity unknown)	14%	330,619	6.2%
7	Crashes not in Stratum 1-6 in which: involved at least one medium or heavy truck or bus (includes school bus, transit bus, and motor coach) with GVWR 10,000 lbs. or more	6%	302,781	5.7%

[2] Additional details about CRSS and how NHTSA collects this information are available in the supporting statements for the ICR with OMB Control No. 2127-0714.

CRSS PAR Stratum	Description	Target Percent of Sample Allocation	Estimated Population (GES 2011) ****	Population Percent
8	Crashes not in Stratum 1-7 in which: at least one occupant of an older passenger vehicle*** is injured (including injury severity unknown)	12%	800,390	15.0%
9.	Crashes not in Stratum 1-8 in which: involved at least one late model year passenger vehicle**, AND No person in the crash is killed or injured	22%	1,511,371	28.4%
10	Crashes not in Stratum 1-9: This includes mostly PDO crashes involving a non-motorist, motorcycle, moped, and passenger vehicles that are not late model year** and any crashes not classified in strata 1-9.	20%	2,078,263	39.0%

*: NTS cases are not in the scope of CRSS. They are set aside for NTS analysis.
**: Late model year passenger vehicle: passenger vehicle that are 4 years old or newer.
***: Older passenger vehicle: passenger vehicle that are 5 years old or older.
****: 2011 GES estimates were the most recent estimates at the time of the CRSS sample design.

The estimated CRSS population size (strata 2–10 of Table 2) was about 5.3 million in 2020. CRSS selects a sample from the population through a stratified multi-stage cluster scheme as follows:

First Stage (PSU Sampling)

The country is divided into geographic units called Primary Sampling Units (PSUs). A PSU is a county or group of counties and serves as a cluster. PSUs were formed as groups of adjacent counties subject to a minimum measure of size (MOS) condition to ensure enough cases will be sampled from each PSU and weights are approximately equal within each PAR stratum defined in Table 2. The CRSS PSU MOS was defined as:

$$MOS_i = \sum_{s=2}^{10} \frac{n_{++s}}{n} \frac{N_{i+s}}{N_{++s}}$$

where

- s = the PAR stratum defined in Table 2.
- n = the desired total sample size of crashes
- n_{++s} = the desired sample size of crashes in the PAR stratum s
- N_{++s} = the estimated population count of crashes in the PAR stratum s
- N_{i+s} = the estimated population count of crashes in the PAR stratum s and PSU i.

In the formula, n_{++s}/n is the desired PAR strata sample allocation (the "Target Percent of Sample Allocation" column in Table 2), and N_{i+s}/N_{++s} is the relative estimated population counts of PSU i for PAR stratum s. In this way, a PSU with a larger high interest (as defined by the oversampled PAR strata defined in Table 2) combination of estimated population counts of all PAR strata has a larger MOS.

PSU formation respects US Census region and urbanicity boundaries. While 23 remote outlying counties in Alaska and three counties of small islands in Hawaii were excluded, the rest of the country is included in the PSU frame. There are 707 CRSS PSUs in the PSU frame.

The PSU frame was then stratified into eight primary PSU strata by two variables – region (Northeast, West, South, and Midwest) and urbanicity (urban and rural). Within each primary stratum, PSUs were further stratified by secondary stratification variables such as vehicle miles traveled, crash rate, truck miles traveled, and crash rate by road type. PSUs with similar characteristics were grouped into secondary strata with approximately equal MOS sizes. Secondary strata groupings were also based on minimizing the between-PSU variance within a stratum. As the result, 50 PSU strata were formed as indicated in Table 3.

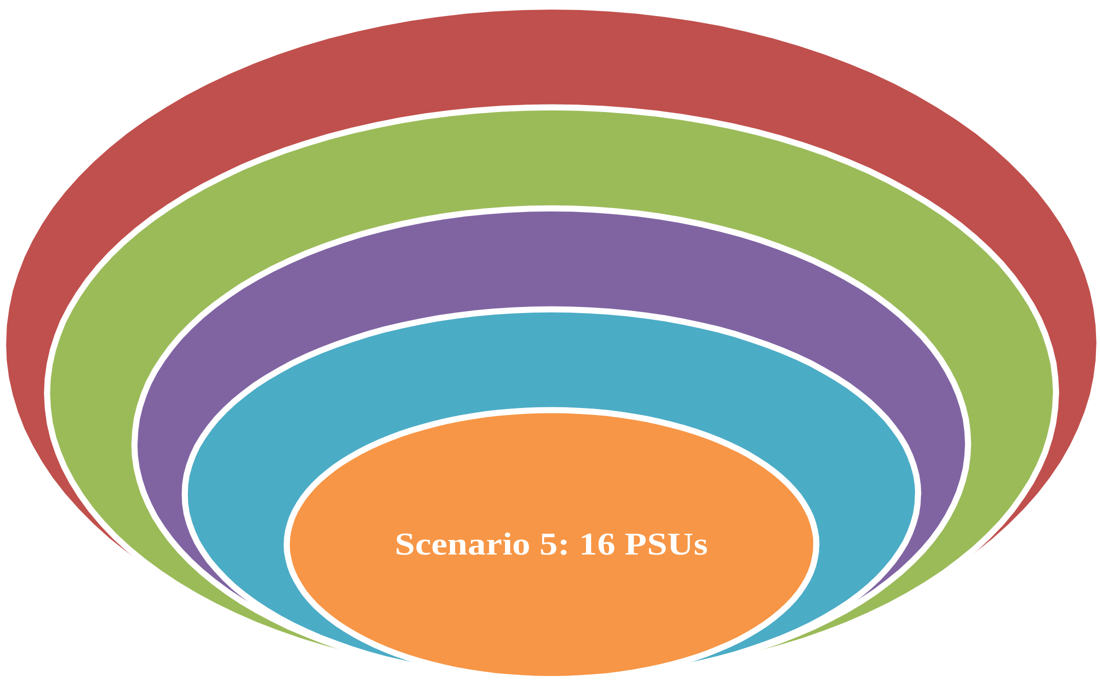
Table 3. CRSS PSU Strata, PSU Population Counts, and Sample Size

PRIMARY STRATA	STRAT ID*	VMT_RATE_IMP**		TOT_CRASH _RATE**		TRK_MI_RATE**		ROAD_TYPE _RATE**		Number of PSUs	PSU Sample Size
		Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower		
1	101	1801	0					359	0	5	2
1	102	4064	1801					359	0	5	2
1	103	7159	4064					359	0	8	2
1	104	5791	0	0.028	0	153756	0	2175	359	6	2
1	105	8040	5791	0.028	0	153756	0	2175	359	7	2
1	106			0.028	0	249918	153756	2175	359	7	2
1	107			0.028	0	591241	249918	2175	359	7	2
1	108			0.039	0.028			2175	359	11	2
2	201					236701	0			22	2
2	202					1027526	236701			22	2
3	301	4135	0			45709	0			3	2
3	302	7465	4135			45709	0			8	2
3	303	9898	7465			45709	0			10	2
3	304					102554	45709			11	2
3	305	4444	0			339758	102554			13	2
3	306	6003	4444			339758	102554			11	2
3	307	11618	6003			339758	102554			10	2
4	401					66171	0	4345	0	28	2
4	402	6045	0			565025	66171	4345	0	27	2
4	403	11623	6045			565025	66171	4345	0	25	2
4	404							17641	4345	30	2
5	501	3620	0	0.048	0	125590	0			5	2
5	502	4530	3620	0.048	0	125590	0			8	2
5	03	4951	4530	0.048	0	125590	0			6	2
5	504	5016	4951	0.048	0	125590	0			3	2
5	505	5277	5016	0.048	0	125590	0			5	2
5	506	5746	5277	0.048	0	125590	0			6	2
5	507	6399	5746	0.048	0	125590	0			5	2
5	508	12826	6399	0.048	0	125590	0			8	2
5	509	5641	0	0.048	0	210430	125590			6	2
5	510	8348	5641	0.048	0	210430	125590			7	2
5	511	13892	8348	0.048	0	210430	125590			10	2
5	512			0.048	0	358684	210430			8	2
5	513			0.048	0	877546	358684			13	2
5	514			0.085	0.048					17	2
6	601					49854	0			35	2
6	602	6353	0			162415	49854			34	2
6	603	14415	6353			162415	49854			35	2
6	604					250190	162415			33	2
6	605	5693	0			1156242	250190			35	2
6	606	16271	5693			1156242	250190			35	2
7	700									1	1
7	701	6477	0	0.027	0	104522	0			7	2
7	702	6921	6477	0.027	0	104522	0			4	2
7	703	7861	6921	0.027	0	104522	0			5	2
7	704	5137	0	0.027	0	249358	104522			3	2
7	705	8070	5137	0.027	0	249358	104522			10	2
7	706			0.048	0.027	92716	0			9	2
7	707			0.048	0.027	186409	92716			7	2
8	801							3938	0	30	2
8	802							18292	3938	41	2

*: STRATID: Secondary PSU ID.
**: VMT_RATE_IMP = imputed vehicle miles traveled / (PSU MOS×1,000,000).
TOT_CRASH_RATE = (imputed 2008 injury crashes + imputed 2008 PDO crashes + 2007-2011 average fatal crashes) / (PSU MOS×1,000,000).
TRK_MI_RATE = Total truck miles / (PSU MOS×1,000,000).
ROAD_TYPE_RATE = (primary road miles + secondary road miles) / (PSU MOS×1,000,000).

A major challenge of the CRSS sample design is the uncertainty of the future operational budget. Due to unknown future funding levels and the need for a stable PSU sample, NHTSA implemented a scalable PSU sample, which allows for the PSU sample size to be decreased or increased with minimum impact to the existing PSU sample and for the selection probabilities to be tracked. To this end, a multi-phase sampling method was used to select the CRSS PSU sample by selecting a sequence of nested PSU samples. In this method, a PSU sample larger than what is actually needed is selected during the first phase of the PSU sample. From the first phase of the PSU sample, a smaller subset of the PSU sample is selected as the second phase of the PSU sample. From the second phase of the PSU sample, another smaller third phase of the PSU sample is selected. This process is continued until the PSU sample size reaches unacceptable levels. In this way, a sequence of nested PSU samples is obtained. Each of these PSU samples is a probability sample and can be used for data collection (see Figure 2). According to the prevailing budget level, a sample with the appropriate sample size is picked from the nested sequence. This allows us to easily track the selection probabilities and minimizes changes to the existing PSU sample.

Figure 2. Nested PSU Samples for CRSS



For the CRSS, five PSU samples were selected under the five scenarios. Table 4 summarizes the number of PSU strata and sampled PSUs for the CRSS PSU sample scenarios.

Table 4. CRSS PSU Sample Scenarios: Number of Strata and Sample Size

Scenario	Number of PSU Strata	Number of Sampled Non-certainty PSUs	Number of Sampled Certainty PSUs	Total Number of Sampled PSUs
1	50	97	4	101
2	37	74	1	75
3	25	50	1	51
4	12	24	0	24

Scenario	Number of PSU Strata	Number of Sampled Non-certainty PSUs	Number of Sampled Certainty PSUs	Total Number of Sampled PSUs
5	8	16	0	16

For scenario 1, with a sample size of 100 and without stratification, one PSU was identified as a certainty PSU by the condition:

$$\pi_i = \frac{100 * MOS_i}{\sum_{i=1}^{N} MOS_i} \geq 1$$

Let N be the total number of PSUs in the PSU frame and i be an index for a PSU. The certainty PSU was selected with certainty[3] and set aside. Then two PSUs were selected using proportional to size (PPS) sampling from each of the 50 scenario-1 strata. With a sample size of two for each PSU stratum, three PSUs were identified as certainty PSUs from three of the 50 scenario-1 strata by the condition:

$$\pi_{hi} = \frac{2 * MOS_{hi}}{\sum_{i=1}^{N_h} MOS_{hi}} \geq 1$$

Let N_h be the total number of PSUs in stratum h. The certainty PSUs were selected with certainty and set aside. The corresponding stratum PSU sample size was reduced by one. Then a PPS sample of non-certainty PSUs was selected using the revised PSU stratum sample size.

Scenario-1 sample has 101 PSUs. For a non-certainty PSU, the selection probability is:

$$\pi_{hi} = \frac{n_{hi} MOS_{hi}}{\sum_{i=1}^{N_h} MOS_{hi}}$$

Let n_h be the non-certainty PSU sample size for PSU stratum h.

For scenario-2, with a sample size of 74 and without stratification, one PSU was identified as a certainty PSU and was set aside. Then 13 of the scenario-1 strata were collapsed with other strata to form the 37 scenario-2 PSU strata. The collapsing of strata follows the following rules:

- Only the secondary strata in the same primary stratum can be collapsed;
- Only the contiguous secondary strata can be collapsed;
- The resulting strata has a similar stratum total MOS within each primary stratum.

In each of the scenario-2 stratum, the sampled scenario-1 PSUs were treated as the sampling frame. Each PSU was assigned a new MOS equal to its scenario-1 stratum total MOS. Then two PSUs were selected from each scenario-2

---

[3] In the probability proportional to size (PPS) sampling, a certainty PSU is identified when the selection probability is equal to or greater than one. If a PSU is identified as certainty, it must be in the sample and its selection probability is set to one. A non-certainty PSU is selected with its selection probability that is greater than 0 and less than 1. If a PSU has a selection probability closer to one, it has more chance to be in the sample. On the other hand, if a PSU has a selection probability closer to zero, it has less chance to be in the sample. For more details, please see Pages 13-28 in the published Technical Report https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812706

stratum using PPS sampling based on the new MOS. In this way, the resulting selection probability of the scenario-2 PSU is still PPS selection probability. Other scenario samples were selected in a similar way.

The current CRSS PSU sample size is 61 (between scenarios 2 and 3) with 60 responding PSUs and one non-responding PSU.

Second Stage (PJ Sampling)

The secondary sampling units (SSU) of CRSS are police jurisdictions. Within each PSU, PARs are grouped by the police jurisdictions (PJ) where PARs are available and PJs become the second stage sampling units. A composite MOS is assigned to each PJ in the selected PSUs. Similar to the PSU MOS definition, it is sensible to assign larger selection probabilities to PJs with more high interest crashes as defined by the oversampled strata in Table 2. For each PJ in the selected PSUs, crash counts from the 9 PAR strata in Table 2 (Stratum 2-10) were estimated from the information collected from the PJs in the selected PSUs. For PJ j in the PJ frame within the sampled PSU i, the composite SSU MOS is defined as the following:

$$MOS_{j \vee i} = \sum_{k=2}^{10} \frac{n_{++s}}{n} \frac{N_{ijs}}{N_{++s}}$$

where

s = the PAR stratum defined in Table 2.

n = the desired total sample size of crashes

n_{++s} = the desired sample size of crashes in the PAR stratum s

N_{++s} = the estimated population count of crashes in PAR stratum s

N_{ijk} = the estimated population count of crashes in PAR stratum s, PJ j and PSU i

PJs are then stratified into two PJ strata by their MOS (large MOS stratum [largest 50%] and small MOS stratum [the rest]) in addition to certainty PJs. A PJ sample is then selected from each PJ stratum using Pareto sampling. The Pareto sampling method produces an approximate PPS sample, handles the frame changes and minimizes the changes to the existing sample at the same time. Pareto sampling method was applied to the PJ sample selection for each of the non-certainty PJ strata (large MOS and small MOS stratum) within the sampled PSUi, as the following:

Step 1: Generate a permanent uniform random number r_{ihj} $U(0,1)$ for each PJ j in the PJ stratum h of PSU i.

Step 2: Identify certainty PJs by the condition:

$$\frac{m_{ih} * MOS_{ihj}}{\sum_{j=1}^{M_{ih}} MOS_{ihj}} \geq 1$$

Let m_{ih} be the PJ sample size and M_{ih} be the PJ frame size for PJ stratum h within PSU i. MOS_{ihj} is the PJ MOS.

Step 3: The identified certainty PJs are set aside. This process is repeated for the remaining PJs based on the reduced PJ sample size until there are no more certainty PJs. Let the total number of certainty PJs be m_{ch}. For the remaining $M_{ih} - m_{ch}$ non-certainty PJs in the frame, calculate the PPS inclusion probability with the non-certainty PJ sample size ($m_{ih} - m_{ch}$¿:

$$p_{ihj} = \frac{(m_{ih} - m_{ch}) MOS_{ihj}}{\sum_{j=1}^{M_{ih} - m_{ch}} MOS_{ihj}}$$

Step 4: Calculate the transformed random numbers and sort the transformed random numbers from the smallest to the largest as following:
¿

Step 5: The m_{ch} certainty PJs plus the first $m_{ih} - m_{ch}$ non-certainty PJs from the above list are the PJ sample for PJ stratum h within PSU i.

Pareto sampling is approximately PPS, and the PJ selection probability is:

$$\pi_{j \vee ih} \approx p_{ihj}$$

The 2022 CRSS PJ sample size is 351.

Third Stage (PAR Sampling)

The tertiary sampling units (TSU) of CRSS are PARs. The CRSS PAR sample is selected by stratified systematic sampling. For each selected SSU (PJ), PARs are periodically obtained by either a technician's visit to the PJ or electronic transmission. All the PARs are listed in the order they become available and are stratified by the PAR strata identified in Table 2. Through this listing process, the PAR sampling frame in each selected PJ is prepared for PAR sample selection.

For a large PJ with too many PARs to be listed, PARs are sub-listed by systematic sampling. For example, only PARs with a PAR number ending in 0 through 4 may be listed if the sub-listing factor is 2 (i.e., 5 PARs among 10 PARs are listed). Or only PARs with a PAR number ending in 0 or 1 are listed if the sub-listing factor is 5 (i.e., 2 PARs among 10 PARs are listed). If L PARs among 10 PARs are sub-listed in PJ j in PSU i, the sub-listing probability for all sub-listed PARs are:

$$\pi_{l \vee ij} = \frac{L}{10}$$

After PARs are listed, a PAR sample is selected by systematic sampling from the listed (or sub-listed) PARs by PAR stratum within each selected PJ. PAR k selection probability within a PAR stratum is:

$$\pi_{k \vee ijl} = \frac{n_{ijl}}{N_{ijl}}$$

Let n_{ijl} be the number of selected PARs and N_{ijl} be the number of listed PARs from each PAR stratum in PJ j of PSU i.

The overall selection probability is:

$$\pi_{ijlk} = \pi_i * \pi_{j \vee i} * \pi_{l \vee ij} * \pi_{k \vee ijl}$$

The design weight is the inverse of π_{ijlk}.

Sample Allocation

For a three-stage sample design as this program, the PSU, SSU and TSU sample sizes can be estimated using optimization by minimizing the variance subject to cost assuming a three-stage simple random sampling without replacement.

The optimization model consists of the objective function, cost constraint, and variance constrains as the following:

$$Minimize:\sum_{g=1}^{G} V\left(\overline{\overline{\overline{y}}}_g\right)=\sum_{g=1}^{G}\left\{\frac{S_{1,g}^2}{n}\left(1-\frac{n}{N}\right)+\frac{S_{2,g}^2}{nm}\left(1-\frac{m}{M}\right)+\frac{S_{3,g}^2}{nmk}\left(1-\frac{k}{K}\right)\right\}$$

$$Subject ¿ : C=C_0+nC_1+nmC_2+nmkC_3,$$

$$V_{CRSS}\left(\overline{\overline{\overline{y}}}_g\right)=\frac{S_{1,g}^2}{n}\left(1-\frac{n}{N}\right)+\frac{S_{2,g}^2}{nm}\left(1-\frac{m}{M}\right)+\frac{S_{3,g}^2}{nmk}\left(1-\frac{k}{K}\right)\le V_{GES}\left(\overline{\overline{\overline{y}}}_g\right),$$

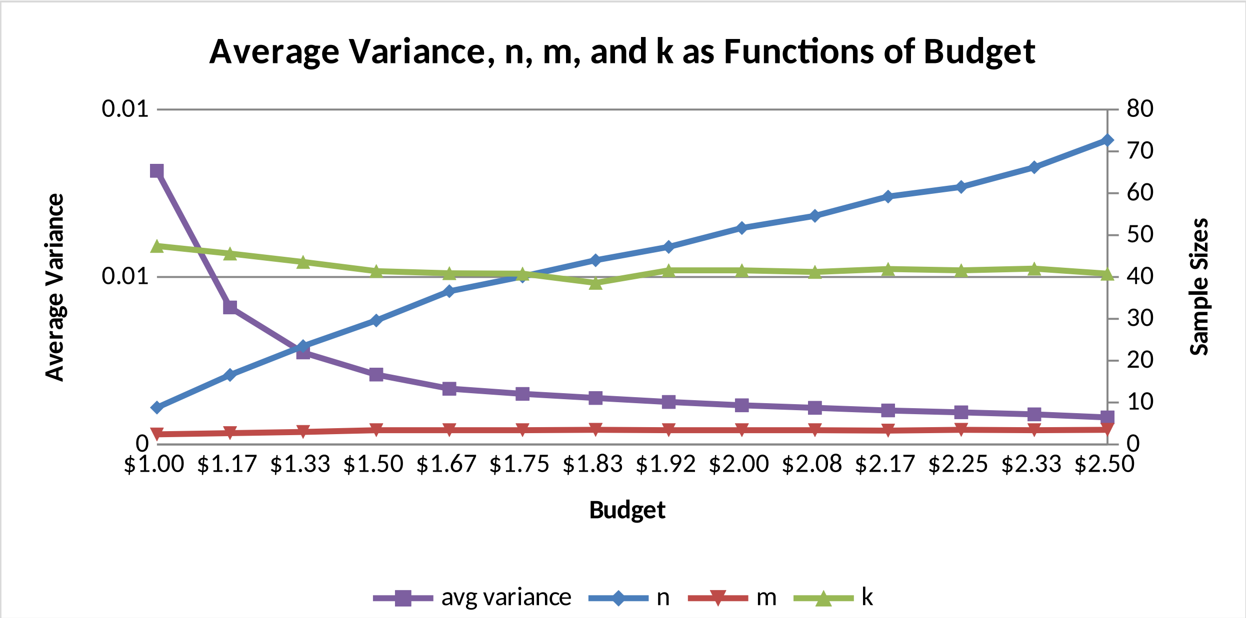
$$for\ g=1,\cdots,G.$$

where

- g: Subscript of the identified key estimate, $g=1,\cdots G$.
- $\overline{\overline{\overline{y}}}_g$: Identified key proportion estimate.
- n,m,k: Optimal sample sizes of PSUs, SSUs per PSU, and TSUs per SSU to be determined.
- N: Population size of PSUs
- M: Average population size of SSUs.
- K: Average population size of TSUs.
- V_{CRSS}: Variance of the identified key estimate $\overline{\overline{\overline{y}}}_g$.
- $S_{1,g}^2, S_{2,g}^2, S_{3,g}^2$: Variance component at PSU-, SSU-, and TSU-level.
- C, C_0, C_1, C_2, C_3: Total, fixed, PSU-, SSU-, and TSU-level cost coefficients.
- $V_{GES}\left(\overline{\overline{\overline{y}}}_g\right)$: Variance of the identified key estimate $\overline{\overline{\overline{y}}}_g$ in General Estimates System (GES)[4].

Figure 3 displays the optimization results. As the rescaled budget increases, the PJ sample size m and the PAR sample size k tend to be stable while the PSU sample size n increases and the average variance decreases. Since there are 9 PAR domains to be estimated (strata 2-10 in Table 1), the final PAR sample size is 9*k.

Figure 3: Average Variance, PSU, PJ and PAR Sample Size as Functions of Budget



Note: All costs are rescaled, so the lowest cost starts from $1.

[4] https://www.nhtsa.gov/national-automotive-sampling-system/nass-general-estimates-system

Each year, CRSS's target sample is 50,000 crashes. Below is a table summarizing the CRSS sample. The 2016 CRSS data collection year had only 53 PSUs instead of the full 60 PSUs due to constraints with receiving police crash reports, thus the sample was less than the anticipated 50,000 crashes.

Table 5: Unweighted Summary Statistics

Year	Crashes	Vehicles (in Transport)	People	Drivers	Occupants	Pedestrians	Pedalcyclists
2016	46,511	82,149	117,759	82,000	113,405	2,257	1,576
2017	54,969	97,625	138,913	97,388	133,408	2,881	1,946
2018	48,443	86,105	120,230	85,916	115,774	2,444	1,436
2019	54,409	96,717	135,410	96,488	129,980	2,949	1,802
2020	54,745	94,718	131,962	94,500	126,460	2,882	1,923

1. Describe collection of information procedures.

CRSS data collection efforts are dependent on the method in which the crash reports are accessed. The crash reports are accessed through NHTSA's Electronic Data Transfer (EDT) program, data feeds, secure email, State websites, or manually by contract staff that physically visit the police jurisdiction.

The EDT program consists of a routine automated transfer of crash data from the State crash database to NHTSA. EDT reduces the level of effort required to share crash data because data is automatically shared nightly from the State to NHTSA. States may also provide crash reports to NHTSA through secure web service portals on a routine basis. The State designates the frequency with which they share data with NHTSA under this protocol. Crash report accessed via EDT and secure web portal are uploaded into the Police Accident Report Sampling Engine (PARSE). The PARSE application is a centralized, web-based repository in which CRSS applicable crash reports are listed, categorized, and selected for further coding. In 2022, one third (20 out of 60) CRSS PSUs are EDT PSUs.

Alternatively, States may provide access to their crash data collection websites. States provide log-in credentials to view crash reports for the sample PJs. The sampler would then list, categorize, and sample the crash reports for sample agencies within the PARSE application.

When States are not able to provide electronic access to crash reports, NHTSA seeks manual access to crash reports from individual police jurisdictions identified in the CRSS sample. Generally, this includes visiting the office to access paper or electronic files, uploading crash reports on an encrypted thumb drive, linking crash reports to a secure email or copying crash reports and sending the crash reports via mail courier service. These more manual processes are completed on a schedule established by the police agency. Once the schedule is agreed upon, then the CRSS sampler can view, list, categorize, and sample the crash reports within the PARSE application. In 2020 CRSS, for example, NHTSA obtained cooperation from 308 operational PJs in 40 non-EDT PSUs.

2. Describe methods to maximize response rates and to deal with issues of non-response.

CRSS has a three-stage sample design. The first stage sampling units are counties or groups of counties. A PSU becomes a non-responding PSU only if all selected police jurisdictions (PJs) within the PSU are non-responding PJs. In CRSS, the PJ sample is selected using the Pareto sampling method. The whole PJ frame can be used as replacement sample. Therefore, a PSU becomes a non-responding PSU only if all PJs in the frame are non-responding PJs. In 2017 CRSS, one PSU was non-responding. Since the CRSS PSU sample is scalable, we increased the sample size from 60 to 61 and selected a replacement PSU without changing the original PSU sample. The weight of the non-responding PSU was adjusted.

The second stage sampling units of CRSS are PJs. A sampled PJ becomes non-responding PJ if it refuses to cooperate. To improve PJ cooperation rate, NHTSA visits each selected PJ and meets with local law enforcement officers to gain cooperation. In 2022 CRSS, 23 PJs among the 331 sampled PJs in non-EDT PSUs were non-responding. The weights of non-responding PJs were adjusted.

The third stage sampling units of CRSS are PCRs. First all police crash reports (PCRs) in the selected PJs are listed. Then a systematic sample of PCRs is selected and coded. A PCR is identified as non-responding if it has un-readable pages or missing pages. In 2017 CRSS for example, only 17 PCRs among the 55,274 sampled PCRs were non-responding. The weights of the non-responding PCRs were adjusted.

The CRSS quality control system is designed to produce the most accurate, reliable, and complete database possible within the limits of available resources. Each selected case is reviewed by quality control personnel for accuracy before proceeding to coding. Additionally, the Police Accident Report Sampling Engine (PARSE) automatically selects five percent from each sample of non-selected cases to review. The findings from the listed cases review helps identify any quality control issues and additional training needs for the CRSS Sampler.

4. Describe any tests of procedures or methods to be undertaken.

There is no tests procedure to be undertaken at this time.

5. Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.

Ms. Chou-Lin Chen, National Center for Statistics and Analysis, NHTSA, 202-366-1048 is responsible for CISS survey design and special studies.

NHTSA contracted with Westat (contract DTNH22-12-F-00389) on the CRSS survey design effort.

NHTSA has contracted with KLD Associates Inc. (contract DTNH2214D00366L/0002) for the data collection, coding and quality control for the CRSS data collection effort.