2021 Vehicle Inventory and Use Survey Methodology

INTRODUCTION

The Vehicle Inventory and Use Survey (VIUS) is a joint effort by the U.S. Department of Transportation's (DOT) Bureau of Transportation Statistics (BTS) and the U.S. Department of Commerce's Census Bureau, in partnership with the DOT Federal Highway Administration and the U.S. Department of Energy.

The VIUS provides data on the physical and operational characteristics of the truck population in the United States. Its primary goal is to produce national- and state-level estimates of the total number of vehicles, vehicle miles, and average miles per vehicle by body type and other physical and operational characteristics. Physical characteristics include date of purchase, weight, number of axles, overall length, type of transmission, and body type. Operational characteristics include type of use, lease characteristics, operator classification, base of operation, gas mileage, annual and lifetime miles driven, commodities hauled by type, and hazardous materials carried.

The VIUS includes vehicles that are classified by manufacturers as trucks, minivans, vans, or sport utility vehicles (SUVs), and excludes governments-owned vehicles, ambulances, buses, motor homes, farm tractors, unpowered trailer units, and trucks reported to have been disposed of prior to January 1 of the survey year 2021.

The VIUS was the principal data source on the physical and operational characteristics and configurations of the United States truck population from 1963 through 2002 (from 1963 through 1992 as the "Truck Inventory and Use Survey"). The VIUS was restored after nearly 20 years for the 2021 survey year. Since the VIUS was last conducted in 2002, caution should be exercised when comparing the 2021 VIUS estimates to 2002 VIUS estimates. For more information, please refer to the "Survey Design Changes That Affect the Comparability of the 2002 and 2021 VIUS Estimates" document.

Although the VIUS estimates the U.S. truck population, New Hampshire did not allow the U.S. Census Bureau access to its motor vehicle registration records—therefore, New Hampshire was not included in the survey universe for the VIUS data collection. As a result, U.S. estimates exclude the state of New Hampshire. Tables of national estimates including modeled estimates for New Hampshire are provided in an Excel file as a one-time release to supplement the main 2021 VIUS publication tables.

The VIUS data are used by the BTS, other government agencies (federal, state, and local), and private sector companies. The VIUS data are used for freight movement analysis, highway cost allocation, truck size and weight evaluation, investment and performance analysis, commercial motor vehicle safety analysis, cost allocation studies, energy consumption and demand studies, and system performance analysis.

SCOPE AND FRAME

The target population of the 2021 VIUS includes most privately-owned and commercial vehicles registered with motor vehicle departments in the 50 states and the District of Columbia that are classified by vehicle manufacturers as trucks, minivans, vans, or SUVs in operation during the 2021 calendar year.[1]

Registered vehicle data and analysis for the VIUS sampling frame is provided by R. L. Polk & Co. The sampling frame is constructed by R. L. Polk & Co. from files of vehicle registrations identified as being active as of July 1, 2021. The VIUS frame excludes known vehicles owned by federal, state, and local governments; ambulances; buses; motor homes; farm tractors; unpowered trailer units; and trucks reported to have been disposed of prior to January 1 of survey year 2021. If a vehicle is found to be one of these excluded types during data collection, then the vehicle is also excluded from analysis. Vehicle information is unduplicated by R. L. Polk & Co. to reflect the most recent information available.

All contact information is provided by R. L. Polk & Co. with one exception. California independently delivered contact information for personal and commercial vehicles, which includes address information, but does not include owner information.

SAMPLE DESIGN

The 2021 VIUS has a stratified simple random sample design. The sampling frame is stratified by geography and vehicle characteristics. Geography refers to the state in which the address is provided on the vehicle registration and not the state with whom the vehicle is registered. These may differ, although it is rare. Due to the exclusion of New Hampshire, the 49 remaining states and the District of Columbia make up the 50 geographic strata.

[1] Types of vehicles defined by R. L. Polk & Co. are available at <www.spglobal.com/mobility/en/products/polk-automotive-solutions.html>.

Body type, gross vehicle weight rating (GVWR), and vehicle usage determine the following ten strata:

- Commercial pickups.
- Commercial vans, minivans, SUVs.
- Commercial straight trucks with GVWR equal to or less than 26,000 lbs.
- Commercial straight trucks with GVWR more than 26,000 lbs.
- Commercial truck tractors.
- Personal pickups.
- Personal vans, minivans, SUVs.
- Personal straight trucks with GVWR equal to or less than 26,000 lbs.
- Personal straight trucks with GVWR more than 26,000 lbs.
- Personal truck tractors.

Therefore, the sampling frame is partitioned into 50 x 10 = 500 geographic-by-vehicle strata. Within each stratum, a simple random sample of vehicles is selected without replacement. Each vehicle was selected using a unique vehicle identification number (VIN). This produces a total sample of approximately 150,000 vehicles.

For many programs, there is a need to distinguish between the sampling unit, the reporting unit, and the tabulation unit. A sampling unit is an entity selected from the underlying statistical population of similarly constructed units. A reporting unit is an entity from which data are collected. A tabulation unit houses the data used from estimation or tabulation and is the most detailed level of unit used for estimation purposes. For the purposes of the VIUS, the VIN is used as the sampling unit, reporting unit, and tabulation unit.

TARGET COEFFICIENTS OF VARIATION

A sample cannot be designed that estimates each characteristic with the same reliability since it depends on the proportion of vehicles in the population that possess the characteristic (or distribution of miles among vehicles for the mileage question). For a stratified sample, the variance of an estimate of a proportion of a population having a given characteristic is minimized for a given overall sample size using the following formula:

$$n_h = n\frac{N_h\sqrt{P_hQ_h}}{\sum N_h\sqrt{P_hQ_h}}$$

Here, n represents the full sample size, n_h represents the sample size for a specific stratum h, N_h represents the population size for the stratum, P_h represents the proportion of the population in the stratum having the characteristic, Q_h represents the proportion of the population in the stratum that does not have the characteristic ($Q_h = 1 - P_h$), and the sum in the denominator is across all strata (Cochran, 1977a).

For sample allocation in VIUS, the ten vehicle strata within each state are merged into three groups based on vehicle size and given targets for the coefficient of variation (CV). This design optimization was done to ensure that similarly sized vehicles received the same CV target. Then the allocation is calculated to meet the CV targets assuming a hypothetical characteristic shared by 10 percent of the population of vehicles (with similar proportions in all strata). The merged strata and CV targets are as follows:

- State-level joint CV target for strata 1, 2, 6, and 7 is 12 percent.
- State-level joint CV target for strata 3 and 8 is 10 percent.
- State-level joint CV target for strata 4, 5, 9, and 10 is 8 percent.

The CV targets are smaller for larger vehicles, which was a focus of the questionnaire design.

The contents and sampling frame is under the sole authority of R. L. Polk & Co. As a result, sample sizes are calculated and given to R. L. Polk & Co. for selection. Using sampling frame counts provided by R. L. Polk & Co., sample sizes are calculated for the merged strata, and the sample is allocated proportionally to the number of vehicles in each stratum. Using the assumptions above, the allocation for each merged stratum, k, and target CV (as a ratio), c, is:

$$n_k = \frac{9}{c^2}\left(1 + \frac{9}{N_k c^2}\right)^{-1}$$

This process results in sample sizes ranging from 2,575 to 3,033 vehicles for each state, except for the District of Columbia, which has a sample size of 1,513. After receiving sample sizes for each stratum, R. L. Polk & Co. selects the sample via simple random sampling without replacement.

DATA COLLECTION

The 2021 VIUS used an electronic reporting system as the main source of data collection. All respondents received an initial letter with instructions to create an account in the respondent portal and log into the electronic instrument to complete the survey for the vehicle(s) listed on the letter. A paper version of the questionnaire was sent as part of the second follow-up operation; however, respondents had the option to request a paper questionnaire at any time.

Table 1.
Data Collection Schedule

Operation	Date
Advance letter mailout	2/9/2022
Initial letter mailout	2/23/2022
Open electronic instrument to respondents	2/23/2022
Due date reminder mailout	3/23/2022
Due date reminder email	3/28/2022
First mail follow-up	4/20/2022
First email follow-up	5/9/2022
Second mail follow-up	6/6/2022
Second email follow-up	6/27/2022
Begin telephone follow-up	7/5/2022
End telephone follow-up	8/30/2022
Third email follow-up	9/6/2022
Respondent closeout	10/31/2022

Source: U.S. Census Bureau.

There were two versions of the questionnaire that were provided to respondents depending on the vehicle's GVWR class and vehicle type.[2] Form TC-9501 was the light vehicle questionnaire for light trucks including pickups, SUVs, minivans, and light vans. Form TC-9502 was the heavy vehicle questionnaire for heavy trucks including straight trucks over 10,000 pounds and truck tractors.

Businesses with large truck fleets may have multiple vehicles selected for the survey. Respondents with three or more vehicles in the sample had the option to download an Excel spreadsheet of the questionnaire, fill in the responses for all of their sampled vehicles, and upload the completed spreadsheet. This option reduced the burden on fleet managers who would otherwise have to complete the electronic instrument multiple times for their sampled vehicles.

Because the VIUS data were last collected in 2002, collection operations included an advance letter to introduce respondents to the survey. This was followed by an initial letter and emails with instructions for completing the survey. Follow-up operations included letters, emails, and robocalls. A targeted telephone follow-up operation was conducted to encourage nonresponding units to log in and complete the surveys. For a schedule of 2021 VIUS data collection activities, refer to Table 1.

[2] For more information regarding questionnaires, refer to <www.census.gov/programs-surveys/vius/information.html>.

EDITS

The 2021 VIUS data were edited through a combination of automated corrections and direct analyst review to ensure the data were reasonable and consistent with the available administrative data. Administrative data were received from R. L. Polk & Co. who used decoded VIN information along with proprietary knowledge to create vehicle specific values.

Automated corrections were used to revise responses with inconsistencies among data items. For example, if a respondent failed to indicate if they still own their vehicle but later reported a date of disposal, automated corrections were used to fill in the missing information. Automated correction edits included (but were not limited to) the following: filling in select blank responses, correcting reporting errors, blanking out unreasonable responses, and ensuring that respondent "write-in" text was recoded to the proper category. This process was done before direct analyst review.

In direct analyst review, an analyst could manually review the entirety of specific cases. Direct data review edits included (but were not limited to) comparing respondent answers against administrative data from R. L. Polk & Co. and against information encoded in the VIN and reviewing responses for reasonableness and validity.

Table 2 shows the 11 categories of edits and provides category descriptions, examples, and how they were resolved. This table provides examples of how item response data can fail the edit process; this is not an exhaustive list.

IMPUTATION

Surveys generally do not yield complete responses from every sampled unit. In certain situations, nonresponse can bias survey estimates if appropriate adjustments are not made. There are two types of nonresponse applicable to the VIUS: unit nonresponse and item nonresponse.

To be considered a "response," the following items must be reported: (1) whether the respondent still owns the vehicle, and (2) if the previous answer is "No," had the respondent disposed of the vehicle prior to January 1, 2021. Unit nonresponse occurs when these response criteria are not met and is resolved through a nonresponse weighting adjustment (for more information, refer to Weighting and Estimation).

Item nonresponse occurs either when a question is unanswered or when the response is unusable. Nonresponse to the following items is corrected by imputation: (1) the number of months a vehicle was operated in the 2021 survey year, (2) the number of miles driven during the

Table 2.
Edit Check Summary

Category	Definition	Examples	Resolution
Blank paper form receipt	Check if the paper form returned was entirely blank.	A paper form was returned without response to any items.	Flag as blank form; exclude from further edit checks.
Post mail out-of-scope	If a response indicates the vehicle was never owned by the respondent or disposed of prior to January 1, 2021.	A respondent reported having never owned the vehicle or reported having disposed of the vehicle prior to January 1, 2021.	Flag response as out-of-scope; exclude from further edit checks.
Recode write-in response	Map write-in responses to valid responses in lookup lists.	A respondent writes that their vehicle was fueled at a gas station with the brand name.	Map write-in response to lookup list entry; otherwise, flag for analyst review.
Sum to 100 percent	Check if an item response appropriately sums to 100 percent.	A respondent reports that over 105 percent or less than 95 percent of their range of operations occurred between zero and 500 miles.	Automatically adjust response to sum to 100 percent, unless it less than 95 percent or greater than 105 percent, then flag it for analyst review.
Electronic response	Validate responses from paper forms or electronic responses that ignored error messages.	A respondent reports that towing capacity is zero, or reports feet for cab height is less than 5 or more than 15 feet.	Flag for analyst review.
Key variables	Ensure the respondent provided answers to the "must have" responses to count as a response.	Did a respondent provide responses to certain items like type of trailer pulled or whether the vehicle was still owned on January 1, 2021?	Flag for analyst review.
Administrative data consistency	Identify item responses that can be checked against data provided by R. L. Polk & Co. or decoded vehicle identification number (VIN).	Did a respondent provide responses consistent with information obtained from R. L. Polk & Co. or decoded VINs such as brakes and body class?	Flag for analyst review.
Other consistency	Respondents are reporting correctly for cases where a response could be used to inform another response.	A respondent reports that the cost of extensive repairs was more than $500 but did not answer if the vehicle had extensive repairs in 2021.	Apply consistent response systematically.
Skip-pattern	Check that the skip pattern is followed correctly.	A respondent answers an item that should not have been answered.	Item response that should have been skipped is blanked.
Reasonableness check	Does item response seem reasonable for the type of vehicle and characteristic provided by respondent?	A respondent reports having driven more miles in 2021 than the lifetime miles for the vehicle.	Flag for analyst review.

Source: U.S. Census Bureau.

2021 survey year, (3) the number of miles driven since the vehicle was manufactured, (4) the total length of a vehicle configuration, and (5) the total average weight of a vehicle configuration. Note that "1" and "4" are categorical responses, whereas "2," "3," and "5" are numeric. Also note that the number of miles driven since the vehicle was manufactured is not provided on the 2021 VIUS data tables—this item will only appear on the Public Use File, which was released in December 2023.

For the 2021 VIUS, the imputation methodology consists of a combination of value substitution, model-based methods (mean and mode), and hot deck donor imputation. These methods were selected to be consistent with the imputation methods used in the 2002 VIUS. Only units classified as responses are eligible for item imputation.

When the number of months a vehicle was operated in the 2021 survey year is not reported, information regarding the vehicle's acquisition and disposal are used to determine length of ownership. This derived length of ownership is used as a proxy for the number of months the vehicle was operated.

The number of miles driven during the 2021 survey year leverages the mean imputation method, where response units are divided into a finite number of mutually exclusive cells based on vehicle model year and other related vehicle characteristics including vehicle type and the number of trailers pulled. For each cell, estimates of the weighted average mileage are computed based on those vehicles in the cell for which number of miles driven has been reported. Missing values are then replaced with the appropriate average values. The post-stratified sampling weight is used to construct weighted averages (for more information on post-stratified sampling weights, refer to Weighting and Estimation). The number of miles driven since the vehicle was manufactured also leverages this same mean imputation method. This method was preferred because it avoided imputing values of extreme usage of a vehicle.

The total length and the average weight of a vehicle configuration are imputed using hot deck donor imputation. For these data items, if a response unit is missing either total length or average weight, both data items are replaced with data from a vehicle with similar characteristics (defined by registration state, truck type or body type, number of trailers pulled, and GVWR) for which total length and average weight have been reported. This method was preferred since it leveraged the ability to mirror other vehicles with similar physical characteristics.

For all other data items, no imputation is performed. Instead, separate estimates are published in a "Not reported" category. Users of the estimates should exercise caution when allocating the estimate for the "Not reported" category to the estimates for the reported categories in the proportions observed for the reported categories. This is because the characteristics of the vehicles for which we obtain data may differ significantly from those vehicles for which we obtain no data.

WEIGHTING AND ESTIMATION

Three major estimates are published for the 2021 VIUS: number of vehicles, miles traveled in 2021 (vehicle miles), and average miles per vehicle. Vehicle miles are annualized to represent usage for the vehicle in all of 2021, regardless of owner. Estimates of the number of vehicles and vehicle miles are computed as the sum of weighted (reported or imputed) data. Estimates of average miles per vehicle are derived by dividing the estimate of the total vehicle miles by the estimated number of vehicles.

Each vehicle is assigned a single tabulation weight that is used in computing all estimates to which the vehicle contributed. The tabulation weight for a given vehicle is the product of two factors: a post-stratified sampling weight and a nonresponse adjustment factor. Post-stratified sampling weights (inverse probability of selection within the sample universe) are calculated for individual (or groups) of GVWRs within each of the 500 sampling strata; these adjusted sampling weights help to ensure the distribution of GVWRs within the survey universe is preserved in estimates derived from the selected sample.

For example, assume for a given state that the survey universe is comprised of 5,000 GVWR class 7 vehicles and 5,000 GVWR class 8 vehicles in sampling stratum 4. Then assume that 200 vehicles from class 7 and 300 vehicles from class 8 are selected into the sample. If all the vehicles receive the same sampling weight of 10,000 / 500 = 20, then (barring the identification of any out-of-scope vehicles among the response units) one would estimate that there are 4,000 GVWR class 7 vehicles and 6,000 GVWR class 8 vehicles in the stratum. While this is a reasonable estimate from the sample, it ignores the known information about the universe. Instead, separate weights are calculated for each GVWR class (or group of classes) to preserve the GVWR class distribution. In this example, the stratum 4, GVWR class 7 weight is 5,000 / 200 = 25 and the stratum 4, GVWR class 8 weight is 5,000 / 300 = 16.666667. For information regarding stratum definition, refer to Sample Design.

Table 3 shows how the sampling strata are substratified.

Table 3.
Substratification Levels

Sampling Strata	Substrata (gross vehicle weight rating)
Strata 1, 2, 6, 7	1. Class 1 2. Class 2 and above
Strata 3, 8	1. Classes 1 and 2 2. Class 3 3. Class 4 4. Class 5 5. Class 6
Strata 4, 9	1. Class 7 2. Class 8
Strata 5, 10	No substratification

Source: U.S. Census Bureau.

Substrata are (rarely) collapsed further if there are not enough response units in the defined substrata.

The nonresponse weighting addresses unit nonresponse by refining the original weights to account for units that did not respond to the survey. After applying this weighting, the response units correctly sum to the full sample total. Nonresponse weighting is performed at the substratum level. Consider a substratum with 20 vehicles in the sample, representing a universe of 4,000. This gives a first-stage weight of 4,000 / 20 = 200. In this hypothetical substratum, 12 of the sampled vehicles are classified

as responses, but two of them provide information to indicate that they are out-of-scope to the VIUS. The nonresponse weight is calculated by taking the substratum sample size and dividing by the total number of response units (including the out-of-scope response units). That leads to a nonresponse weight of 20 / 12 = 1.666667 and a tabulation weight of 200 x 1.666667 = 333.333333.

Estimates on VIUS publication Table 5A (Mileage by Registration State and Body Type for the U.S. [excluding New Hampshire] and States: 2021) and Table 5B (Mileage for Business-use Trucks by Registration State and Body Type for the U.S. [excluding New Hampshire] and States: 2021) represent the proportion of vehicle miles that meet some extra condition(s).[3] These mileage estimates are calculated by first multiplying the total vehicle miles by the proportion of miles that meet the given condition, then applying the tabulation weight.

Most of the estimates published in the VIUS come directly from responses to the questionnaire. There are a few estimates that leverage administrative data from VIN decoding. These estimates include model year, GVWR class, cylinders, and cubic inch displacement.

SAMPLING ERROR

The sampling error of an estimate based on a sample survey is the difference between the estimate and the result that would be obtained from a complete census conducted under the same survey conditions. This error occurs because characteristics differ among sampling units in the population and only a subset of the population is measured in a sample survey. The sample used in this survey is one of many samples of the same size that could have been selected using the same sample design. Because each unit in the sampling frame has a known probability of being selected into the sample, it is possible to estimate the sampling variability of the survey estimates.

Common measures of the variability among these estimates are the sampling variance, the standard error, and the CV (which is also referred to as the relative standard error). The sampling variance is defined as the squared difference, averaged over all possible samples of the same size and design, between the estimator and its average value. The standard error is the square root of the sampling variance. The CV expresses the standard error as a percentage of the estimate to which it refers.

For example, an estimate of 200 units that has an estimated standard error of 10 units has an estimated CV of 5 percent. The sampling variance, standard error, and CV of an estimate can be estimated from the selected sample because the sample is selected using probability sampling. Note that measures of sampling variability, such as the standard error and CV, are estimated from the sample and are also subject to sampling variability. It is also important to note that the standard error and CV only measure sampling variability. They do not measure any systematic biases in the estimates.

The Census Bureau recommends that individuals using these estimates incorporate sampling error information into their analyses, as this could affect the conclusions drawn from the estimates.

Variance estimates for the VIUS are computed using sample design-based formulas. The formulas evaluate the dispersion of item values within each substratum. For a given geography, *s*, and substratum, *h*, the variance for the estimated number of vehicles or vehicle miles in 2021 can be expressed as:

$$v(a) = \frac{N_{s,h}(N_{s,h} - r_{s,h})}{r_{s,h}(r_{s,h} - 1)} \left[\sum_k a_k^2 - \frac{(\sum_k a_k)^2}{r_{s,h}} \right]$$

In this formula, $N_{s,h}$ is the substratum population size for geography *s*; $r_{s,h}$ is the number of response units in the substratum (including response units determined to be out-of-scope) for geography *s; k* is a specific vehicle; a_k is the item value in question for the specific vehicle; and the sums are across all response units in the substratum (Cochran, 1977b). For estimates of number of vehicles, a_k equals one if the qualifying conditions were met; otherwise, a_k equals zero. For estimates of vehicle miles, a_k equals the vehicle's miles driven (total or distribution) if the qualifying conditions are met; otherwise, it equals zero. Substratum variance estimates are summed to obtain state-level variance estimates. Similarly, state-level variance estimates are summed to obtain U.S.-level variance estimates.

Variance estimates for the ratio of average miles per vehicle are derived by finding variance estimates for each item and an estimated covariance between the two items; then, a ratio variance estimation formula is applied. The covariance estimate between vehicle miles and number of vehicles in a substratum is:

$$cov(a,b) = \frac{N_{s,h}(N_{s,h} - r_{s,h})}{r_{s,h}(r_{s,h} - 1)} \left[\sum_k a_k b_k - \frac{(\sum_k a_k)(\sum_k b_k)}{r_{s,h}} \right]$$

In the previous formula, a_k and b_k are an indicator for the vehicle meeting the conditions and the vehicle miles (or

[3] For VIUS publication estimate tables, refer to <https://data.census.gov/table?q=vius21>.

zero if the conditions are not met), respectively. The covariance estimates can be summed across substrata and state, as well. Finally, the variance estimate for the ratio is:

$$v\left(\frac{b}{a}\right) = \left(\frac{b}{a}\right)^2 \left[\frac{v(a)}{a^2} + \frac{v(b)}{b^2} - \frac{2cov(a,b)}{ab}\right]$$

In the previous formula, *a* and *v(a)* represent estimates of number of vehicles and associated variances for a given level, and *b* and *v(b)* represent estimates of vehicle miles and associated variances at the same level (Lee, 2006).

In this publication, estimates that have high CVs (50 percent or greater) are suppressed and denoted by an "S." Some of these suppressed estimates can be derived directly from the tables by subtracting published estimates from their respective totals. However, the suppressed estimates obtained by such subtraction would be subject to poor response, high sampling variability, or other factors that may make them potentially misleading. Estimates derived in this manner should not be attributed to the Census Bureau.

The sample estimate and an estimate of its standard error allow us to construct interval estimates with prescribed confidence that the interval includes the average result of all possible samples with the same size and design. To illustrate, if all possible samples are surveyed under same or nearly identical conditions and an estimate and its standard error are calculated from each sample, then:

- For approximately 90 percent of the possible samples, the interval from 1.645 standard errors below the estimate to 1.645 standard errors above the estimate would include the population parameter as obtained from a complete enumeration of the sampling frame conducted under the same survey conditions.
- For approximately 95 percent of the possible samples, the interval from 1.96 standard errors below the estimate to 1.96 standard errors above the estimate would include the population parameter as obtained from a complete enumeration of the sampling frame conducted under the same survey conditions.

To illustrate the computation of a confidence interval for an estimate of the number of vehicles, assume that an estimate is 3,377,800 (published as 3,377.8 thousand) vehicles and the coefficient of variation for this estimate is 2.9 percent, or 0.029. First, obtain the standard error by multiplying the estimated number of vehicles by its coefficient of variation. For this example, multiply 3,377,800 by 0.029 yields a standard error of 97,956.2. The upper and lower bounds of the 90 percent confidence interval are computed as 3,377,800 plus or minus 1.645 times 97,956.2. Consequently, the 90 percent confidence interval is 3,216,700 to 3,538,900 (after rounding). If corresponding confidence intervals are constructed for all possible samples of the same size and design, approximately 9 out of 10 (90 percent) of these intervals would contain the result obtained from a complete enumeration of all vehicles on the sampling frame.

NONSAMPLING ERROR

Nonsampling error encompasses all factors other than sampling error that contribute to the total error associated with an estimate. This error may also be present in censuses and other nonsurvey programs. Nonsampling error arises from many sources: inability to obtain information on all units in the sample; response errors; differences in the interpretation of the questions; mismatches between sampling units and reporting units, requested data and data available or accessible in response units' records, or with regard to reference periods; mistakes in coding or keying the data obtained; and other errors of collection, response, coverage, and processing.

Although no direct measurement of nonsampling error was obtained, precautionary steps were taken in all phases of the collection, processing, and tabulation of the data to minimize its influence. Precise estimation of the magnitude of nonsampling errors would require special experiments or access to independent data and, consequently, the magnitudes are often unavailable.

The Census Bureau recommends that individuals using these estimates factor in this information when assessing their analyses of these data, as nonsampling error could affect the conclusions drawn from the estimates.

RESPONSE RATES

Economic programs at the Census Bureau are required to compute two different types of response rates: a unit response rate and a weighted item response rate. For information regarding item nonresponse, refer to Edits and Imputation. For information regarding unit nonresponse, refer to Weighting and Estimation.

The unit response rate (URR) is defined as the unweighted proportion of eligible, or unknown eligible, reporting units that responded to the survey. URRs are indicators of the performance of data collection for obtaining usable responses. National-level URRs for all vehicle strata were similar, ranging from 46.9 to 52.5 percent. Refer to Imputation for the definition of a response.

The total quantity response rate (TQRR) is defined as the percentage of the weighted estimated total of a given item reported by the active tabulation units in the statistical period or from sources determined to be equivalent-to-reported data. The TQRR is an item-level indicator of the "quality" of each estimate. In contrast to the URR,

these weighted response rates are computed for individual data items.

The URRs and the TQRRs for number of vehicles and vehicle miles at the registration geography level are listed in Table 4.

NATIONAL ESTIMATES INCLUDING NEW HAMPSHIRE (MODELED)

New Hampshire did not allow the Census Bureau access to its motor vehicle registration records and, therefore, New Hampshire was not included in the VIUS data collection. The main VIUS sample-based survey estimates do not include estimates for New Hampshire and, similarly, the main national-level estimates exclude all vehicles with registration addresses in New Hampshire.

An experimental data product provides the national-level estimates, which include modeled estimates of the characteristics of vehicles with registration addresses in New Hampshire. Estimates of vehicles with registration addresses in New Hampshire were modeled using data collected from three nearby states (Maine, Massachusetts, and Vermont). These modeled estimates were then combined with estimates of the survey data collected from the remaining 49 U.S. states and the District of Columbia.

Tables of national estimates including modeled estimates for New Hampshire are provided as an Excel file as a one-time release to supplement the main 2021 VIUS publication tables.[4]

DISCLOSURE AVOIDANCE

Disclosure is the release of data that reveals information or permits deduction of information about a particular survey unit through the release of either tables or microdata. Disclosure avoidance is the process used to protect each survey unit's identity and data from disclosure. Using disclosure avoidance procedures, the Census Bureau modifies or removes the characteristics that put information at risk of disclosure. Although it may appear that a table shows information about a specific survey unit, the Census Bureau has taken steps to disguise or suppress a unit's data that may be "at risk" of disclosure while making sure the results are still useful.

[4] For more information regarding New Hampshire modeling for the 2021 VIUS including methodology and data tables, refer to <www.census.gov/data/experimental-data-products/vius-national-estimates.html>.

Table 4.
Unit Response Rates (URR) and Total Quantity Response Rates (TQRR)
(In percent)

Registration geography	URR	TQRR number of vehicles	TQRR vehicle miles
United States[1]	48.8	46.7	41.2
Alabama	47.2	45.6	40.5
Alaska	39.2	37.2	34.0
Arizona	42.1	39.5	28.5
Arkansas	50.5	49.6	44.1
California	26.1	32.1	27.7
Colorado	47.0	47.7	43.6
Connecticut	54.6	57.3	50.8
Delaware	44.0	47.7	42.8
District of Columbia	31.8	42.0	36.1
Florida	46.1	46.8	40.7
Georgia	41.1	40.3	35.3
Hawaii	47.8	46.6	41.8
Idaho	54.5	54.1	49.7
Illinois	50.8	53.1	46.7
Indiana	46.1	53.2	43.6
Iowa	60.5	56.1	51.4
Kansas	48.4	50.7	45.6
Kentucky	50.7	53.2	47.3
Louisiana	39.6	40.3	37.0
Maine	55.7	51.4	46.5
Maryland	48.2	47.6	43.0
Massachusetts	52.4	50.1	44.4
Michigan	52.1	49.8	45.3
Minnesota	55.4	58.1	54.4
Mississippi	37.8	38.9	33.4
Missouri	53.1	56.1	49.5
Montana	47.7	48.6	43.0
Nebraska	52.2	53.4	48.6
Nevada	44.5	43.7	37.7
New Jersey	41.6	49.2	44.4
New Mexico	45.6	45.6	40.3
New York	49.7	50.0	42.7
North Carolina	49.5	50.5	46.2
North Dakota	53.9	46.5	43.1
Ohio	52.9	52.1	46.0
Oklahoma	38.7	39.2	34.1
Oregon	55.1	53.9	49.8
Pennsylvania	66.8	58.5	44.9
Rhode Island	46.6	54.4	48.0
South Carolina	45.5	45.1	41.2
South Dakota	53.2	53.7	50.2
Tennessee	48.4	48.1	43.6
Texas	37.8	33.3	30.5
Utah	49.8	52.8	47.4
Vermont	57.2	57.2	50.7
Virginia	54.6	58.4	51.4
Washington	54.0	57.8	52.2
West Virginia	53.0	54.4	46.4
Wisconsin	59.4	61.1	56.1
Wyoming	51.9	51.6	47.7

[1] Excludes New Hampshire.
Source: U.S. Census Bureau, 2021 Vehicle Inventory and Use Survey.

The disclosure protections for the 2021 VIUS publication tables are described as a combination of the following methods:

Aggregation: For example, lowest published geography level is state.

Methods of grouping: For example, due to rarity, model years prior to a specified year are combined) or recoding (e.g., to ranges) unusual or rare characteristics/values.

Rounding: For example, number of vehicles estimates are rounded to the nearest 100; vehicle miles estimates are rounded to the nearest 100,000; and percentages are rounded to the nearest tenth so that estimates of the number of vehicles equal less than 50; mileages less than 50,000; and percentages less than 0.05 are replaced with a "Z" indicating these estimates round to zero. These data are still included in higher level aggregate totals.

Weighting: Examples of weighting include the application of sampling, nonresponse adjustment, and post-stratification weights.

Further, any estimate with fewer than three individual contributors in the sample is suppressed on publication tables.

The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data (Project No. P-7527235, Disclosure Review Board [DRB] approval number: CBDRB-FY23-0032).

REFERENCES

Cochran, W. G., *Sampling Techniques*, Third ed., John Wiley & Sons, 1977a, Chapter 5.10, "Stratified Sampling for Proportions," pp. 107–109.

Cochran, W. G., *Sampling Techniques*, Third ed., John Wiley & Sons, 1977b, Chapter 5A.14, "Estimating Totals and Means Over Subpopulations," pp. 142–143.

Lee, Eun Sul, *Analyzing Complex Survey Data*, Second ed., Vol. 71, Series: Quantitative Application in the Social Sciences, 2006, Chapter 4, "Strategies for Variance Estimation. The Taylor Series Method (Linearization)," pp. 36–39.