**March 2023**

# National Crime Victimization Survey: State-Level Analysis User's Guide

Prepared for

**U.S. Department of Justice**

**Bureau of Justice Statistics**

810 7th Street, Northwest
Washington, DC 20531

Prepared by

**Andrew Moore, MStat**

**Marcus Berzofsky, DrPH**

RTI International
3040 E. Cornwallis Road
Research Triangle Park, NC 27709

RTI Project Number 0217713.000.006.003

# CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# INTRODUCTION

The National Crime Victimization Survey (NCVS), sponsored by the Bureau of Justice Statistics (BJS), estimates the incidence of and describes the characteristics of criminal victimization in the United States. The purpose of this document is to describe the guidelines developed by BJS for producing state-level estimates with NCVS data and additional considerations and recommendations that researchers should bear in mind when conducting state-level analyses. Strategies for avoiding potential pitfalls and ensuring the reliability of results and conclusions from subnational analyses are also provided. Example code in SAS is also included to illustrate methods for generating common estimates and comparisons produced with NCVS data.

Beginning in 2016, the NCVS sample size was increased in large states to enable the creation of direct estimates for the 22 most populous states and large local areas within those states (as of the 2010 Census, all 22 states have a population of 5,000,000 or more persons). This sample boost, which was designed to support both national and subnational estimates, required significant changes to the NCVS sample design. In addition to increasing the overall sample size, the redesigned NCVS included adjustments to the sample allocation process and weighting methodology to ensure estimates of victimization based on the NCVS sample would be representative of the population within these states. The 22 states identified for state-level estimates are shown below. State-level victimization estimates can be found in *Criminal Victimization in the 22 Largest States, 2017–2019* (NCJ 305402, BJS, March 2023). Additional information on the history of the NCVS subnational estimation program is also provided in Appendix C.

| **State** | **State** |
|---|---|
| Arizona | Missouri |
| California | New Jersey |
| Colorado | New York |
| Florida | North Carolina |
| Georgia | Ohio |
| Illinois | Pennsylvania |
| Indiana | Tennessee |
| Maryland | Texas |
| Massachusetts | Virginia |
| Michigan | Washington |
| Minnesota | Wisconsin |

Although the sample redesign was implemented beginning in 2016, data from the 2016 survey should not be used to produce state-level estimates. Because of the large increase in the 2016 sample size and a switch to using population counts based on the 2010 decennial census, the 2016 sample included a large percentage of new households and first-time interviewees. These interviews were also more likely to be conducted by inexperienced interviewers. It is currently unknown to what extent these factors affected estimates of victimization based on the

2016 data. Therefore, state-level estimates should only utilize NCVS data from 2017 and later. For additional information on the 2016 redesign and potential impacts on estimates, see *Criminal Victimization 2016, Revised* (Morgan & Kena, 2018).

The goal of the sample redesign developed by BJS, in collaboration with the Census Bureau, was to enable the creation of reliable state-level estimates of violent victimization with a relative standard error (RSE) of no more than 10% using 3 years of data. For higher-crime states, states with the largest sample sizes, or crime types with higher rates (e.g., property crime), fewer years of data may be sufficient to reach this threshold. For smaller areas or for subgroup analyses of violent or property crime, more than 3 years of data may be necessary. With current sample sizes, a minimum of 3 years of data should be used to ensure estimates achieve adequate precision.

The remainder of this document provides additional resources for analysts conducting state-level analyses with NCVS data. Section 1 provides an overview of the appropriate use of the restricted-use data files. Section 2 includes evaluation criteria that analysts should utilize to evaluate the reliability of their analyses. Section 3 describes steps that researchers can take to remediate issues with state-level analyses. Finally, Section 4 and Section 5 provide guidance and example code for producing victimization estimates and conducting statistical tests to detect significant differences, respectively.

# 1. NCVS SUBNATIONAL ESTIMATES—
## OVERVIEW OF USES, LIMITS, AND CAVEATS

In redesigning the NCVS to fulfill the dual purpose of supporting both national and subnational estimates, BJS had to balance the increased costs associated with a larger and more geographically dispersed sample with the benefits of enhanced utility for local stakeholders and others interested in producing subnational estimates. The sample boost approach offers many benefits, including the use of direct observation from sample members within the areas of interest, uniformity between the national and state-level methodologies, a reduction in coverage error for states included in the sample boost, and the ability to implement the methodology for both the core NCVS and supplemental surveys. Pursuing a direct estimation approach also simplifies analyses for data users. The complex modeling or reweighting approaches that have been used to produce subnational estimates prior to the redesign are no longer needed to produce state-level estimates for areas included in the sample boost. However, the sample boost approach limits the number of years and areas for which subnational estimates can be produced.

Although geographic identifiers (e.g., state) are available on the restricted-use data files for all survey years, the sampling and weighting methodologies implemented prior to 2016 were not designed to produce state-level estimates. That is, prior to 2016, the sample was selected and weighted to be representative of the nation as a whole. Samples within states are unlikely to be representative of the population within those states in terms of geographic (e.g., rural or urban) and demographic (e.g., age, race) characteristics for 2015 and earlier. Therefore, analysts should not use data from survey years prior to 2017 to produce direct state-level estimates with the sample boost data.[1,2]

Similarly, geographic identifiers on the restricted-use data files should not be used to generate victimization estimates for states that were not part of the sample boost. The sample within states that were not part of the sample boost was selected and weighted to ensure national estimates could still be produced. That is, the combined data that includes respondents within the 22 most populous states and respondents in non-boost states are weighted to represent the nation as a whole. However, respondents from non-boost states are not representative of the population within those individual states and direct estimates of victimization within non-boost states would be subject to bias and low precision.

In addition to state-level estimates, which are the focus of this user guide, the redesigned NCVS was intended to allow direct estimates to be produced for large local areas (e.g., metropolitan statistical areas, or MSAs) within the 22 most populous states. However, the validity and reliability of these substate estimates have not been fully examined by BJS. As such, it is currently unknown how various error sources (e.g., coverage, nonresponse, sampling) would affect individual substate estimates and how many years of data would be necessary to produce

---

[1] Although the sample boost was implemented beginning in 2016, data from 2016 should not be used to produce direct state-level estimates with the sample boost data. For more information, see the discussion in the introduction.

[2] For subnational victimization estimates prior to 2017, one of the other subnational approaches outlined in Appendix C (e.g., model-based estimates) should be used.

estimates with adequate precision. While substate estimate development and validation is currently a priority of ongoing research for BJS, guidance on producing victimization estimates for local areas within the 22 most populous states that were part of the sample boost is outside the scope of this user guide.

Further, state-level identifiers are not currently included on NCVS public-use files. Therefore, analysts must use the restricted-use data files available at the U.S. Census Bureau headquarters or at a Census Research Data Center (RDC). As such, any state-level analyses conducted with NCVS data will also be subject to additional Census Bureau restrictions related to disclosure avoidance. Therefore, researchers are encouraged to review the *Federal Statistical Research Data Center's Handbook for Researchers*[3] and to discuss any analyses planned with their Census Bureau sponsor to ensure Disclosure Review Board (DRB) guidelines for the release of data products will be met.

*Table 1* summarizes the data and geographic identifiers available on the NCVS public-use files, available from the Inter-university Consortium for Political and Social Research (ICPSR),[4] and the restricted-use files, available within a Census RDC. Information about the use of this data to produce victimization estimates is also provided.

**Table 1:     NCVS data: Summary of uses, limits, and caveats**

| File and Geography | Availability of Geographic Identifiers | Survey Years Available for Analysis | Require DRB Review? | Minimum # of Years to Include in Analysis |
|---|---|---|---|---|
| Public-use data files | | | | |
| National | X | 1992–present | N/A | 1 |
| Boost states | | | | |
| Non-boost states | | State identifiers not available for analysis | | |
| Substate areas | MSA[a] | 2000–2015 | N/A | Sliding scale |
| Restricted-use data files | | | | |
| National | X | 2005–present[b] | X | 1 |
| Boost states | X | 2017–present | X | 3 |
| Non-boost states | X | | | |
| Substate areas | X | Direct estimation not recommended | | |

MSA=metropolitan statistical area; N/A=not applicable.
[a] MSA identifiers available on *National Crime Victimization Survey: MSA Public-Use Data, 2000-2015* (Bureau of Justice Statistics, 2022).
[b] Additional years of data may be available upon request.

---

[3] At the time of this report, the most recent researcher's handbook, *The Researcher Handbook* (U.S. Census Bureau, 2020) is available for download at https://trdc.cpc.unc.edu/wp-content/uploads/2021/08/Researcher_Handbook_1208020.pdf. Analysts should check with their Census Bureau sponsor to determine if guidelines have changed.
[4] NCVS public-use data files are available at https://www.icpsr.umich.edu/web/ICPSR/series/95.

## 2. EVALUATION CRITERIA FOR STATE-LEVEL ESTIMATES

Although the NCVS has been redesigned to allow direct estimates of criminal victimization to be produced for the 22 most populous states, analysts should exercise prudence when generating and interpreting such estimates. The sample sizes and estimated precision of estimates may vary across various victimization types, from state to state, and over time. For example, although the redesigned NCVS aimed to allow state-level violent crime estimates to be produced with an RSE of no more than 10% using 3 years of data, only five states met this goal based on NCVS data from 2017 to 2019.

BJS, in consultation with RTI International, has evaluated the redesigned NCVS from a total survey error perspective to identify potential sources of error that may compromise the validity of state-level estimates (Moore et al. (2023). *National Crime Victimization Survey: Validation of state-level estimates* (NCJ 305403). RTI International.). This research concluded that victimization estimates meeting BJS's standards for quality and precision can be produced and released for the 22 most populous states beginning with the 2017 collection year and utilizing a minimum of 3 years of data. However, this evaluation was focused on identifying overarching issues that would call into question the validity of the underlying methodology. It was not possible to cover all potential analyses that could be undertaken by a researcher. As such, it is the responsibility of the data user to ensure the validity of any analyses conducted using state-level data. The remainder of this section discusses potential error sources that should be considered when producing state-level estimates, particularly for analyses within subdomains (e.g., age group, race/ethnicity).

## 2.1 Coverage Quality for Critical Subdomains

Coverage error occurs when the sampling frame excludes a portion of the target population or when certain groups of the target population are overrepresented on the frame. If the excluded population or overrepresented groups are systematically different from the target population, then estimates could be biased. Coverage rates can be calculated by comparing the sum of the weights in the NCVS data for the state(s) and time period of interest across various characteristics (e.g., age, race/ethnicity, education, marital status) to the same population estimates from a gold-standard source such as the American Community Survey or the U.S. Census Bureau's Population Estimates Program.

For person-level crimes (e.g., violent crime), the bias due to coverage error is expected to be minimal for overall state-level victimization estimates. Similarly, state-level estimates by sex, age group, race, or marital status of the victim are unlikely to be affected by coverage error as the coverage ratios are near 1 for most states in the sample boost from 2017 to 2019 (see ***Table A.1*** in Appendix A). However, within a few states (e.g., GA, IN), the coverage ratios indicate that younger persons are underrepresented while persons classified as "other/more than one race" may be underrepresented in some states (e.g., MI, MN) and overrepresented in others (e.g., NY, NC). Persons who rent their current residence and the unemployed also tend to be overrepresented. Coverage rates by educational attainment can also vary considerably across

states, though respondents with a high school education or some college generally tend to be underrepresented.

For household-level crimes (e.g., property crime), overall household-level coverage ratios are near 1 for all states from 2017 to 2019 (see **Table A.2** in Appendix A). However, coverage ratios within subdomains defined by characteristics of the household differ from 1 more often than those defined by person-level characteristics. In particular, the following groups tend to be overrepresented in many states (e.g., CA, WI): householders age 30 or younger; householders who are Black or classified as "other/more than one race;" households that are being rented; householders with a high school diploma or less education; and households in multiunit structures. Undercoverage by more than 10% was less common at the household level but was observed for some characteristics and states (e.g., householders ages 50 to 64 in MD and PA).

Analysts should be aware of potential coverage issues when calculating subdomain estimates within states as coverage error can result in biased estimates. Additional caution should be used when comparing subgroups or evaluating changes over time as differences in the coverage rates for subgroups or time periods could result in erroneous conclusions about whether groups are significantly different or whether victimization rates are changing over time for a particular subgroup. This is particularly true if comparing victimization totals as this type of analysis can be more susceptible to differences in coverage rates. If the analysis includes subdomains not included in **Tables A.1** and **A.2**, analysts are encouraged to evaluate the domain(s) of interest for possible under- or overcoverage.[5]

## 2.2 Response Rates for Critical Subdomains

Nonresponse error quantifies the error due to a portion of the sample not participating in the survey. If the nonresponding units are different from responding units with respect to the outcome of interest, then estimates can be subject to nonresponse bias. Note, however, that the bias due to nonresponse cannot be measured directly because the values for nonresponding units are unknown. That is, nonresponse bias is a function of both the response rate and the difference between respondents and nonrespondents for the outcome being measured (Biemer, 2010). Therefore, a low response rate, by itself, is not necessarily an indication of bias. However, the precision of estimates is likely to be affected by nonresponse due to smaller sample sizes, because response rates tend to be inversely related to the likelihood or risk of nonresponse bias.

**Tables A.3** and **A.4** in Appendix A show person-level response rates by various characteristics and overall household-level response rates by state, respectively, for the aggregate period covering 2017 to 2019. Household-level response rates are generally 70% or higher for most states. Within responding households, person-level response rates are also 70% or higher for most states and subdomains. Note, however, that because person-level response rates are calculated within responding households, the overall response rate (i.e., the product of the household-level response rate and the person-level response rate) would be considerably lower. Response rates tend to be lower for younger persons, persons who have never been married, and

---

[5] An example of performing a coverage analysis can be found in *Assessing the Coverage and Reliability of Subnational Geographic Identifiers in the NCVS Public-Use File* (Bureau of Justice Statistics, 2015).

persons with less than a high school diploma or equivalent. Analysts should exercise caution when analyzing subgroups with low response rates. As noted previously, however, a low response rate is not necessarily indicative of nonresponse bias. To evaluate how respondents and nonrespondents differ, a relative bias analysis can also be conducted, as described in Section 2.3.

## 2.3    Relative Bias

Because the true values for victimization outcomes are not measured or known for nonresponding households or persons, the actual bias due to nonresponse cannot be measured. A proxy for this bias can, however, be estimated by comparing characteristics of respondents with the overall sample. The relative bias of the responding sample is calculated by comparing the percentage of respondents with various characteristics to the percentage of persons or households in the overall sample with the same characteristics, and then calculating the relative difference. For example, if 50% of the sample is male for a given state but only 48% of respondents are male, then the percent relative bias is calculated as (48%–50%)/50%*100 = −4%. A negative relative bias indicates the respondents underrepresent a particular group, whereas a positive value indicates overrepresentation. Substantial differences between respondents and the overall sample may be an indicator of nonresponse bias.

*Tables A.5* and *A.6* in Appendix A show the percent relative bias by person- and household-level characteristics, respectively, by state from 2017 to 2019. Relative to the sample, persons ages 12 to 17 and 18 to 24 are generally underrepresented among respondents. Non-White respondents are also underrepresented in many states (e.g., MN, PA, WA) relative to the sample within those states. Analysts should exercise caution when analyzing subgroups with large relative bias values as a large relative bias is associated with less precise estimates.

## 2.4    Unweighted Sample Size for Critical Subdomains

The standard reporting convention used by BJS is to flag estimates that are based on 10 or fewer sample cases to alert readers that the estimate may be unreliable. Data users are also encouraged to follow this same practice by flagging estimates based on 10 or fewer sample cases. However, due to rounding rules imposed by the U.S. Census Bureau's DRB for unweighted counts, a higher threshold of 15 or fewer sample cases should be used to flag estimates within a Census RDC. In addition, victimization estimates based on a small number of sample cases must be suppressed prior to submission to the Census DRB.[6]

## 2.5    Standard Error and Relative Standard Error

Sampling errors arise when a sample, rather than a census, is used to estimate characteristics of a target population. The difference between a statistic generated from the sample and the unknown population parameter is the sampling error. These types of errors affect both the bias and variance of statistics generated from a sample and are dependent on characteristics of the sample design, including the methods used to select the sample and the size

---

[6] Note that the suppression and rounding rules imposed by the DRB may change over time. Data users are encouraged to check with their RDC sponsor to ensure all rules related to the release of data products are followed.

of the sample (Biemer, 2010). Because the true population parameters are unknown (e.g., the rate of violent crime victimization in Arizona), it is not possible to estimate the contribution of bias to the sampling error of state-level estimates. However, the RSE of state-level estimates produced from the sample is one way to quantify the sampling error.

The RSE of an estimate, which is the ratio of the standard error and the estimate, is typically expressed as a percentage. RSEs provide a measure of the precision of an estimate, relative to the size of the estimate, and can also be used to compare the relative precision of different estimates. That is, large RSE values are associated with less precise estimates and small RSE values are indicative of more precise estimates.

*Table A.7* in Appendix A shows the RSEs of victimization rates for various crime types by state based on the 2017 to 2019 NCVS data. *Tables A.8* and *A.9* present the RSEs of the violent crime rate and property crime rate, respectively, by various subdomains and state. BJS typically flags estimates with an RSE of 50% or greater to indicate that the estimate may be unreliable. Data users are also encouraged to calculate standard errors (see Section 4) and RSEs for state-level victimization estimates. As with estimates based on a small number of cases, it is recommended that state-level estimates with an RSE of 50% or greater be flagged as unreliable. However, a lower threshold (e.g., 10%, 30%) may also be used, if desired.[7]

## 2.6    Outlier Identification and Assessment

As part of the NCVS weighting process, a small number of respondents each year are classified as outliers. These individuals are outliers in the sense that they have a victimization profile that is dissimilar from most other individuals in the target population (based on the number of victimizations reported), yet they also represent a lot of individuals in the target population (based on the size of their analysis weight). As a result, this small number of individuals can have a large impact on some annual estimates. To mitigate their influence and this issue, the weights assigned to these outlier cases are trimmed and the excess weight is redistributed to other respondents.[8]

This outlier detection and remediation process is implemented at the national level, rather than within individual states, and within each individual data collection year. Therefore, it is possible that within an individual state, and after aggregating multiple years of data, additional respondents could meet the "outlier" threshold. Although data users are not expected to reweight the data when producing state-level estimates, users are encouraged to evaluate whether a single individual, or a small number of individuals, are accounting for a large proportion of the estimate. While the definition of "large" is subjective, it is recommended that analysts flag estimates where one respondent within a reporting period contributes 20% or more of the total number of victimizations and give pause to presenting an estimate when a single contribution accounts for 50% or more of the estimate. For example, if in a given state and time period, the total number of rape/sexual assault (RSA) victimizations is estimated at 60,000 and one

---

[7] The standard practice followed by the Census Bureau is to flag estimates with an RSE of 30% or greater.
[8] For more information on the outlier detection and remediation process, see *National Crime Victimization Survey: Assessment of Outlier Weights* (Bureau of Justice Statistics, 2021).

respondent reported 12 RSA incidents during a single interview and had a victimization weight of 1,500, then that individual would account for 18,000, or 30%, of the 60,000 estimated RSA victimizations. If situations like this are identified, it is recommended that analysts flag the estimate and consider implementing some of the strategies discussed in Section 3.2.

## 2.7    Summary of Evaluation Criteria

*Table 2* summarizes the main findings from the *National Crime Victimization Survey: Validation of State-Level Estimates* report and recommendations for data analysts to follow when producing state-level estimates.

**Table 2:    Summary of validation report findings and recommendations**

| Findings from State-Validation Report | Recommendation Action when Producing State-Level Estimates |
|---|---|
| Estimates in some states may not be representative if underrepresented groups (e.g., persons with a high school education) or overrepresented groups (e.g., persons with a college education) are systematically different with respect to victimization. | Evaluate coverage for population subgroups included in analysis as described in Section 2.1 |
| Higher levels of nonresponse for some population subgroups (e.g., persons ages 12 to 15) may negatively impact precision and increase nonresponse bias. | Evaluate response rates and relative bias for population subgroups as described in Sections 2.2 and 2.3 |
| Victimization estimates, particularly for rarer crime types, may be significantly influenced by series weights. State-level estimates are more susceptible than national-level estimates to being influenced by respondents with a large series weight because of the smaller sample sizes. | Evaluate whether a single respondent is accounting for a large proportion of the total victimization rate or total as described in Section 2.6 |
| Estimate precision generally failed to meet stated goal, but few estimates were flagged as unreliable. | Flag estimates based on a small number of sample cases and estimates with a large RSE as described in Sections 2.4 and 2.5 |

# 3. INTERPRETATION OF EVALUATION CRITERIA AND MITIGATION STRATEGIES

## 3.1    Interpretation of Evaluation Criteria

Although BJS found no overarching issues that would call into question the validity of state-level victimization estimates produced from the redesigned NCVS sample, it is not possible to anticipate every type of analysis that may be conducted by data users. Further, due to the complexity of the NCVS and differences in the sample designs employed across similar large national surveys, identifying benchmarks for many of the recommended quality measures (e.g., coverage rates, response rates) can be challenging. Therefore, researchers are encouraged to take a holistic view of the various indicators of quality when evaluating analysis options.

## 3.2    Mitigation Strategies

If the recommended steps and considerations outlined in Section 2 elicit concerns about the validity and reliability of a particular estimate or analysis, one or more of the following actions should be taken in an attempt to address those concerns.

- **Aggregate additional years of data**: Combining additional years of data (i.e., more than 3 years) will increase the number of respondents included in the analysis. This strategy will be particularly useful when dealing with suppressed estimates and large RSE values.

- **Collapse subdomain categories**: Sample sizes may be insufficient to analyze many subdomains (e.g., age group categories) within states, particularly for rare crime victimization types. Collapsing small subgroups can increase the sample size for the combined category and may mitigate issues related to poor coverage, low response rates, suppression, large RSEs, and outliers.

- **Exclude the state, crime victimization type, or domain from the analysis**: The methodological assumptions guiding the redesign of the NCVS were focused on ensuring the reliability of state-level victimization estimates for the entire population of households and persons of interest within those states. Analysis of some subdomains and rare crime victimization types may not be feasible with current sample sizes and the number of years of data available. Therefore, if the first two strategies are ineffective at resolving a data user's concerns, it is recommended that the particular state(s), crime victimization type(s), or subdomain(s) in question be excluded from the analysis.

After implementing one or more of these strategies, the quality indicators should be reevaluated using the evaluation criteria described in Section 2 to determine if the issues that provoked the initial concerns have been adequately addressed. Data users are also encouraged to disclose these limitations within published reports to clarify why certain analyses or subgroups may have been excluded from the results.

*Table 3* summarizes potential issues with state-level estimates and guidelines for implementing the various mitigation strategies. However, the proposed order of implementation may require modifications based on the goals of the individual analysis.

**Table 3:    Summary of mitigation strategies and recommended order of implementation**

| Issue Noted with Evaluation Criteria | Mitigation Strategy | | |
|---|---|---|---|
| | Include More Years of Data | Collapse Subdomains | Exclude State, Crime Type, or Subdomain |
| Coverage quality | — | 1 | 2 |
| Response rates | — | 1 | 2 |
| Relative bias | — | 1 | 2 |
| Small Sample Size | 1 | 2 | 3 |
| Low precision/High RSE | 1 | 2 | 3 |
| Outliers | 1 | 2 | 3 |

Note: Mitigation strategies marked as "—" are unlikely to be effective at remediating the issue.

# 4. GENERATING VICTIMIZATION ESTIMATES AND STANDARD ERRORS

Much of the material in this section is taken from the User's Guide to the National Crime Victimization Survey (NCVS) Direct Variance Estimation report (Shook-Sa, Couzens, & Berzofsky, 2015) and the National Crime Victimization Survey: MSA Public-Use Data, 2000-2015 codebook (Bureau of Justice Statistics, 2022). The material has been included here for convenience and, where applicable, to provide additional details specific to working with the restricted-use data files in a Census Bureau RDC.

## 4.1    Data Structure

The NCVS data files available within the Census Bureau RDC are available in SAS format and are organized into three file types: household-level files, person-level files, and incident-level files. In contrast to the NCVS public-use files (PUFs), where all data of a particular type (e.g., household-level) is concatenated into a single annual file, the restricted-use files for a given year and type are typically separated into either two (i.e., semiannual) or four (i.e., quarterly) files. To create an annual file, the quarterly or semiannual files can be combined using a SET statement in SAS.

### 4.1.1    Household-Level File

The household-level file contains one record for each sampled household per interview period and is most commonly used to calculate property victimization rates. This file contains information about the geographic (e.g., state) and demographic characteristics of the household as well as information about the principal person and reference person. Information from the household screening interview, which assesses whether the household experienced any property victimizations during the previous 6 months, is also included in the household-level file.

Because the NCVS is a panel survey with sampled households being interviewed every 6 months over a 3-year period, most households will have two records within a given year. Therefore, both the household identifier (CTRLNUM) and the year/quarter identifier (YEARQUARTER) must be used to uniquely identify households by reporting period when merging the household-level file with the person- or incident-level files. Household-level victimization estimates utilize the household weight (HHWEIGHT). Note, however, that the variable HHWEIGHT is equivalent to V2116 on NCVS PUFs and must be divided by 2 to create a variable equivalent to the adjusted collection-year household weight (i.e., WGTHHCY on NCVS PUFs). Two additional variables on the household-level file that are important for calculating the standard errors of victimization estimates include the pseudo-stratum (UCF_PSEUDOSTR) and half-sample code (UCF_HALFSAMPCD). A crosswalk between the restricted-use files and public-use files for select household-level variables is provided in *Table D.1* in Appendix D.

### *4.1.2 Person-Level File*

The person-level file contains one record for each household member age 12 or older in responding households per interview period and is most commonly used to calculate person-level victimization rates (e.g., violent crime rate). This file contains information about each sampled person (e.g., age, race, sex) as well as information from the person-level screening interview which determines whether a person experienced a personal victimization during the previous 6 months.

As with the household-level file, the person-level file includes 2 records for most individuals within a given year. Therefore, in addition to the household identifier (CTRLNUM) and year/quarter identifier (YEARQUARTER), a person-within-household identifier (LINENUM) must also be used to uniquely identify a person within a reporting period when merging the person-level file with the incident-level file. Person-level victimization estimates utilize the person weight (PERSONWEIGHT). This weight is equivalent to variable V3080 on NCVS PUFs and must be divided by 2 to create a weight that is equivalent to the adjusted collection-year person weight (i.e., WGTPERCY on NCVS PUFs). A crosswalk between the restricted-use files and public-use files for select person-level variables is provided in ***Table D.2*** in Appendix D.

### *4.1.3 Incident-Level File*

The incident-level file contains more detailed information about victimizations identified during the household screening interview and the person screening interview. This includes information that can be used to classify the incident into a specific crime type (e.g., burglary, robbery, aggravated assault) as well as additional details of the incident (e.g., presence of a weapon, victim-offender relationship).

Generally, the incident file will contain one record for each victimization reported by NCVS respondents. In some instances, however, a single incident record is used to represent multiple victimizations. This occurs when the respondent reports experiencing six or more similar incidents during the previous 6 months, but they are unable to recall specific details about each individual incident. These types of victimizations are referred to as series crimes and are represented by a single record in the incident file.

NCVS respondents (households or persons) who report multiple victimizations during a reporting period will have multiple records in the incident file for that period. However, if the respondent does not report any victimizations, then they will not have a corresponding record in the incident-level file. Property crimes in the incident-level file can be linked to the household-level file using the household identifier (CTRLNUM) and year/quarter identifier (YEARQUARTER). Person-level crimes in the incident-level file can be linked to the person-level file using the household identifier (CTRLNUM), year/quarter identifier (YEARQUARTER), and the person-within-household identifier (LINENUM).

The incident-level file includes the victimization weight (VWGT) as well as a victimization weight that has been adjusted for series crimes (SERIESWGT). An incident weight adjusted for series crimes (SERIESINCWGT) is also available on the restricted-use files. A

crosswalk between the restricted-use files and public-use files for select incident-level variables is provided in ***Table D.3*** in Appendix D.

### 4.1.4   *Victimization Summary File*

Although not a standard NCVS file, a victimization summary file is used in many of the examples provided in Sections 4 and 5. The victimization summary file is used to calculate victimization rates and contains data about reported victimizations from the incident-level file. For person-level crimes, the summary file contains one record for each person (CTRLNUM and LINENUM) and reporting period (YEARQUARTER). For property crimes, the summary file contains one record for each household (CTRLNUM) and reporting period (YEARQUARTER). For both person-level crimes and household-level crimes, the victimization summary file contains the number of series-adjusted victimizations with the specified characteristic(s) of interest (e.g., violent crimes involving a weapon) and the victimization weight. The summary file is then merged with the person-level or household-level file for the calculation of victimization rates and standard errors.

## 4.2   Variance Estimation

Variance is a measure of the sampling error of an estimate and provides an indication of the reliability of the estimate produced from a sample. Two general methods exist for calculating the variance of estimates that account for the complex sample design of the NCVS: (1) generalized variance functions (GVFs), and (2) direct variance estimation. Two direct variance estimation approaches, balanced repeated replication (BRR) and Taylor series linearization (TSL), can be used with the NCVS PUFs. However, the restricted-use data files available within the RDC do not include BRR weights available on the PUF. Further, GVF parameters are only available at the national level and should not be used for subnational estimates. Therefore, only the TSL variance estimation method will be discussed in the examples below.

## 4.3   Victimization Rates

The victimization rate is the ratio of the number of victimizations and the population size (typically expressed per 1,000 households or persons). Since the number of victimizations is calculated using the incident-level file (and victimization weight) and the population size is calculated using the household-level file (and household weight) or the person-level file (and person weight), additional steps are necessary to prepare the data for analysis. After concatenating the quarterly or semiannual files to create an incident-level file and a person or household-level file with the years of interest, the process for calculating victimization rates generally involves the following steps.[9] A graphical representation of this process is also provided in ***Figure B.1*** of Appendix B.

---

[9] For additional information, see *User's Guide to National Crime Victimization Survey (NCVS) Direct Variance Estimation* (Shook-Sa et al., 2015).

1. Identify records on the incident file with the victimization characteristic(s) of interest.
2. Create a victimization summary file from the incident file. The summary file includes the series-adjusted number of victimizations for the crime type(s) of interest within a reporting period.
3. Merge the victimization summary file onto the household-level file, for property crimes, or merge the victimization summary file onto the person-level file, for personal crimes. If the crime type being estimated is a person-level crime, additional information from the household-level file (i.e., sample design variables and state indicators) will also need to be merged with the person-level file.
4. Calculate a victimization adjustment factor. The victimization adjustment factor is equal to the ratio of the victimization weight and the adjusted household or person weight. If the respondent did not report any victimizations for the crime type of interest, the adjustment factor will equal 0. The series-adjusted number of victimizations from Step 2 is then multiplied by the victimization adjustment factor and 1,000 (to express the rate per 1,000 households or persons) to create the variable needed for analysis.
5. Calculate the victimization rate and TSL standard error.

### 4.3.1 Example SAS Code

The following example will illustrate the process for calculating the violent victimization rate for California from 2017 to 2019 using SAS.[10]

**Step 1: Identify Records with Victimization Characteristic(s) of Interest**

```
data ex1_incident;
 set incident1719; *Concatenated incident file;
* Create an indicator of violent crime: 0=Not a violent crime; 1=violent crime;
 VIOLENT=(1 <= TOCNEW <= 20);
* Create an indicator for crimes that occurred outside the United States: 1=crime occurred outside United States;
 EXCLUDE_OUTUS=(INCIDENTPLACE='1');
* Create a series crime indicator: 0=Not a series crime; 1=Series crime *;
 if (SIXORMOREINCIDENTS in ('1','8')) or (INCIDENTSSIMILAR in ('2','8')) or (RECALLDETAILS in ('1','8'))
then series=0;
 else series=1;
* Calculate the number of incidents for series crimes;
 SERIESWEIGHT=SERIESWGT/VWGT;
run;
```

**Step 2: Create a Victimization Summary File**

```
proc sort data=ex1_incident; by YEARQUARTER CTRLNUM LINENUM; run;

proc means data=ex1_incident noprint;
where EXCLUDE_OUTUS=0 and VIOLENT=1; * Exclude crimes occurring outside the United States and subset file to crime type of interest. This also ensures the appropriate weight (VWGT) is kept on the file if a respondent reported both property crimes and personal crimes;
weight SERIESWEIGHT;
id VWGT;
by YEARQUARTER CTRLNUM LINENUM;
var VIOLENT;
output out=ex1_victimization_summary sum=;
```

---

[10] For assistance with SAS procedures, please reference the SAS documentation available at https://support.sas.com/en/documentation.html.

```
run;
```

**Step 3, Part 1: Merge the Victimization Summary File onto the Household-Level File (for Property Crimes) or Person-Level File (for Person-Level Crimes)**

```
proc sort data=person1719 out=ex1_person;
by YEARQUARTER CTRLNUM LINENUM;
run;

data ex1_merged_file;
 merge ex1_person ex1_victimization_summary;
 by YEARQUARTER CTRLNUM LINENUM;
* The incident count variable is missing for persons not included on the victimization summary file, so they are set
to '0' (no victimizations of this type);
 if VIOLENT=. then VIOLENT=0;
run;
```

**Step 3, Part 2: Merge Design Variables and State Indicator onto the Person-Level File**

```
data ex1_hhld;
 set hhld1719; /* Concatenated household-level file */
 * Calculate a numeric PSEUDOSTRATA;
 PSEUDOSTRATA=UCF_PSEUDOSTR*1;
 * Calculate a numeric HALFSAMPLE code;
 HALFSAMPLE=UCF_HALFSAMPCD*1;
 * Calculate a numeric state indicator: STATE is a character variable indicating the state FIPS code;
 STATENUM=STATE*1;
 keep YEARQUARTER CTRLNUM STATENUM PSEUDOSTRATA HALFSAMPLE;
run;

proc sort data=ex1_hhld; by YEARQUARTER CTRLNUM; run;

data ex1_merged_file2;
 merge ex1_merged_file(in=in1) ex1_hhld;
 by YEARQUARTER CTRLNUM;
 if in1;
run;
```

**Step 4: Calculate the Victimization Adjustment Factor**

```
data ex1_analysis_file;
 set ex1_merged_file2;
* Adjust the person weight;
 PERSONWEIGHT2=PERSONWEIGHT/2;
* Calculate the adjustment factor;
 if VWGT > 0 then ADJINC_WT=VWGT/PERSONWEIGHT2;
 else ADJINC_WT=0;
* Create an analysis variable equal to the victimization count multiplied by the adjustment factor multiplied by
1,000 (to express the rate per 1,000 persons);
 ANALYSISVAR=VIOLENT*ADJINC_WT*1000;
run;
```

**Step 5: Calculate the Victimization Rate and Standard Error**

```
proc surveymeans data=ex1_analysis_file varmethod=taylor mean stderr;
strata PSEUDOSTRATA;
cluster HALFSAMPLE;
domain statenum('6'); * California: STATENUM is a numeric variable indicating the state FIPS code (see Step 3);
```

```
weight PERSONWEIGHT2; * Adjusted person weight - collection year;
var ANALYSISVAR;
run;
```

## 4.4     Victimization Totals

Victimization totals are weighted counts of the number of criminal victimizations for a particular crime type of interest. In contrast to victimization rates, which require both the incident file (for the numerator) and the household or person file (for the denominator), victimization totals are calculated using only the incident file and the series-adjusted victimization weight (SERIESWGT).

Because the incident file only includes respondents who reported a victimization, additional steps are necessary to ensure standard errors properly account for the complex sample design. The incident file must contain at least one record for each NCVS primary sampling unit (PSU) identified by the pseudo-stratum (UCF_PSEUDOSTR) and half-sample code (UCF_HALFSAMPCD). This requires creating dummy records for each combination of pseudo-stratum and half-sample codes, and these dummy records must have a positive weight. An indicator of whether the record is a dummy record should also be created to allow these cases to be excluded from calculations.

Further, because multiple years of data must be aggregated to produce state-level victimization estimates, the victimization weight will also need to be adjusted. Within a single year, the weights, after being divided by 2 to account for the semiannual weighting process, are designed to be representative of the entire population of households or persons within that state. When creating state-level victimization estimates with sample boost data from the NCVS, a minimum of 3 years of data must be aggregated. Without an adjustment, the estimate would represent the total number of victimizations over the entire period rather than an annual average. Therefore, if 3 years of data are being aggregated (e.g., 2017–2019), for example, then the series-adjusted victimization weight would be divided by 3. The process for calculating victimization totals typically involves the following steps. A graphical representation of this process is also provided in *Figure B.2* of Appendix B.

1.  Merge the design variables and state indicators from the household-level file onto the incident-level file. Additional variables (e.g., age, sex) from the household or person file may need to be merged depending on the analysis.
2.  Using the household-level file, create a file with a dummy record for each PSU.
3.  Append the file from Step 2 with the incident-level file from Step 1.
4.  Create an indicator for the victimization characteristic(s) of interest (e.g., violent crime) and adjust the victimization weight to account for aggregating multiple years of data. A domain indicator should also be created to exclude dummy records and crimes occurring outside the United States. Additional information can also be included in the domain indicator (e.g., crime type) or can be added to the domain statement in the SAS SURVEYMEANS procedure.
5.  Calculate the victimization total and the TSL standard error.

### *4.4.1   Example SAS Code*

The following example will illustrate the process for calculating the annual average number of violent victimizations experienced by males and females in Arizona from 2017 to 2019 using SAS.

**Step 1: Merge Design Variables and State Indicators onto the Incident-Level File**

```
data ex2_hhld;
 set hhld1719; /* Concatenated household-level file */
 * Calculate a numeric PSEUDOSTRATA;
 PSEUDOSTRATA=UCF_PSEUDOSTR*1;
 * Calculate a numeric HALFSAMPLE code;
 HALFSAMPLE=UCF_HALFSAMPCD*1;
 * Calculate a numeric state indicator: STATE is a character variable indicating the state FIPS code;
 STATENUM=STATE*1;
 keep YEARQUARTER CTRLNUM STATENUM PSEUDOSTRATA HALFSAMPLE;
run;

proc sort data=ex2_hhld; by YEARQUARTER CTRLNUM; run;

proc sort data=incident1719; by YEARQUARTER CTRLNUM; run;

data ex2_incident;
merge incident1719(in=inINC) /* Concatenated incident file */
    ex2_hhld;
by YEARQUARTER CTRLNUM;
if inINC;
run;
```

**Step 1, Part 2 (Optional): Add Respondent Characteristic(s) to the Incident-Level File**

```
data ex2_person;
set person1719; * Concatenated person file;
* Create indicator of respondent gender;
if SEXALLOC='1' then gender=1; * male;
else if SEXALLOC='2' then gender=2; * female;
keep YEARQUARTER CTRLNUM LINENUM GENDER;
run;

proc sort data=ex2_person; by YEARQUARTER CTRLNUM LINENUM; run;

proc sort data=ex2_incident; by YEARQUARTER CTRLNUM LINENUM; run;

data ex2_incident2;
merge ex2_incident(in=inINC) ex2_person;
by YEARQUARTER CTRLNUM LINENUM;
if inINC;
run;
```

**Step 2: Create a Dummy Record for Each PSU**

```
proc freq data=ex2_hhld noprint;
 tables pseudostrata*halfsample/out=ex2_psu;
run;

data ex2_dummy;
 set ex2_psu;
```

```
 * Assign a positive weight to each dummy record;
 SERIESWGT=1;
 * Create a dummy indicator for each record;
 DUMMY=1;
run;
```

**Step 3: Append the Dummy File to the Incident-Level File**

```
data ex2_incident3;
 set ex2_psu
    ex2_incident2(in=inINC); * Concatenated incident file ;
 if inINC then dummy=0;
run;
```

**Step 4: Identify Records with Victimization Characteristic(s) of Interest**

```
data ex2_analysis_file;
 set ex2_incident3;
* Create an indicator of violent crime;
 VIOLENT=(1 <= TOCNEW <= 20);
* Create an indicator for crimes that occurred outside the United States;
 EXCLUDE_OUTUS=(INCIDENTPLACE=1);
* Adjust the victimization weight for aggregating multiple years of data;
 SERIESWGT3=SERIESWGT/3;
* Create a domain indicator to exclude crimes occurring outside the United States and dummy records and to subset
to the state(s) of interest. STATENUM is a numeric variable indicating the state FIPS code (see Step 1);
 sub=(EXCLUDE_OUTUS=0 and DUMMY=0 and STATENUM=4); * Arizona;
run;
```

**Step 5: Calculate the Victimization Totals and TSL Standard Errors**

```
proc surveymeans data=ex2_analysis_file varmethod=taylor sum stderr;
strata PSEUDOSTRATA;
cluster HALFSAMPLE;
weight SERIESWGT3; * Series-adjusted victimization weight;
domain SUB('1')*GENDER;
var VIOLENT;
run;
```

## 4.5    Victimization Proportions

Victimization proportions estimate the proportion of victimizations or victims with a particular characteristic (e.g., proportion of victimizations reported to police or proportion of victimizations involving a weapon). The process for calculating victimization proportions is very similar to the process outlined for victimization totals in Section 4.4. Calculating victimization proportions involves the following steps. A graphical representation of this process is also provided in *Figure B.3* of Appendix B.

1. Merge the design variables and state indicators from the household-level file onto the incident-level file. Additional variables (e.g., age, sex) from the household- or person-level file may need to be merged depending on the analysis.
2. Using the household-level file, create a file with a dummy record for each PSU.
3. Append the file from Step 2 with the incident-level file from Step 1.

4. Create an indicator for the victimization characteristic(s) of interest (e.g., violent crime). A domain indicator should also be created to exclude dummy records and crimes occurring outside the United States. Additional information can also be included in the domain indicator (e.g., crime type) or can be added to the domain statement in the SAS SURVEYMEANS procedure.
5. Calculate the victimization proportion and TSL standard error.

### 4.5.1 *Example SAS Code*

The following example will illustrate the process for calculating the proportion of violent victimizations reported to the police in California from 2017 to 2019 using SAS.

**Step 1: Merge the Design Variables and State Indicators to Incident-Level File**

```
data ex3_hhld;
 set hhld1719; /* Concatenated household-level file */
 * Calculate a numeric PSEUDOSTRATA;
 PSEUDOSTRATA=UCF_PSEUDOSTR*1;
 * Calculate a numeric HALFSAMPLE code;
 HALFSAMPLE=UCF_HALFSAMPCD*1;
 * Calculate a numeric state indicator: STATE is a character variable indicating the state FIPS code;
 STATENUM=STATE*1;
 keep YEARQUARTER CTRLNUM STATENUM PSEUDOSTRATA HALFSAMPLE;
run;

proc sort data=ex3_hhld; by YEARQUARTER CTRLNUM; run;

proc sort data=incident1719; by YEARQUARTER CTRLNUM; run;

data ex3_incident;
merge incident1719(in=inINC) /* Concatenated incident file */
    ex3_hhld;
by YEARQUARTER CTRLNUM;
if inINC;
run;
```

**Step 2: Create a Dummy Record for Each PSU**

```
proc freq data=ex3_hhld noprint;
 tables pseudostrata*halfsample/out=ex3_psu;
run;

data ex3_dummy;
 set ex3_psu;
 * Assign a positive weight to each dummy record;
 SERIESWGT=1;
 * Create a dummy indicator for each record;
 DUMMY=1;
run;
```

**Step 3: Append the Dummy File to the Incident-Level File**

```
data ex3_incident2;
 set ex3_psu
    ex3_incident(in=inINC); * Concatenated incident file;
 if inINC then dummy=0;
```

20

```
run;
```

**Step 4: Identify Records with Victimization Characteristic(s) of Interest**

```
data ex3_analysis_file;
 set ex3_incident2;
* Create an indicator of violent crime;
 VIOLENT=(1 <= TOCNEW <= 20);
* Create an indicator of reporting to police: 0=No; 1=Yes;
 REPORT_POLICE=(POLICEINFORMED = 1);
* Create an indicator for crimes that occurred outside the United States;
 EXCLUDE_OUTUS=(INCIDENTPLACE=1);
* Create a domain indicator to exclude crimes occurring outside the United States and dummy records and to subset
to the crime type and state(s) of interest (i.e., violent victimizations in California). STATENUM is a numeric
variable indicating the state FIPS code (see Step 1);
 sub=(EXCLUDE_OUTUS=0 and DUMMY=0 and STATENUM=6 and VIOLENT=1);
run;
```

**Step 5: Calculate the Victimization Proportion and TSL Standard Errors**

```
proc surveymeans data=ex3_analysis_file varmethod=taylor mean stderr;
strata PSEUDOSTRATA;
cluster HALFSAMPLE;
weight SERIESWGT; * Series-adjusted victimization weight;
domain SUB('1');
var REPORT_POLICE;
run;
```

# 5. SIGNIFICANCE TESTING

## 5.1    Comparisons Between Subgroups

The process of performing comparisons between subgroups is very similar to the process outlined in Section 4 for generating estimates except that the "diffmeans" option will be used on the domain statement in the SAS SURVEYMEANS procedure. The following example demonstrates the process of comparing the violent victimization rate between males and females in California from 2017 to 2019. A graphical representation of this process is also provided in *Figure B.4* of Appendix B.

**Step 1: Identify Records with Victimization Characteristic(s) of Interest**

```
data ex4_incident;
 set incident1719; *Concatenated incident file;
* Create an indicator of violent crime: 0=Not a violent crime; 1=violent crime;
 VIOLENT=(1 <= TOCNEW <= 20);
* Create an indicator for crimes that occurred outside the U.S.: 1=crime occurred outside the United States;
 EXCLUDE_OUTUS=(INCIDENTPLACE='1');
* Create a series crime indicator: 0=Not a series crime; 1=Series crime *;
 if (SIXORMOREINCIDENTS in ('1','8')) or (INCIDENTSSIMILAR in ('2','8')) or (RECALLDETAILS in ('1','8'))
then series=0;
 else series=1;
* Calculate the number of incidents for series crimes;
 SERIESWEIGHT=SERIESWGT/VWGT;
run;
```

**Step 2: Create a Victimization Summary File**

```
proc sort data=ex4_incident; by YEARQUARTER CTRLNUM LINENUM; run;

proc means data=ex4_incident noprint;
where EXCLUDE_OUTUS=0 and VIOLENT=1; * Exclude crimes occurring outside the United States and subset
file to crime type of interest. This also ensures the appropriate weight (VWGT) is kept on the file if a respondent
reported both property crimes and personal crimes.;
weight SERIESWEIGHT;
id VWGT;
by YEARQUARTER CTRLNUM LINENUM;
var VIOLENT;
output out=ex4_victimization_summary sum=;
run;
```

**Step 3: Merge the Victimization Summary File onto Household File (for Property Crimes) or Person File (for Person-Level Crimes)**

```
proc sort data=person1719 out=ex4_person;
by YEARQUARTER CTRLNUM LINENUM;
run;

data ex4_merged_file;
 merge ex4_person ex4_victimization_summary;
 by YEARQUARTER CTRLNUM LINENUM;
* The incident count variable is missing for persons not included on the victimization summary file, so they are set
to '0' (no victimizations of this type;
 if VIOLENT=. then VIOLENT=0;
run;
```

**Step 3, Part 2: Merge Design Variables and State Indicator onto Person-File**

```
data ex4_hhld;
 set hhld1719; /* Concatenated household-level file */
 * Calculate a numeric PSEUDOSTRATA;
 PSEUDOSTRATA=UCF_PSEUDOSTR*1;
 * Calculate a numeric HALFSAMPLE code;
 HALFSAMPLE=UCF_HALFSAMPCD*1;
 * Calculate a numeric state indicator: STATE is a character variable indicating the state FIPS code;
 STATENUM=STATE*1;
 keep YEARQUARTER CTRLNUM STATENUM PSEUDOSTRATA HALFSAMPLE;
run;

proc sort data=ex4_hhld; by YEARQUARTER CTRLNUM; run;

data ex4_merged_file2;
merge ex4_merged_file(in=in1) ex4_hhld;
by YEARQUARTER CTRLNUM;
if in1;
run;
```

**Step 4: Calculate the Victimization Adjustment Factor**

```
data ex4_analysis_file;
 set ex4_merged_file2;
* Adjust the person weight;
 PERSONWEIGHT2=PERSONWEIGHT/2;
* Calculate the adjustment factor;
 if VWGT > 0 then ADJINC_WT=VWGT/PERSONWEIGHT2;
 else ADJINC_WT=0;
* Create indicator of respondent gender;
if SEXALLOC='1' then gender=1; * male;
else if SEXALLOC='2' then gender=2; * female;
* Create an analysis variable equal to the victimization count multiplied by the adjustment factor multiplied by
1,000 (to express rates per 1,000 persons);
 ANALYSISVAR=VIOLENT*ADJINC_WT*1000;
run;
```

**Step 5: Compare the Victimization Rate by Gender**

```
proc surveymeans data=ex4_analysis_file varmethod=taylor mean stderr;
strata PSEUDOSTRATA;
cluster HALFSAMPLE;
domain statenum('6')*GENDER/diffmeans; * DIFFMEANS compares the groups defined by the domain variables.
STATENUM is a numeric variable indicating the state FIPS code (see Step 3, Part 2);
weight PERSONWEIGHT2; * Adjusted person weight - collection year;
var ANALYSISVAR;
run;
```

## 5.2    Comparisons Between States

The process of performing comparisons between states is very similar to the process outlined in Section 5.1 for comparing subgroups except that the states being compared will replace the victim characteristic (i.e., GENDER in the Section 5.1 example) on the domain statement. The following example demonstrates the process of comparing the violent victimization rate in California with the violent victimization rate in Arizona from 2017 to 2019

in SAS. Only Step 5 is shown below as Steps 1 through 4 are identical to the example in Section 5.1. A graphical representation of this process is also provided in ***Figure B.5*** of Appendix B.

**Step 5: Compare the Victimization Rate Between States**

```
proc surveymeans data=ex4_analysis_file varmethod=taylor mean stderr;
strata PSEUDOSTRATA;
cluster HALFSAMPLE;
domain statenum('6','4')/diffmeans; * DIFFMEANS compares the groups defined by the domain variables.
STATENUM is a numeric variable indicating the state FIPS code.;
weight PERSONWEIGHT2; * Adjusted person weight - collection year;
var ANALYSISVAR;
run;
```

## 5.3    Comparisons Over Time

Making comparisons over time requires additional steps to set up the data file properly for analysis. Because multiple years of data will be included in each time period and these periods will often overlap (e.g., 2017 to 2019 vs. 2018 to 2020), the data must be stacked with an indicator of the time period (see Step 5 below). The following example demonstrates the process for comparing the violent victimization crime rate in California from 2017 to 2019 and from 2018 to 2020 in SAS. Analysts should note that because of the overlapping time periods, and because of the longitudinal nature of the NCVS, estimates from different time periods may be correlated. A graphical representation of this process is also provided in ***Figure B.6*** of Appendix B.

**Step 1: Identify Records with Victimization Characteristic(s) of Interest Across All Years**

```
data ex6_incident;
 set incident1720; *Concatenated incident file, 2017 - 2020;
* Create an indicator of violent crime: 0=Not a violent crime; 1=violent crime;
 VIOLENT=(1 <= TOCNEW <= 20);
* Create an indicator for crimes that occurred outside the United States: 1=crime occurred outside the United States;
 EXCLUDE_OUTUS=(INCIDENTPLACE='1');
* Create a series crime indicator: 0=Not a series crime; 1=Series crime *;
 if (SIXORMOREINCIDENTS in ('1','8')) or (INCIDENTSSIMILAR in ('2','8')) or (RECALLDETAILS in ('1','8'))
then series=0;
 else series=1;
* Calculate the number of incidents for series crimes;
 SERIESWEIGHT=SERIESWGT/VWGT;
run;
```

**Step 2: Create a Victimization Summary File**

```
proc sort data=ex6_incident; by YEARQUARTER CTRLNUM LINENUM; run;

proc means data=ex6_incident noprint;
where EXCLUDE_OUTUS=0 and VIOLENT=1; * Exclude crimes occurring outside United States and subset file
to crime type of interest. This also ensures the appropriate weight (VWGT) is kept on the file if a respondent
reported both property crimes and personal crimes;
weight SERIESWEIGHT;
id VWGT;
by YEARQUARTER CTRLNUM LINENUM;
var VIOLENT;
```

```sas
 output out=ex6_victimization_summary sum=;
run;
```

**Step 3: Merge Victimization Summary File onto Household File (for Property Crimes) or Person File (for Person-Level Crimes)**

```sas
proc sort data=person1720 out=ex6_person; *Concatenated person file, 2017 - 2020;
by YEARQUARTER CTRLNUM LINENUM;
run;

data ex6_merged_file;
 merge ex6_person ex6_victimization_summary;
 by YEARQUARTER CTRLNUM LINENUM;
* The incident count variable is missing for persons not included on the victimization summary file, so they are set
to '0' (no victimizations of this type;
 if VIOLENT=. then VIOLENT=0;
run;
```

**Step 3, Part 2: Merge Design Variables and State Indicator onto Person-File**

```sas
data ex6_hhld;
 set hhld1720; /* Concatenated household-level file, 2017 - 2020 */
* Calculate a numeric PSEUDOSTRATA;
PSEUDOSTRATA=UCF_PSEUDOSTR*1;
* Calculate a numeric HALFSAMPLE code;
HALFSAMPLE=UCF_HALFSAMPCD*1;
* Calculate a numeric state indicator: STATE is a character variable indicating the state FIPS code;
STATENUM=STATE*1;
 keep YEARQUARTER CTRLNUM STATENUM PSEUDOSTRATA HALFSAMPLE;
run;

proc sort data=ex6_hhld; by YEARQUARTER CTRLNUM; run;

data ex6_merged_file2;
 merge ex6_merged_file(in=in1) ex4_hhld;
 by YEARQUARTER CTRLNUM;
 if in1;
run;
```

**Step 4: Calculate the Victimization Adjustment Factor**

```sas
data ex6_merged_file3;
 set ex6_merged_file2;
* Adjust the person weight;
 PERSONWEIGHT2=PERSONWEIGHT/2;
* Calculate the adjustment factor;
 if VWGT > 0 then ADJINC_WT=VWGT/PERSONWEIGHT2;
 else ADJINC_WT=0;
* Create indicator of respondent gender;
if SEXALLOC='1' then gender=1; * male;
else if SEXALLOC='2' then gender=2; * female;
* Create an analysis variable equal to the victimization count multiplied by the adjustment factor multiplied by
1,000 (to express rates per 1,000 persons);
 ANALYSISVAR=VIOLENT*ADJINC_WT*1000;
run;
```

**Step 5: Stack the Data for the Year Groups Being Compared**

```
data ex6_temp1 ex6_temp2;
 set ex6_merged_file3;
* Create a separate file with each year grouping;
 if YEAR in (2017,2018,2019) then output ex6_temp1;
 if YEAR in (2018,2019,2020) then output ex6_temp2;
run;


data ex6_analysis_file;
 set ex6_temp1(in=in1)
    ex6_temp2(in=in2);
* Create an indicator of the year grouping;
 if in1 then YEARGRP=1;
 if in2 then YEARGRP=2;
run;
```

**Step 6: Compare the Time Periods of Interest**

```
proc surveymeans data=ex6_analysis_file varmethod=taylor mean stderr;
strata PSEUDOSTRATA;
cluster HALFSAMPLE;
domain statenum('6')*YEARGRP/diffmeans; * DIFFMEANS compares the groups defined by the domain variables.
STATENUM is a numeric variable indicating the state FIPS code (see Step 3, Part 2);
weight PERSONWEIGHT2; * Adjusted person weight - collection year;
var ANALYSISVAR;
run;
```

# REFERENCES

Biemer, P. P. (2010). Total survey error: Design, implementation, and evaluation. *Public Opinion Quarterly, 74*(5), 817-848.

Brick, J. M., Lohr, S., Giambo, P., Broene, P., Edwards, W. S., Jones, R., & Lim, Y. (2020). *National Crime Victimization Survey local-area crime survey: Field test methodology report (NCJ 254519)*. Office of Justice Programs, U.S. Justice Department. Retrieved from https://www.ojp.gov/library/publications/national-crime-victimization-survey-local-area-crime-survey-field-test

Bureau of Justice Statistics, U.S. Department of Justice. (2015). *Assessing the coverage and reliability of subnational geographic identifiers in the NCVS public-use file*. Retrieved from https://www.ojp.gov/pdffiles1/bjs/grants/249467.pdf

Bureau of Justice Statistics, U.S. Department of Justice. (2021). *National Crime Victimization Survey: assessment of outlier weights*. Retrieved from https://www.ojp.gov/pdffiles1/bjs/grants/302186.pdf

Bureau of Justice Statistics, U.S. Department of Justice. (2022). *National Crime Victimization Survey: MSA public-use data, 2000–2015*. Inter-university Consortium for Political and Social Research. Retrieved from https://doi.org/10.3886/ICPSR38321.v1

Fay, R. E. (2021). *Constructing and disseminating small area estimates from the National Crime Victimization Survey, 2007–2018 (NCJ 300603)*. Office of Justice Programs, U.S. Department of Justice. Retrieved from https://www.ojp.gov/library/publications/constructing-and-disseminating-small-area-estimates-national-crime

Kena, G. & Morgan, R. E. (2023). *Criminal victimization in the 22 largest states, 2017–2019 (NCJ 305402)*. Bureau of Justice Statistics, U.S. Department of Justice. Retrieved from https://bjs.ojp.gov/library/publications/criminal-victimization-22-largest-us-states-2017-2019

Lauritsen, J. L. (2022). *Victimization in different types of areas in the United States: Subnational findings from the National Crime Victimization Survey, 2010–2015 (NCJ 252630)*. Office of Justice Programs, U.S. Department of Justice. Retrieved from https://www.ojp.gov/library/publications/constructing-and-disseminating-small-area-estimates-national-crime

Moore, A., Zimmer, S., Berzofsky, M., Trotta, M., & Krebs, C. (2023). *National Crime Victimization Survey: Validation of state-level estimates (NCJ 305403)*. RTI International. Retrieved from https://www.ojp.gov/pdffiles1/bjs/grants/305403.pdf

Moore, A., Couzens, G. L., & Berzofsky, M. (2021). *Estimating crime victimization in large states and MSAs through reweighting: Evaluation and methodology (NCJ 252701)*. Bureau of Justice Statistics, U.S. Department of Justice. Retrieved from

https://bjs.ojp.gov/library/publications/estimating-crime-victimization-large-states-and-msas-through-reweighting

Morgan, R. E. & Kena, G. (2018). *Criminal victimization, 2016: Revised (NCJ 252121)*. Bureau of Justice Statistics, U.S. Department of Justice. Retrieved from https://www.ojp.gov/ncjrs/virtual-library/abstracts/criminal-victimization-2016

Shook-Sa, B., Couzens, G., & Berzofsky, M. (2015). *User's guide to National Crime Victimization Survey (NCVS) direct variance estimation*. Bureau of Justice Statistics, U.S. Department of Justice. Retrieved from https://bjs.ojp.gov/sites/g/files/xyckuh236/files/media/document/ncvs_variance_user_guide_11.06.14.pdf

U.S. Census Bureau. (2020). *The researcher handbook: Federal Statistical Research Data Center's handbook for researchers*. Retrieved from https://trdc.cpc.unc.edu/wp-content/uploads/2021/08/Researcher_Handbook_1208020.pdf

**APPENDIX A: SUPPLEMENTAL TABLES**

**Table A.1:   Estimated person-level coverage of NCVS sample by state and characteristic, 2017–2019**

| Characteristic | Coverage by State: −3 = Less than 0.5; −2 = 0.5 to less than 0.75; −1 = 0.75 to less than 0.90; 0 = 0.90 to 1.1; 1 = Greater than 1.1 to 1.25; 2 = Greater than 1.25 to 1.5; 3 = Greater than 1.5 | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
|  | AZ | CA | CO | FL | GA | IL | IN | MD | MA | MI | MN | MO | NJ | NY | NC | OH | PA | TN | TX | VA | WA | WI |
| All persons age 12 or older | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Gender | | | | | | | | | | | | | | | | | | | | | | |
|   Male | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|   Female | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Age | | | | | | | | | | | | | | | | | | | | | | |
|   12–15 | −1 | 0 | 0 | 0 | −1 | 0 | −1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|   16–19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | −1 | 0 | 0 | 0 | 0 | 0 |
|   20–24 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
|   25–34 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|   35–49 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|   50–64 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|   65 or older | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Race/ethnicity | | | | | | | | | | | | | | | | | | | | | | |
|   White | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|   Black | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|   Hispanic | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|   Other/more than one race | −1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | −1 | −1 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| Household tenure | | | | | | | | | | | | | | | | | | | | | | |
|   Own | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|   Rent/no cash rent | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| Educational attainment | | | | | | | | | | | | | | | | | | | | | | |
|   Less than high school | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
|   High school | 0 | 0 | 0 | 0 | 0 | −1 | 0 | −1 | −1 | 0 | −1 | 0 | 0 | 0 | −1 | 0 | −1 | 0 | −1 | 0 | 0 | 0 |
|   Some college/associate's degree | 0 | −1 | 0 | 0 | −1 | 0 | 0 | −1 | −1 | 0 | 0 | 0 | 0 | −1 | −1 | −1 | −1 | 0 | 0 | −1 | 0 | 0 |
|   College degree or greater | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| Marital status | | | | | | | | | | | | | | | | | | | | | | |
|   Married | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|   Never married | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|   Previously married | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Employed last week | | | | | | | | | | | | | | | | | | | | | | |
|   Employed | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|   Unemployed | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
|   Minor (age < 18) | 0 | 0 | 0 | 0 | 0 | 0 | −1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

*Notes.* The following scale was used to summarize the coverage ratios: −3 = Less than 0.5; −2 = 0.5 to less than 0.75; −1 = 0.75 to less than 0.90; 0 = 0.90 to 1.1; 1 = Greater than 1.1 to 1.25; 2 = Greater than 1.25 to 1.5; 3 = Greater than 1.5.

A coverage ratio less than 1 indicates that the NCVS total undercovers the target population relative to the American Community Survey population totals. A coverage ratio greater than 1 indicates that the NCVS total overcovers the population relative to the American Community Survey population totals.

**Table A.2: Estimated household-level coverage of NCVS sample by state and characteristic, 2017–2019**

| Characteristic | Coverage by State: −3 = Less than 0.5; −2 = 0.5 to less than 0.75; −1 = 0.75 to less than 0.90; 0 = 0.90 to 1.1; 1 = Greater than 1.1 to 1.25; 2 = Greater than 1.25 to 1.5; 3 = Greater than 1.5 | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | AZ | CA | CO | FL | GA | IL | IN | MD | MA | MI | MN | MO | NJ | NY | NC | OH | PA | TN | TX | VA | WA | WI |
| All households | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Age of householder | | | | | | | | | | | | | | | | | | | | | | |
|   Younger than 30 | 2 | 3 | 2 | 2 | 3 | 2 | 2 | 2 | 3 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 3 | 1 | 2 |
|   30–34 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
|   35–49 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|   50–64 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | −1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | −1 | 0 | 0 | 0 | 0 | 0 |
|   65 or older | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Race/ethnicity of householder | | | | | | | | | | | | | | | | | | | | | | |
|   White | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|   Black | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 |
|   Hispanic | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
|   Other/more than one race | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 2 | 2 | 0 | 1 | 0 | 0 |
| Household tenure | | | | | | | | | | | | | | | | | | | | | | |
|   Own | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|   Rent/no cash rent | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Educational attainment of householder | | | | | | | | | | | | | | | | | | | | | | |
|   Less than high school | 0 | 0 | 2 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 3 | 0 | 0 | 1 | 2 | 0 | 1 | 1 | 2 | 0 | 1 |
|   High school | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | −1 | 0 | 0 | 0 | 0 | 0 | −1 | 0 | 0 | 0 | 0 | 0 |
|   Some college/associate's degree | 0 | 0 | 0 | 0 | 0 | 0 | 0 | −1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|   College degree or greater | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Number of housing units in structure | | | | | | | | | | | | | | | | | | | | | | |
|   1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|   2 or more | 0 | 1 | 1 | 1 | 2 | 1 | 1 | 0 | 0 | 1 | 2 | 0 | 0 | 0 | 1 | 0 | 2 | 1 | 1 | 2 | 0 | 1 |
| Number of motor vehicles | | | | | | | | | | | | | | | | | | | | | | |
|   0 | 0 | 0 | 1 | −1 | 0 | −1 | 2 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
|   1 | 0 | 0 | −1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|   2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | −1 |
|   3 or more | 2 | 0 | 2 | 1 | 0 | 1 | 2 | 0 | 0 | 1 | 2 | 1 | 0 | 0 | −1 | 1 | 1 | 1 | 0 | 1 | 2 | 2 |

*Notes.* The following scale was used to summarize the coverage ratios: −3 = Less than 0.5; −2 = 0.5 to less than 0.75; −1 = 0.75 to less than 0.90; 0 = 0.90 to 1.1; 1 = Greater than 1.1 to 1.25; 2 = Greater than 1.25 to 1.5; 3 = Greater than 1.5.

A coverage ratio less than 1 indicates that the NCVS total undercovers the target population relative to the American Community Survey population totals. A coverage ratio greater than 1 indicates that the NCVS total overcovers the population relative to the American Community Survey population totals.

**Table A.3:   Person-level response rates by state and characteristic, 2017–2019**

| Characteristic | Response Rates by State: 1 = Less than 50%; 2 = 50% to less than 60%; 3 = 60% to less than 70%; 4 = 70% to less than 80%; 5 = 80% or greater | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | AZ | CA | CO | FL | GA | IL | IN | MD | MA | MI | MN | MO | NJ | NY | NC | OH | PA | TN | TX | VA | WA | WI |
| All persons age 12 or older | 4 | 5 | 4 | 5 | 5 | 4 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 |
| Gender | | | | | | | | | | | | | | | | | | | | | | |
| Male | 4 | 5 | 4 | 5 | 5 | 4 | 4 | 5 | 5 | 5 | 5 | 5 | 4 | 4 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 |
| Female | 4 | 5 | 4 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| Age | | | | | | | | | | | | | | | | | | | | | | |
| 12–15 | 1 | 3 | 1 | 3 | 3 | 1 | 1 | 2 | 1 | 3 | 2 | 2 | 2 | 2 | 4 | 3 | 1 | 2 | 3 | 3 | 3 | 2 |
| 16–19 | 2 | 3 | 1 | 3 | 3 | 1 | 1 | 2 | 2 | 4 | 2 | 2 | 2 | 2 | 4 | 3 | 1 | 2 | 3 | 2 | 3 | 3 |
| 20–24 | 3 | 3 | 3 | 4 | 4 | 3 | 3 | 4 | 3 | 5 | 4 | 4 | 3 | 3 | 5 | 4 | 3 | 4 | 4 | 3 | 3 | 4 |
| 25–34 | 4 | 4 | 4 | 5 | 5 | 4 | 4 | 4 | 4 | 5 | 5 | 5 | 4 | 4 | 5 | 5 | 4 | 5 | 5 | 5 | 4 | 5 |
| 35–49 | 4 | 5 | 4 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| 50–64 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| 65 or older | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| Race/ethnicity | | | | | | | | | | | | | | | | | | | | | | |
| White | 5 | 5 | 4 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 |
| Black | 4 | 5 | 4 | 5 | 5 | 4 | 4 | 5 | 4 | 5 | 4 | 5 | 4 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 4 | 5 |
| Hispanic | 4 | 4 | 4 | 5 | 5 | 4 | 3 | 4 | 5 | 5 | 4 | 4 | 4 | 5 | 5 | 4 | 5 | 5 | 5 | 4 | 4 | 4 |
| Other/more than one race | 4 | 4 | 4 | 5 | 5 | 4 | 4 | 5 | 4 | 5 | 4 | 5 | 4 | 5 | 5 | 5 | 4 | 5 | 5 | 4 | 4 | 4 |
| Household tenure | | | | | | | | | | | | | | | | | | | | | | |
| Own | 4 | 5 | 4 | 5 | 5 | 4 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 4 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 |
| Rent/no cash rent | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| Educational attainment | | | | | | | | | | | | | | | | | | | | | | |
| Less than high school | 3 | 4 | 2 | 5 | 4 | 3 | 2 | 4 | 3 | 5 | 3 | 4 | 3 | 4 | 5 | 4 | 3 | 4 | 4 | 4 | 4 | 4 |
| High school | 4 | 5 | 4 | 5 | 5 | 4 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| Some college/associate's degree | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| College degree or greater | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| Marital status | | | | | | | | | | | | | | | | | | | | | | |
| Married | 5 | 5 | 4 | 5 | 5 | 5 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| Never married | 3 | 4 | 3 | 5 | 5 | 4 | 3 | 4 | 4 | 5 | 4 | 4 | 4 | 4 | 5 | 4 | 3 | 4 | 4 | 4 | 4 | 4 |
| Previously married | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |

*Notes.* The following scale was used to summarize the response rates: 1 = Less than 50%; 2=50% to less than 60%; 3 = 60% to less than 70%; 4 = 70% to less than 80%; 5 = 80% or greater. Person-level response rates are calculated within responding households.

## Table A.4: Household-level response rates by state, 2017–2019

| Characteristic | Response Rates by State: 1 = Less than 50%; 2 = 50% to less than 60%; 3 = 60% to less than 70%; 4 = 70% to less than 80%; 5 = 80% or greater | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | AZ | CA | CO | FL | GA | IL | IN | MD | MA | MI | MN | MO | NJ | NY | NC | OH | PA | TN | TX | VA | WA | WI |
| All households | 3 | 4 | 3 | 4 | 4 | 4 | 4 | 3 | 3 | 4 | 4 | 4 | 2 | 4 | 5 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |

*Note.* The following scale was used to summarize the response rates: 1 = Less than 50%; 2 = 50% to less than 60%; 3 = 60% to less than 70%; 4 = 70% to less than 80%; 5 = 80% or greater.

## Table A.5: Person-level percent relative bias by characteristic and state, 2017–2019

| Characteristic | Percentage of Relative Bias by State | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | AZ | CA | CO | FL | GA | IL | IN | MD | MA | MI | MN | MO | NJ | NY | NC | OH | PA | TN | TX | VA | WA | WI |
| Gender | | | | | | | | | | | | | | | | | | | | | | |
| Male | −2% | −1% | −2% | −1% | −1% | −1% | −1% | −1% | −1% | −1% | −1% | −2% | −1% | −2% | −1% | −2% | −2% | −1% | −2% | −1% | −1% | −1% |
| Female | 2 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 | 0 | 2 | 2 | 1 | 1 | 1 | 1 | 1 |
| Age | | | | | | | | | | | | | | | | | | | | | | |
| 12–17 | −38 | −23 | −39 | −22 | −23 | −39 | −49 | −30 | −42 | −23 | −32 | −36 | −30 | −38 | −21 | −28 | −45 | −32 | −25 | −27 | −24 | −32 |
| 18–24 | −20 | −14 | −13 | −13 | −11 | −17 | −17 | −16 | −19 | −10 | −15 | −15 | −23 | −17 | −9 | −13 | −21 | −14 | −10 | −14 | −16 | −10 |
| 25–34 | 1 | −1 | 1 | −2 | 2 | 2 | 6 | −3 | −1 | 1 | 1 | 2 | −4 | −2 | 2 | 1 | −1 | 0 | 1 | 0 | −2 | 1 |
| 35–49 | 1 | 3 | 1 | 0 | 2 | 4 | 4 | 3 | 4 | 2 | 2 | 2 | 3 | 3 | 2 | 3 | 5 | 2 | 2 | 1 | 1 | 1 |
| 50–64 | 7 | 4 | 6 | 3 | 4 | 5 | 8 | 5 | 6 | 3 | 4 | 5 | 6 | 6 | 3 | 4 | 6 | 5 | 5 | 4 | 5 | 4 |
| 65 or older | 15 | 12 | 16 | 9 | 8 | 13 | 17 | 12 | 14 | 7 | 14 | 12 | 14 | 15 | 6 | 9 | 17 | 12 | 10 | 12 | 14 | 11 |
| Race/ethnicity | | | | | | | | | | | | | | | | | | | | | | |
| White | 4 | 5 | 2 | 1 | 0 | 1 | 1 | 3 | 1 | 0 | 2 | 1 | 2 | −1 | 1 | 0 | 2 | 1 | 2 | 3 | 3 | 1 |
| Black | 0 | 2 | 2 | −4 | 1 | −1 | 3 | −2 | −5 | 2 | −8 | −1 | −2 | 5 | −1 | −2 | −5 | −3 | 1 | 1 | −7 | −2 |
| Hispanic | −5 | −4 | −5 | 1 | −4 | −4 | −10 | −7 | 0 | −1 | −10 | −8 | −5 | 0 | −5 | 1 | −9 | −4 | −3 | −9 | −3 | −6 |
| Other/more than one race | −7 | −3 | −3 | −5 | −2 | −1 | −2 | −3 | −3 | 2 | −12 | −3 | −2 | 1 | 0 | 1 | −5 | −1 | 1 | −11 | −9 | −9 |

## Table A.6: Household-level percent relative bias by characteristic and state, 2017–2019

| Characteristic | Percentage of Relative Bias by State | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | AZ | CA | CO | FL | GA | IL | IN | MD | MA | MI | MN | MO | NJ | NY | NC | OH | PA | TN | TX | VA | WA | WI |
| MSA status | | | | | | | | | | | | | | | | | | | | | | |
| Central city of an MSA | 0% | 0% | 0% | −1% | 0% | −6% | −5% | −11% | −10% | 1% | −2% | −11% | −5% | 0% | 0% | −4% | −10% | −6% | 0% | 3% | −2% | −5% |
| In MSA but not central city | 2 | 0 | 0 | 0 | −1 | 3 | 3 | 2 | 2 | −1 | 1 | 4 | 1 | 0 | −1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 |
| Not in MSA | −37 | 14 | 1 | 0 | 3 | 3 | 5 | 17 | 22 | 4 | −1 | −1 | 0 | −1 | 3 | 0 | 9 | 5 | −2 | −9 | 5 | 4 |
| Urban/rural status | | | | | | | | | | | | | | | | | | | | | | |
| Urban | 0 | 0 | 0 | 0 | −1 | −1 | −2 | −2 | 0 | 0 | 0 | −1 | 0 | 0 | 0 | −1 | −2 | −1 | 0 | 0 | 0 | −1 |
| Rural | −3 | 3 | −2 | 2 | 2 | 5 | 7 | 11 | 5 | 0 | 0 | 4 | 4 | 1 | 0 | 5 | 9 | 3 | 2 | −2 | 2 | 3 |

**Table A.7: Relative standard error by crime type and state, 2017–2019**

| Crime Type | AZ | CA | CO | FL | GA | IL | IN | MD | MA | MI | MN | MO | NJ | NY | NC | OH | PA | TN | TX | VA | WA | WI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Violent crime** | 14% | 10% | 9% | 12% | 14% | 15% | 12% | 15% | 27% | 14% | 9% | 10% | 21% | 14% | 17% | 11% | 14% | 14% | 10% | 11% | 9% | 21% |
| Rape/sexual assault | 26 | 40 | 35 | 23 | 44 | 43 | 41 | 40 | 43 | 37 | 21 | 44 | 47 | 38 | 50 | 35 | 45 | 32 | 30 | 27 | 27 | 41 |
| Robbery | 29 | 17 | 24 | 27 | 26 | 26 | 37 | 47 | 38 | 38 | 20 | 20 | 58 | 25 | 26 | 31 | 19 | 27 | 23 | 23 | 26 | 36 |
| Assault | 17 | 11 | 11 | 14 | 15 | 17 | 11 | 17 | 31 | 16 | 11 | 10 | 23 | 15 | 16 | 11 | 15 | 15 | 11 | 12 | 11 | 23 |
| Aggravated assault | 23 | 16 | 12 | 22 | 23 | 38 | 20 | 21 | 27 | 23 | 14 | 20 | 53 | 22 | 18 | 24 | 25 | 20 | 14 | 19 | 24 | 30 |
| Simple assault | 19 | 12 | 13 | 18 | 16 | 15 | 12 | 20 | 35 | 18 | 13 | 12 | 24 | 17 | 18 | 13 | 19 | 18 | 14 | 13 | 9 | 26 |
| **Violent crime excluding simple assault** | 15 | 12 | 17 | 16 | 19 | 24 | 21 | 21 | 22 | 14 | 11 | 18 | 32 | 16 | 27 | 18 | 16 | 16 | 13 | 14 | 15 | 22 |
| **Selected characteristics of violent crime** | 30 | 25 | 33 | 22 | 45 | 27 | 17 | 30 | 41 | 37 | 25 | 24 | 48 | 38 | 39 | 25 | 27 | 26 | 18 | 24 | 25 | 19 |
| Domestic violence | 35 | 37 | 39 | 26 | 46 | 32 | 20 | 39 | 43 | 41 | 31 | 30 | 56 | 40 | 61 | 28 | 39 | 39 | 17 | 34 | 39 | 25 |
| Intimate partner violence | 16 | 11 | 17 | 21 | 18 | 22 | 21 | 22 | 15 | 19 | 10 | 17 | 26 | 15 | 22 | 16 | 17 | 25 | 16 | 14 | 14 | 24 |
| Stranger violence | 28 | 19 | 18 | 19 | 32 | 26 | 21 | 25 | 39 | 15 | 18 | 24 | 39 | 21 | 24 | 25 | 19 | 20 | 16 | 20 | 22 | 26 |
| Violent crime involving injury | 20 | 13 | 19 | 21 | 21 | 31 | 17 | 27 | 28 | 23 | 12 | 18 | 46 | 22 | 19 | 21 | 22 | 18 | 15 | 19 | 22 | 27 |
| Violent crime involving a weapon | 14 | 10 | 9 | 12 | 14 | 15 | 12 | 15 | 27 | 14 | 9 | 10 | 21 | 14 | 17 | 11 | 14 | 14 | 10 | 11 | 9 | 21 |
| **Property crime** | 5 | 5 | 4 | 7 | 7 | 8 | 5 | 6 | 7 | 7 | 7 | 10 | 7 | 7 | 6 | 6 | 6 | 6 | 4 | 7 | 5 | 7 |
| Burglary | 10 | 9 | 11 | 21 | 17 | 13 | 9 | 20 | 23 | 13 | 13 | 20 | 21 | 14 | 11 | 16 | 20 | 11 | 11 | 18 | 12 | 16 |
| Motor vehicle theft | 22 | 10 | 20 | 23 | 16 | 26 | 19 | 29 | 35 | 26 | 22 | 22 | 31 | 29 | 23 | 24 | 21 | 21 | 15 | 24 | 13 | 20 |
| Other theft | 6 | 6 | 4 | 7 | 7 | 8 | 5 | 6 | 8 | 8 | 8 | 9 | 9 | 7 | 7 | 6 | 6 | 7 | 5 | 7 | 4 | 7 |

*Note.* The relative standard error is calculated as the ratio of the standard error and the victimization rate multiplied by 100.

**Table A.8:   Relative standard error of violent crime rate by victim characteristic and state, 2017–2019**

| Characteristic | Relative Standard Error by State | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | AZ | CA | CO | FL | GA | IL | IN | MD | MA | MI | MN | MO | NJ | NY | NC | OH | PA | TN | TX | VA | WA | WI |
| All persons age 12 or older | 14% | 10% | 9% | 12% | 14% | 15% | 12% | 15% | 27% | 14% | 9% | 10% | 21% | 14% | 17% | 11% | 14% | 14% | 10% | 11% | 9% | 21% |
| Gender | | | | | | | | | | | | | | | | | | | | | | |
| Male | 21 | 13 | 17 | 18 | 16 | 22 | 17 | 14 | 20 | 15 | 12 | 15 | 27 | 17 | 20 | 16 | 21 | 21 | 13 | 17 | 13 | 24 |
| Female | 18 | 13 | 13 | 16 | 20 | 20 | 15 | 23 | 37 | 19 | 15 | 15 | 27 | 20 | 23 | 14 | 15 | 18 | 14 | 15 | 11 | 20 |
| Age | | | | | | | | | | | | | | | | | | | | | | |
| 12–17 | 42 | 30 | 15 | 42 | 28 | 37 | 29 | 34 | 43 | 24 | 26 | 29 | 79 | 30 | 44 | 33 | 46 | 36 | 23 | 36 | 31 | 34 |
| 18–24 | 39 | 21 | 41 | 29 | 28 | 35 | 34 | 36 | 31 | 34 | 20 | 30 | 49 | 28 | 33 | 21 | 20 | 31 | 23 | 19 | 31 | 25 |
| 25–34 | 28 | 14 | 16 | 23 | 24 | 24 | 24 | 23 | 32 | 29 | 18 | 29 | 38 | 34 | 34 | 22 | 21 | 23 | 26 | 29 | 14 | 38 |
| 35–49 | 22 | 15 | 11 | 24 | 38 | 38 | 20 | 25 | 68 | 27 | 25 | 22 | 29 | 27 | 30 | 19 | 23 | 32 | 16 | 17 | 20 | 29 |
| 50–64 | 23 | 16 | 18 | 24 | 25 | 27 | 25 | 30 | 59 | 16 | 18 | 22 | 50 | 31 | 24 | 22 | 22 | 27 | 18 | 30 | 19 | 36 |
| 65 or older | 30 | 24 | 33 | 30 | 27 | 36 | 26 | 29 | 37 | 41 | 34 | 28 | 42 | 35 | 33 | 28 | 22 | 24 | 22 | 57 | 20 | 36 |
| Race/ethnicity | | | | | | | | | | | | | | | | | | | | | | |
| White | 19 | 14 | 15 | 16 | 19 | 21 | 14 | 23 | 30 | 14 | 10 | 13 | 26 | 20 | 22 | 13 | 13 | 19 | 19 | 15 | 12 | 25 |
| Black | 41 | 31 | 48 | 21 | 22 | 36 | 22 | 21 | 40 | 23 | 33 | 29 | 52 | 26 | 37 | 25 | 43 | 19 | 21 | 28 | 68 | 35 |
| Hispanic | 24 | 13 | 13 | 30 | 44 | 31 | 27 | 30 | 43 | 28 | 38 | 68 | 54 | 19 | 27 | 38 | 30 | 29 | 13 | 28 | 20 | 50 |
| Other/more than one race | 33 | 24 | 42 | 36 | 75 | 49 | 38 | 32 | 88 | 70 | 36 | 24 | 48 | 43 | 41 | 35 | 39 | 31 | 44 | 34 | 31 | 48 |
| Educational attainment | | | | | | | | | | | | | | | | | | | | | | |
| Less than high school | 32 | 24 | 14 | 28 | 32 | 21 | 22 | 27 | 28 | 19 | 21 | 21 | 42 | 24 | 26 | 22 | 34 | 23 | 17 | 24 | 25 | 31 |
| High school | 20 | 16 | 28 | 19 | 19 | 39 | 24 | 33 | 38 | 35 | 18 | 23 | 40 | 24 | 28 | 17 | 21 | 19 | 20 | 31 | 23 | 19 |
| Some college/ associate's degree | 17 | 16 | 26 | 16 | 24 | 24 | 24 | 20 | 30 | 16 | 19 | 19 | 38 | 23 | 32 | 18 | 16 | 29 | 16 | 16 | 20 | 34 |
| College degree or greater | 22 | 14 | 21 | 21 | 16 | 17 | 23 | 16 | 43 | 26 | 13 | 15 | 35 | 26 | 31 | 25 | 21 | 27 | 28 | 17 | 16 | 28 |
| Marital status | | | | | | | | | | | | | | | | | | | | | | |
| Married | 22 | 16 | 15 | 26 | 14 | 31 | 22 | 18 | 62 | 32 | 15 | 14 | 30 | 26 | 25 | 23 | 23 | 29 | 16 | 16 | 18 | 42 |
| Never married | 21 | 13 | 17 | 17 | 17 | 18 | 20 | 19 | 31 | 13 | 12 | 17 | 26 | 16 | 21 | 13 | 21 | 19 | 15 | 16 | 16 | 21 |
| Previously married | 23 | 15 | 15 | 20 | 24 | 31 | 20 | 22 | 32 | 24 | 24 | 20 | 41 | 25 | 28 | 18 | 17 | 18 | 15 | 18 | 20 | 30 |
| Household income | | | | | | | | | | | | | | | | | | | | | | |
| Less than $25,000 | 19 | 18 | 10 | 17 | 22 | 25 | 24 | 28 | 22 | 29 | 17 | 17 | 47 | 18 | 28 | 20 | 25 | 12 | 14 | 24 | 18 | 28 |
| $25,000–$49,999 | 25 | 16 | 13 | 20 | 27 | 26 | 24 | 23 | 17 | 12 | 22 | 21 | 39 | 28 | 35 | 27 | 16 | 22 | 13 | 20 | 15 | 18 |
| $50,000–$74,999 | 28 | 19 | 21 | 34 | 30 | 25 | 30 | 32 | 35 | 24 | 23 | 19 | 53 | 35 | 33 | 30 | 26 | 40 | 20 | 22 | 27 | 42 |
| $75,000 or more | 29 | 16 | 18 | 21 | 20 | 28 | 11 | 18 | 45 | 27 | 14 | 25 | 27 | 24 | 22 | 19 | 21 | 31 | 27 | 19 | 13 | 29 |

*Note.* The relative standard error is calculated as the ratio of the standard error and the victimization rate multiplied by 100.

**Table A.9:   Relative standard error of property crime rate by household characteristic and state, 2017–2019**

| Characteristic | Relative Standard Error by State | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | AZ | CA | CO | FL | GA | IL | IN | MD | MA | MI | MN | MO | NJ | NY | NC | OH | PA | TN | TX | VA | WA | WI |
| All households | 5% | 5% | 4% | 7% | 7% | 8% | 5% | 6% | 7% | 7% | 7% | 10% | 7% | 7% | 6% | 6% | 6% | 6% | 4% | 7% | 5% | 7% |
| Household tenure | | | | | | | | | | | | | | | | | | | | | | |
| Own | 6 | 5 | 6 | 7 | 8 | 9 | 6 | 8 | 9 | 7 | 9 | 11 | 9 | 9 | 7 | 10 | 8 | 7 | 5 | 9 | 6 | 7 |
| Rent/no cash rent | 9 | 6 | 5 | 9 | 11 | 10 | 8 | 9 | 9 | 9 | 9 | 10 | 12 | 12 | 9 | 7 | 10 | 8 | 7 | 9 | 6 | 9 |
| Household income | | | | | | | | | | | | | | | | | | | | | | |
| Less than $25,000 | 11 | 8 | 12 | 9 | 10 | 15 | 8 | 12 | 11 | 13 | 13 | 13 | 18 | 11 | 12 | 9 | 14 | 8 | 9 | 15 | 17 | 14 |
| $25,000–$49,999 | 9 | 12 | 9 | 13 | 14 | 12 | 8 | 14 | 16 | 13 | 15 | 13 | 24 | 15 | 10 | 10 | 8 | 9 | 7 | 10 | 7 | 11 |
| $50,000–$74,999 | 9 | 8 | 13 | 13 | 13 | 13 | 11 | 13 | 18 | 10 | 10 | 11 | 14 | 19 | 13 | 11 | 10 | 12 | 8 | 14 | 9 | 9 |
| $75,000 or more | 10 | 5 | 6 | 10 | 9 | 10 | 9 | 9 | 10 | 9 | 7 | 10 | 10 | 12 | 10 | 10 | 8 | 11 | 6 | 10 | 6 | 11 |

*Note.* The relative standard error is calculated as the ratio of the standard error and the victimization rate multiplied by 100.

# APPENDIX B: SUPPLEMENTAL FIGURES

## Figure B.1: Process for calculating victimization rates
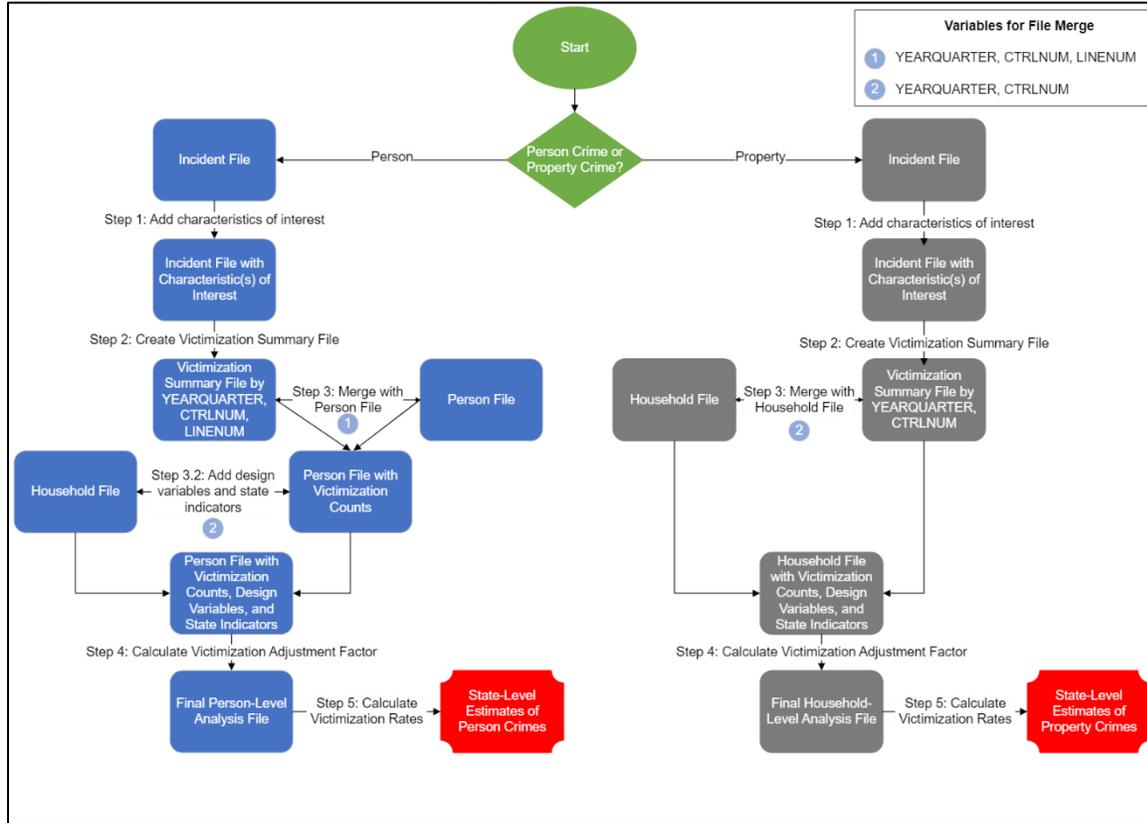
**Figure B.2:  Process for calculating victimization totals**
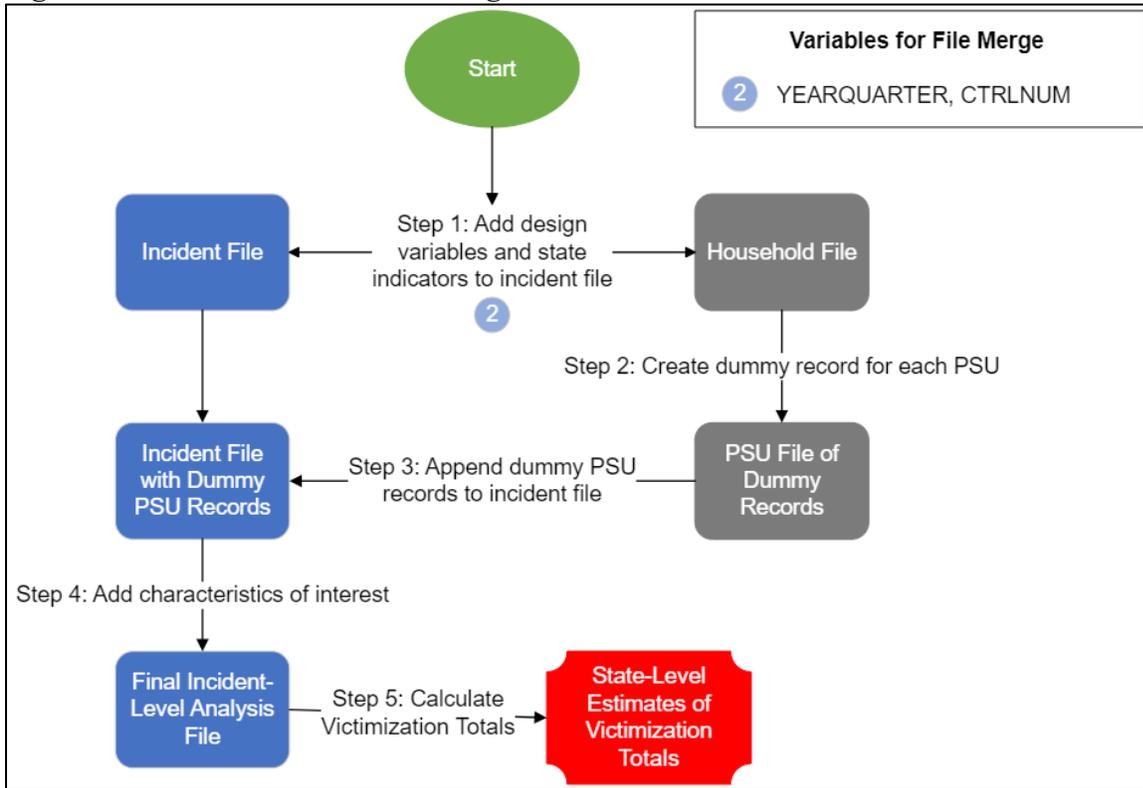


**Figure B-3:  Process for calculating victimization proportions**
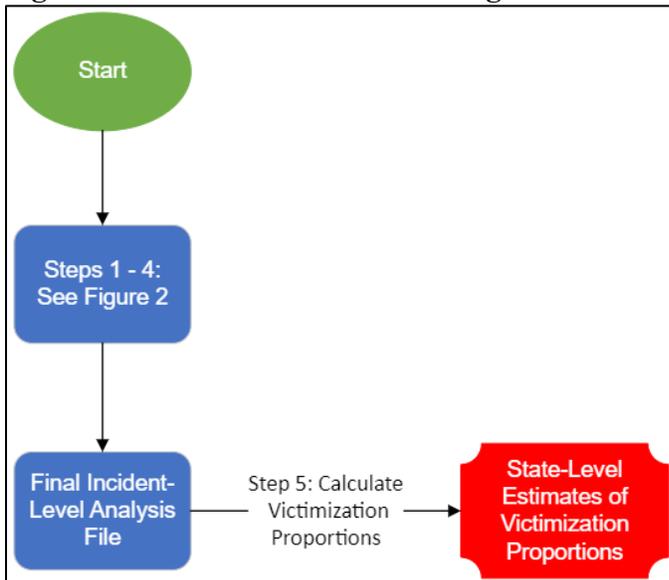
**Figure B-4:    Process for comparing victimization rates for subgroups within states**
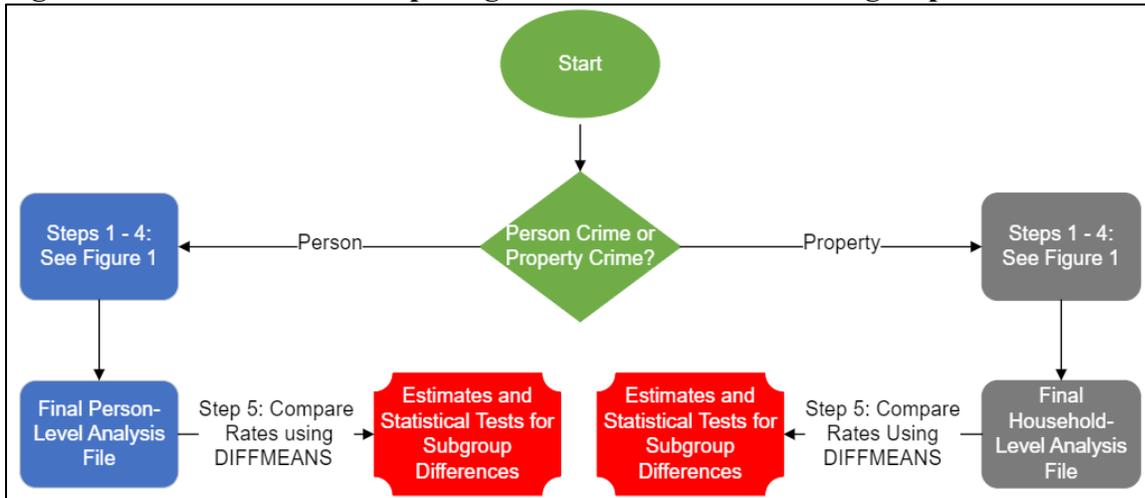


**Figure B-5:    Process for comparing victimization rates across states**
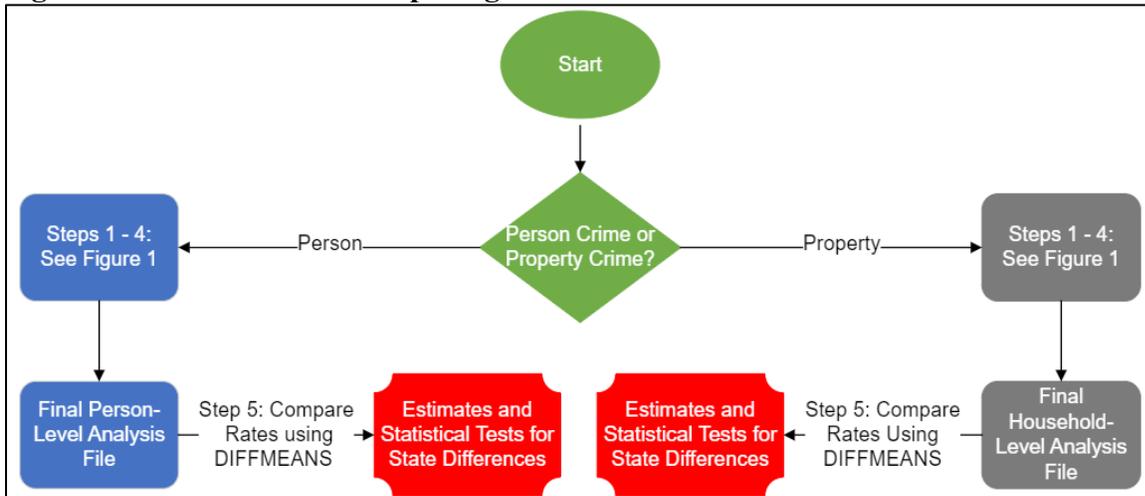
**Figure B.6: Process for comparing victimization rates within a state over time**
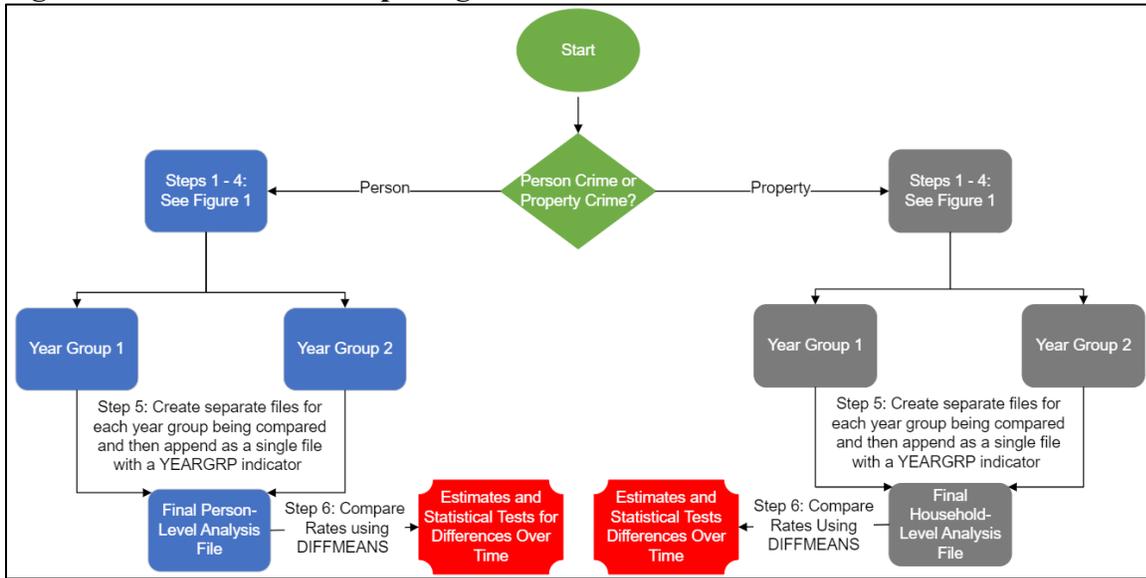
**APPENDIX C: OVERVIEW OF NCVS SUBNATIONAL ESTIMATION PROGRAM**

The National Crime Victimization Survey (NCVS), sponsored by the Bureau of Justice Statistics (BJS), estimates the incidence and describes the characteristics of criminal victimization in the United States. Since its inception in the early 1970s, the NCVS has been a rich source of national-level information about criminal victimization, including victimizations that go unreported to the police. Before 2016, the primary objective of the NCVS was to provide national estimates of criminal victimization in the United States. Under this directive, the sampling and weighting procedures were designed to ensure that the NCVS sample of households and persons was nationally representative. This national focus has historically limited the ability to leverage fully the abundance of information collected through the NCVS to understand the patterns and correlates of crime victimization within subnational areas. The purpose of this guide is to provide a brief overview of BJS's subnational estimation program and to aid data users with producing valid and reliable state-level victimization estimates with NCVS data.

BJS has evaluated several approaches for producing subnational victimization estimates to fill this gap.[11] These methods include:

1. Direct estimates in subnational areas using reweighting methodologies and NCVS data collected under the national design (Moore, Couzens, & Berzofsky, 2021).
2. Model-based estimates that leverage auxiliary data from external sources such as the FBI's Uniform Crime Reporting program and NCVS data from different time periods to effectively increase the sample size in an effort to represent small areas (Fay, 2021).
3. Direct estimates using generic area typologies based on various geographic, social, economic, or demographic characteristics (Lauritsen, 2022).
4. Administration of the Local Area Crime Survey to obtain estimates of crime victimization within states, metropolitan statistical areas, cities, or police jurisdictions (Brick et al., 2020).
5. Boosting the NCVS sample in large states to produce direct state-level estimates of criminal victimization (e.g., *Criminal Victimization in the 22 Largest States, 2017–2019*, NCJ 305402).

---

[11] For more information on the NCVS subnational estimation program, see https://bjs.ojp.gov/subnational-estimates-program.

# APPENDIX D: CROSSWALK BETWEEN RESTRICTED-USE FILES AND PUBLIC-USE FILES FOR SELECT VARIABLES

**Table D.1: Crosswalk between restricted-use files and public-use files for select variables: Household-level file**

| RDC Variable Name | Description | Analogous PUF Variable | Notes |
|---|---|---|---|
| YEARQUARTER | Year and quarter of interview | YEARQ | YEARQUARTER and CTRLNUM uniquely identify a household interview. |
| CTRLNUM | Control number | Not available on PUF | |
| UCF_PSEUDOSTR | Pseudo-stratum Code | V2117 | Used to denote the sample design for TSL variance estimation. |
| UCF_HALFSAMPCD | Half-sample code | V2118 | |
| STATE | State FIPS code | Not available on PUF | Used to identify respondent's state of residence. |
| HHWEIGHT | Household weight | V2116 | HHWEIGHT must be divided by 2 to equal the adjusted household weight (WGTHHCY on NCVS PUFs) used to determine the total number of households during the collection year. |
| TENUREALLOC | Household tenure | V2015 | N/A |
| NUMBEROFUNITS | Number of housing units in structure | V2024 | N/A |
| HHINCOMEA | Household income | SC214A | N/A |
| PPIAGEALLOC | Principal person age | V2033 | Demographics of principal person. The principal person's demographic characteristics are used to construct the household weight. |
| PPIMARITAL | Principal person marital status | V2034 | |
| PPISEXALLOC | Principal person sex | V2036 | |
| PPIEDUCATION | Principal person educational attainment | V2038 | |
| RACEPPRECODE | Principal person race | V2040A | |
| PPIHISPANIC | Principal person Hispanic origin | V2041 | |
| NUMBEROFADULTS | Number of household members age 12 or older | V2071 | These variables can be combined to calculate the total household size. |
| NUMBEROFCHILDREN | Number of household members younger than 12 | V2072 | |
| FAMILYSTRUCTURE | Family structure code | V2122 | N/A |
| INTERVIEWCODE | Type interview code | V2123 | Denotes whether the household interview was completed or the reason for nonresponse. |
| UCF_CBSAMSASTAT | CBSA MSA status | V2129 | RDC variable name may change as sample design changes |

**Table D.2:   Crosswalk between restricted-use files and public-use files for select variables: Person-level file**

| RDC Variable Name | Description | Analogous PUF Variable | Notes |
|---|---|---|---|
| YEARQUARTER | Year and quarter of interview | YEARQ | YEARQUARTER, CTRLNUM, and LINENUM uniquely identify a person interview. |
| CTRLNUM | Control number | Not available on PUF | |
| LINENUM | Person line number | V3010 | |
| PERSONWEIGHT | Person weight | V3080 | PERSONWEIGHT must be divided by 2 to equal the adjusted person weight (WGTPERCY on NCVS PUFs) used to determine the total number of persons age 12 or older during the collection year. |
| TYPEINTERVIEW | Type of interview | V3011 | Demographic characteristics of NCVS sample members. |
| AGEALLOC | Age | V3014 | |
| MARITALSTATUS | Marital status | V3015 | |
| SEXALLOC | Sex | V3018 | |
| EDUCATION | Educational attainment | V3020 | |
| RACEARECODE | Race | V3023A | |
| HISPANICA | Hispanic origin | V3024A | |
| JOBLASTWEEK | Have job or work last week | V3071 | |
| MONTHSATADDRESS | How long at this address (months) | V3031 | N/A |
| TIMEATADDRESS | How long at this address (years) | V3032 | N/A |
| TIMESMOVEDIN5YEARS | How many times moved in the last 5 years | V3033 | N/A |

**Table D.3:   Crosswalk between restricted-use files and public-use files for select variables: Incident-level file**

| RDC Variable Name | Description | Analogous PUF Variable | Notes |
|---|---|---|---|
| YEARQUARTER | Year and quarter of interview | YEARQ | - YEARQUARTER and CTRLNUM uniquely identify a household interview |
| CTRLNUM | Control number | Not available on PUF | - YEARQUARTER, CTRLNUM, and LINENUM uniquely identify a person interview |
| LINENUM | Person line number | V4010 | |
| VWGT | Adjusted victimization weight | WGTVICCY | Used to weight the incident data to count victimizations |
| SERIESWGT | Series-adjusted victimization weight | SERIES_ WEIGHT | Equal to the victimization weight multiplied by the number of victimizations in a series |
| TOCNEW | Type of crime code | V4529 | Incident records with TOCNEW = 99 are unclassified crimes and should not be used |
| INCIDENTPLACE | The city, town, or village where the incident occurred | V4022 | Used to exclude victimizations that occurred outside the United States |
| SIXORMOREINCIDENTS | How many incidents | V4017 | Used to define whether an incident is a series crime |
| INCIDENTSSIMILAR | Incidents similar in detail | V4018 | |
| RECALLDETAILS | Enough detail to distinguish incidents | V4019 | |
| INCIDENTLOCATION | Where incident occurred | V4024 | N/A |
| HHMEMPRESALLOC | Respondent present | V4045 | N/A |
| WHICHMEMBERALLOC | Household members who were present | V4047 | N/A |
| WEAPONPRESENT | Whether offender had a weapon | V4049 | Used to determine the presence and type of weapon used during incident |
| WEAPON_L, WEAPON1-WEAPON7, WEAPON_R | Type of weapon | V4050 – V4058 | |
| ATTACKALLOC | Whether offender was hit or attacked | V4060 | Used to determine whether victim was injured and the type of injury |
| TRYATTACKALLOC | Whether offender tried to attack | V4062 | |
| THREATENALLOC | Whether offender threatened with harm | V4064 | |
| INJURY_L, INJURY1-INJURY11, INJURY_R | Type of injury | V4111-V4122 | |
| POLICEINFORMED | Reported to police | V4399 | Used to determine whether crimes were reported to police and the most important reason for reporting or not reporting an incident |
| POLICEFINDOUT | How police found out | V4400 | |
| NOTREPORTIMPORTANT | Most important reason not reported to police | V4422 | |

| RDC Variable Name | Description | Analogous PUF Variable | Notes |
|---|---|---|---|
| REPORTIMPORTANT | Most important reason for reporting to police | V4437 | |
| AGENCYHELP | Help from victim agencies | V4467 | N/A |
| DOINGATINCIDENTTIME | Activity at time of incident | V4478 | N/A |