

**Department of Transportation
National Highway Traffic Safety Administration
Supporting Statements: Part B
National Survey of the Use of Booster Seats
OMB Control Number: 2127-0644**

Abstract:¹

The National Survey of the Use of Booster Seats (NSUBS) is a voluntary, biennial survey that collects data on restraint use information for children under 13. Data collectors observe vehicles at locations such as fast-food restaurants, gas stations, daycare centers, and recreation centers, where child occupants are more likely to be present. When a vehicle with child passengers enters a selected location, the data collectors note the age, sex, and restraint use of all occupants. If the vehicle's driver voluntarily participates, they provide additional information, including the race/ethnicity of all occupants and the height and weight of the child passengers. No personal identifying information, such as names, addresses, or phone numbers, is collected during this process.

The contractor processes the collected data and supplies it electronically to the National Highway Traffic Safety Administration (NHTSA). The contractor also generates replicate weights, derives analytic variables, and integrates auxiliary information such as State laws. NHTSA uses the data to estimate restraint use among children under 13 and produces detailed breakouts by age, race/ethnicity, height, and weight. The data also helps assess how often children are prematurely transitioned to restraint types that are inappropriate for their age, height, or weight. In addition, NHTSA collects the age of adult drivers to analyze the relationship between driver age, driver seat belt use, and child restraint use. NHTSA utilizes this information to design programs that target caregivers of children who are either unrestrained or improperly restrained, with the goal of improving child safety in motor vehicles. The findings may also inform State legislatures considering updates to their child restraint laws.

The current collection, approved under the Office of Management and Budget (OMB) Control No. 2127-0644 (expiration date: 09/30/2025), reflects revisions including a transition from paper forms to electronic devices.

The average number of respondents has decreased from 5,300 to 4,600, reducing the estimated burden hours from 376 hours to 326 hours. The contractor obtains Institutional Review Board approval before any data collection begins. NHTSA estimates that the annual burden to respondents is approximately \$8,329.30.

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) if the information collection involves approval by an institutional review board, include a statement to that effect; (8) the purpose of the collection; and (9) if a revision, a description of the revision and the change in burden.

B. JUSTIFICATION

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

The purpose of t NSUBS is to gather information on appropriate restraint use for all child occupants, in particular, the use of booster seats among children ages 4-7. The survey will visit sampled gas stations, recreation centers, day care centers, and seven specific fast food restaurant chains (McDonald's, Taco Bell, Burger King, Wendy's, Kentucky Fried Chicken, Dairy Queen, and Sonic). Data collectors will approach as many vehicles as possible that appear to have at least one child occupant under the age of 13 to allow for data collector observation of restraint use for all occupants, and subsequent interviews to determine age, height, weight, race, and ethnicity for child occupants.

The potential respondent universe consists of all motorists with child passengers (age 12 and under). However, the accessible population available for data collection are those who frequent gas stations, recreation centers, day care centers, and seven specific fast-food restaurants (not located in shopping centers).

These site types (gas stations, recreation centers, day care centers, and fast-food restaurants (those not located in shopping centers)) were chosen because they are frequented by motorists transporting children, and businesses' parking lots are usually small enough that data collectors can likely approach vehicles as they are parking, before child restraints have been unfastened and occupants have entered nearby buildings.

Data collectors will approach as many motorists as possible who appear to have at least one child under the age of 13 in their vehicle for possible participation in the survey.

B.2 Describe the procedures for the collection of information.

Sampling Frame

The sample design is a three-stage design, where geographic areas are sampled as primary sampling units (PSUs), and data collection sites are selected from sampled PSUs at the second stage, and lastly, vehicles with child passengers are observed from the respondent sites.

PSU Formation and Measure of Size (MOS)

The county is the building block of NSUBS as it is one of the most recognizable geographic units with rich auxiliary information including the Census population data of various age groups. PSUs were formed with two basic principles: (1) PSUs should be geographically contiguous and compact so that the travel cost of the data collectors is minimized; (2) yet they should be large enough in terms of the MOS so that data collection efforts are not wasted due to small sample yield.

Considering that NSUBS covers children of ages 0-12, the size of this age group of children could be used as the MOS for selection of NSUBS PSUs by the probability-proportional-to-size method. However, it is hard to get a direct count of these children, so we chose to use the number of children of ages 0-9 from the 2016-2020 American Community Survey as the MOS because this age group is more inclusive of children of booster seat age.

The PSUs were formed in the continental U.S. respecting State boundaries to stratify PSUs by census region and urbanicity. This PSU formation process resulted in 1,601 PSUs that cover the continental U.S. The PSUs consist of 1.94 counties on average and have an average MOS of 24,579 children of ages 0-9.

PSU Stratification

The PSUs are stratified by census region and urban or rural status. Counties are identified as urban or rural based on the National Center for Health Statistics (NCHS) classification, which has six different levels. The six levels were then converted down to two levels at the PSU level, where if at least one county within the PSU has an NCHS level of 1, 2, or 3 the PSU was classified as urban, otherwise it was classified as rural. Combined with the census region, there were eight major strata.

Table 1. Stratification of PSUs Census Region and Urban Status

Census Region	Urban Status	Number of PSUs	Total Child Population
Northeast	Urban	103	5,514,457
	Rural	74	716,890
Midwest	Urban	180	5,767,974
	Rural	289	2,568,798
South	Urban	370	12,167,763
	Rural	372	3,254,407
West	Urban	81	7,950,785
	Rural	132	1,410,502

PSU Sample Size and Allocation

The NSUBS includes 30 PSUs using the approximately optimal Neyman allocation to reduce overall variance, called for selecting 4, 6, 15, and 6 PSUs from the Northeast, Midwest, South, and West regions, respectively.

Secondary Sampling Units and Sample Selection

Data are collected at four data collection site types (daycare centers, recreation centers, fast food restaurants, and gas stations) on the basis that using these four site types to conduct the survey does not cause a significant bias. Seven fast food restaurant chains (McDonald's, Burger King, Wendy's, Taco Bell, Kentucky Fried Chicken, Sonic and Dairy Queen) were selected because they are commonly found nationwide. Site selection considerations are that the businesses typically have a good number of customers with children and dedicated parking lots for data collection.

The goal is to obtain 15 respondent sites per PSU. Sites will be selected prior to data collection and will look similar to the 2015 sample selection. Using SafeGraph² data, data for the sites of interest are filtered out into their own dataset. Sites are then processed to determine if they are eligible for including in the sampling frame; are they located within one of the 30 PSUs, located outside of military base or office building, if a daycare, rec center, or fast food restaurant it is not located in a shopping center, if a rec center it is not a park, climbing wall, or senior center, if a daycare it is not home-based.

Sampling strata are created by site type within PSU and sites are independently sampled by stratum within PSU. Sample sizes within PSU are determined, allocating proportionally to each site type according to the site frame size for a total of 450 eligible, responding sites across all PSUs to acquire approximately 15 respondent sites per PSU.

Data Collection Schedule

The NSUBS data collection protocol calls for 2-hour observation periods of restraint use for all passenger vehicles with child occupants who enter selected data collection locations. Using data from the 2025 NSUBS, we will examine whether to reduce the 2-hour data collection time at daycare centers to 1 hour or when all children have arrived. Once the vehicles stop, the drivers are approached for short interviews. Because of the requirement for observations, data are collected only during daylight hours, typically defined as between 7 a.m. and 6 p.m.

Trained data collectors approach passenger vehicles appearing to have child occupants under the age of 13; observe the restraint use of up to nine occupants in the first three rows of seats; and conduct interviews to obtain the race and ethnicity of all occupants; the age of the driver; and the heights, weights, and ages of child occupants appearing to be under age 13. The approximate ages of other occupants (expressed as an age range, such as 16-24 years), and the sex of all occupants are subjectively assessed by the data collectors.

Per OMB standards, the data on race/ethnicity in NSUBS are collected via self-reporting. To capture restraint use before children unfasten the restraints, data collectors observe restraint use prior to or just as the vehicles come to a stop except at fast food drive-through lanes. In that case, restraint use is observed prior to the vehicles reaching the drive-through window.

To reach as wide an audience as possible, NSUBS uses some Spanish-speaking data collectors in PSUs with known populations of Spanish-speaking residents.

Data are recorded in tablets and data sent back to Westat, the Contractor, electronically.

Information Collected

If an adult in the approached vehicle agrees to take part in the survey, data collectors will ask the adult for the following information:

² SafeGraph is a data company that provide spatial data including points of interest and business listings. Its datasets include attributes such as location details and business categories.

- Children's ages
- Children's heights
- Children's weights
- Children's races/ethnicities
- Adults' races/ethnicities
- Driver's age

In addition, data collectors will collect the following information by observation only, and not by interview:

- Survey site information
- Vehicle type
- Seating position of each occupant
- Restraint use for each occupant, specifying the types of child restraints used
- Sex of each occupant
- Estimated age range for adult occupants other than driver

Data Collection Form

Screenshots from the data collection application for the interview portion used in the survey are shown in the attachments.

Data collectors will pick their data collection site, verify their name and their partner's name, and select the weather condition.

Data collectors will then select the vehicle type when an eligible vehicle is observed, even if the occupants refuse to agree to participate in the survey. This is because the observation data are retained and used, even in the absence of related interview data. Data collectors will input observed data for each eligible car. Information on the age category, sex, and restraint use of all occupants will be filled out by the data collector based on observation.

Data collectors will introduce themselves and ask drivers if they will participate in the survey. Data collectors tell all potential respondents that their participation is voluntary, and that no personal identifying information will be collected.

For motorists who agree to participate, data collectors will interview an adult motorist for answers to fill out the application's information on restraint use.

Data collectors will keep track of the number of vehicles that they missed and the number of vehicles whose occupants declined to participate in the survey, and who refused due to language barriers.

Data collectors receive extensive training in protocols for interviewing motorists and observing restraint use in a manner that is professional and as unobtrusive as possible.

Statistical Editing, Imputation, Estimation, and Variance Estimation

Simple range edits will be performed on the data to improve data quality. For instance, the data will be edited to ensure that children's ages fall between 0 and 12 years. Data that fall out of range will be treated as missing.

Restraint use will be estimated by the following:

Base Weighting

To calculate the PSU-level base weight, a simulation method is used since the sample was selected using sequential Poisson sampling. The simulation method for calculating the inclusion probabilities is computed as follows:

1. For each PSU i in the NSUBS frame, generate X_{ij} independently from a Uniform(0,1) distribution for $j=1, \dots, B$, where B is the number of simulations to be done.
2. For each $j=1, \dots, B$, draw the NSUBS 2015 sample following the same procedure used to select the PSU in this design (e.g., stratified systematic sample with the PSUs ordered by 2015 MOS. Define $Z_{2015,ij}=1$ if PSU i is selected to be in the 2015 sample in simulation j , and 0 otherwise.
3. Generate the retrospective permanent random number (RPRN) for the NSUBS 2015 frame accounting for the sample selected in (1).
4. Using the RPRN and the 2025 MOS, select the 2025 PSU sample using the new strata. The selection is based on the systematic. Define $Z_{2025,ij}=1$ if PSU i is selected in the 2025 sample in simulation j , and 0 otherwise.
5. Determine the empirical inclusion probability for PSU i , for the 2015 design as $\pi_{2015,i} = \sum_{j=1}^B Z_{2015,ij} / B$. The PSU-level weight for PSU i is then $1/\pi_{2015,i}$.
6. Determine the empirical probability that PSU i is selected in the 2025 samples as $\pi_{2025,i} = \sum_{j=1}^B Z_{2025,ij} / B$. The PSU-level weight for PSU i is then $1/\pi_{2025,i}$.

Then, the site probability of selection within the PSU is used to compute the conditional site base weight. The base weights are adjusted for site nonresponse, followed by weighting adjustments for the observation time difference, missed observations, and missed interviews.

Site Level Nonresponse Adjustment

An eligible site is considered “responding” if data collection was conducted at the site for more than 60 minutes.

We define site-level nonresponse adjustment cells by site type within PSUs. Within each cell, the nonresponse adjustment factor is the sum of the weights for all eligible sites divided by the sum of the weights for responding sites only. This essentially scales up responding sites in the cell so that they represent all eligible sites in the cell. Theoretically, there are 120 (= 4 site types times 30 PSUs) possible cells, which collapsed as needed due to small sample sizes and/or large adjustment factors. The final number of cells is denoted by L_s and varies by survey year. We also denote the nonresponse adjustment cells for site-level adjustment as

$$C_a^{(S)} \text{ for } 1 \leq a \leq L_s,$$

where the superscript (S) symbolizes the “site” sample. Each cell can consist of up to three kinds of sites: (1) eligible and participated with an observation duration that lasted more than 60 minutes (responding); (2) eligible but did not participate or had an observation duration that did not last 60 minutes (nonresponding); (3) ineligible; and (4) unknown eligibility³.

These categories are denoted, respectively, by $C_{1a}^{(S)}$, $C_{2a}^{(S)}$, and $C_{3a}^{(S)}$. Some cells may have no nonrespondents and/or no ineligible sites. For simplicity, the stratum index is dropped in much of the following notation.

³ The “unknown eligibility” adjustment category is a theoretical category that is not used in NSUBS. In NSUBS, the eligibility of all sample sites is usually known, so $C_{4a}^{(S)}$ are usually empty for all $1 \leq a \leq L_s$. Since we are able to determine the eligibility of every site, we have removed any reference to unknown eligibility adjustment in previous reports.

Using this notation, the site level nonresponse adjustment factor for cell $C_a^{(S)}$ is defined as:

$$A_a^{(S)} = \frac{\sum_{ij \in C_{1a}^{(S)}} w_{ij} + \sum_{ij \in C_{2a}^{(S)}} w_{ij}}{\sum_{ij \in C_{1a}^{(S)}} w_{ij}}, \text{ for } 1 \leq a \leq L_s.$$

If $ij \in C_{1a}^{(S)}$ (i.e., site ij is a respondent in the nonresponse adjustment cell $C_{1a}^{(S)}$), then the site level nonresponse adjusted weight for site ij is given by

$$A_a^{(S)} w_{ij}.$$

For a handful of sites, data were collected for a period of time that is longer or shorter than the scheduled 2 hours. In order to account for this variation, a duration adjustment is applied that effectively standardizes the duration to exactly 2 hours for all sites. The duration adjusted weight for respondent

site ij in the nonresponse adjustment cell $C_{1a}^{(S)}$ is defined by

$$w_{ij}^{(S)} = \frac{120}{D_{ij}} A_a^{(S)} w_{ij},$$

where D_{ij} is the time in minutes for which the data collectors collected data from site ij . Sites with $D_{ij} < 60$ are treated as nonrespondents and included in $C_{2a}^{(S)}$ starting in 2015.

Vehicle Level Nonresponse Adjustment

At the site level, a vehicle is considered to be responding if it is observed, regardless of interview status. The vehicle-level nonresponse adjustment cells are created within each responding site with at least one observed

vehicle, denoted as $C_a^{(V)}$ where a is the index for the participating sites. Each

cell is divided into observed (responding) vehicles (denoted as $C_{1a}^{(V)}$) and missed (nonresponding) vehicles (denoted as $C_{2a}^{(V)}$). Each observed vehicle k in $C_a^{(V)}$ is given the vehicle level nonresponse adjusted weight defined as

$$w_{ijk}^{(V)} = A_a^{(V)} w_{ij}^{(S)},$$

where

$$A_a^{(V)} = \frac{\sum_{ij \in C_{1a}^{(V)}} w_{ij}^{(S)} + \sum_{\tilde{ij} \in C_{2a}^{(V)}} w_{\tilde{ij}}^{(S)}}{\sum_{ij \in C_{1a}^{(V)}} w_{ij}^{(S)}}.$$

This same nonresponse adjusted weight is given to all occupants in the vehicle, so it can also be called the observation weight.

These weights are adjusted for estimates based on interviewed occupants only to account for interview-level nonresponse. Because restraint use estimates for interviewed child occupants are produced separately from estimates for other interviewed occupants, two sets of interview weights are developed: weights for interviewed child occupants ages 0-12 and weights for interviewed adult occupants ages 13 and older.

Child Interview Nonresponse Adjustment

Nonresponse adjustment cells for child interview nonresponse are defined by site type within PSU, resulting in 120 possible cells. After collapsing over cells with small sample sizes and/or large adjustment factors, there are L_C final adjustment cells. Denote each cell by $C_a^{(I)}$ for $1 \leq a \leq L_C$. All interviewed children in vehicle ijk belonging to $C_a^{(I)}$ are given the nonresponse adjusted weight defined by:

$$w_{ijk}^{(I)} = A_a^{(I)} w_{ijk}^{(V)} \quad \text{for } 1 \leq a \leq L_c,$$

where

$$A_a^{(I)} = \frac{\sum_{ijk \in C_{1a}^{(I)}} w_{ijk}^{(V)} + \sum_{ijk \in C_{2a}^{(I)}} w_{ijk}^{(V)}}{\sum_{ijk \in C_{1a}^{(I)}} w_{ijk}^{(V)}},$$

and $C_{1a}^{(I)}$ is the set of interviewed children and $C_{2a}^{(I)}$ is the set of observed but not interviewed children.

Adult Occupant Interview Nonresponse Adjustment

Nonresponse adjustment for adult occupant interview nonresponse is

performed identically to that for children. Let there be L_A nonresponse

adjustment cells formed, which are denoted as $C_a^{(A)}$ for $1 \leq a \leq L_A$. All

interviewed adult occupants in vehicle ijk in cell $C_a^{(A)}$ are given the nonresponse adjusted weight defined by:

$$w_{ijk}^{(A)} = A_a^{(A)} w_{ijk}^{(V)} \quad \text{for } 1 \leq a \leq L_A,$$

where

$$A_a^{(A)} = \frac{\sum_{ijk \in C_{1a}^{(A)}} w_{ijk}^{(V)} + \sum_{ijk \in C_{2a}^{(A)}} w_{ijk}^{(V)}}{\sum_{ijk \in C_{1a}^{(A)}} w_{ijk}^{(V)}},$$

and $C_{1a}^{(A)}$ is the set of interviewed adult occupants and $C_{2a}^{(A)}$ is the set of observed but not interviewed adult occupants.

Variance Estimation

Variance estimates will be computed using Westat's WesVar software, reflecting the jackknife variance estimation method.

B.3 Describe methods to maximize response rates.

The refusal rate by year is described in the following table:

Table 2 Interview Refusal Rate by Drivers and Survey Years

Year	Total Number of Observed Vehicles	Total Number of Interviews	Total Number of Refusals	Refusal Rate (%)
2006	3,489	2,920	548	15.7
2007	4,828	4,199	181	3.7
2008	6,204	4,899	224	3.6
2009	6,033	4,601	286	4.7
2011	6,350	5,191	300	4.7
2013	7,229	6,070	288	4.0
2015	6,060	5,352	243	4.0
2017	7,490	6,396	460	6.1
2019	5,605	5,005	383	6.8
2021	5,249	4,470	400	7.6

We do not expect many missing values in the observed portion of the data, because the data collectors will be well trained, and they should have adequate time to record site information and restraint use.

The interview refusal rate for drivers is on average 5.7 percent for the last 5 years of the survey.

Imputation

For occupant variables including a vehicle variable, Westat's imputation software, AutoImpute, is used. AutoImpute is a convenient tool since it can perform imputation for categorical and continuous variables at the same time. It uses regression modeling to create imputation classes and carries out hot-deck imputation using the created imputation classes through regression modeling. One of its strengths is its ability to use all available variables as predictors in the regression modeling so that it can maintain, in the imputed data, the correlation structure of the data set among variables used in imputation. Since it utilizes hot-deck imputation, AutoImpute performs donor imputation, where for a missing value, a donor is randomly picked from the non-missing cases in the imputation class to which the missing case belongs, and the donor value is used to impute the missing value. This imputation procedure is applied to all variables with missing values, except for a few variables that are not imputed (mostly site characteristics plus observer name, number of refusals, number of missed vehicles, seating position) and a few variables imputed as special cases (site identifiers and administrative data).

Data Quality

Data collectors for the NSUBS undergo extensive training to minimize errors that could arise from their categorizing or recording data incorrectly.

The NHTSA does not believe that there is reliable information with which to adjust the survey results to account for inaccurate responses given by motorists, motorists who choose not to participate in the survey, motorists who do not frequent the site types, or motorists who frequent the site types outside of the observation period. The agency's published report will clearly state that the results are based on motorists who visit the site types and voluntarily chose to participate in the survey.

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

The NSUBS was designed in 2006. A pilot test was conducted at that time, and the results of that test were incorporated into the revised survey procedures. Subsequently, NSUBS has been conducted eleven times (through 2023). In 2015, the survey was redesigned and expanded from 16 PSUs to 30 PSUs (as is described above) to reduce the variance and enhance the statistical precision of the estimates. No changes were made to the data collection protocol because of the 2015 sample redesign. The survey went through another redesign in 2025 to update the same and reduce variance. Using data from the 2025 NSUBS, we will examine whether to reduce the two-hour data collection time at daycare centers to one hour or when all children have arrived.

B.5 Provide the name and telephone number of individuals consulted on statistical aspects of the design.

This survey was designed and will be conducted under Federal contract with Westat, Inc. The Assistant Contracting Officer's Representative is Lacey B. Werth who can be reached at (202) 366-7468. The program manager at Westat, Inc. is Dr. Doreen De Leonardis who can be reached at (301) 315-5963.