This report was prepared by RTI International using federal funding provided by the Bureau of Justice Statistics.

Document Title: National Crime Victimization Survey Historical Trends Project, 1973-2014

Authors: Andrew Moore, RTI International
Marcus Berzofsky, RTI International
Duren Banks, RTI International
Shannan Catalano, Ph.D., Bureau of Justice Statistics

Document No.: 250655

Publication Date: February 14, 2019

Award No.: This project was supported by award number 2011-NV-CX-K068

Abstract:

The National Crime Victimization Survey (NCVS) Historical Trends (NHT) project integrates cross-sectional concatenated data from the historical National Crime Survey (NCS) and the contemporary NCVS into a series of NHT data files covering the entire lifespan of the survey from 1973 to 2014. This technical report describes the linking methodology and considerations that went into integrating the NCS and NCVS for the NHT master data files, extract files, and initial statistical tables. The end of this report briefly describes the other components of the project and information on where these products can be found.

This page intentionally left blank.

National Crime Victimization Survey Historical Trends Project, 1973-2014

Andrew Moore
Marcus Berzofsky
Duren Banks

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

Shannan Catalano, Ph.D.

Bureau of Justice Statistics
810 Seventh Street, NW
Washington, DC 20001

CONTENTS

LIST OF TABLES

LIST OF FIGURES

SECTION 1. NCVS HISTORICAL TRENDS OVERVIEW

1.1 Purpose and Context

As part of ongoing improvements to the National Crime Victimization Survey (NCVS), the Bureau of Justice Statistics (BJS) has invested in the continuous examination of the technical and substantive issues related to the collection and dissemination of victimization statistics. In support of improved dissemination of victimization statistics, BJS initiated the NCVS Historical Trends (NHT) project.

The central task of the NHT project is to integrate cross-sectional concatenated data from the historical National Crime Survey (NCS) and the contemporary NCVS into a series of NHT data files covering the entire lifespan of the survey from 1973 to 2014. Although the victimization survey data is publicly available, the data files can be complex, making it difficult for most persons to use the files or generate historical estimates.

The primary purpose of the NHT data files is to provide BJS analysts, external researchers, and the broader community of BJS stakeholders with readily available access to the entire series of victimization survey data, which began in 1973. To date, researchers or the interested public have not had a single location where they could quickly access information on trends or retrieve data they could then use in their own analyses.

A secondary purpose of the NHT is to provide detailed documentation on the data file development. Historical documentation for NCS and NCVS data files can be difficult to locate or, in some instances, can seem incomplete or inconsistent. In developing the NHT data file, RTI International (RTI) and BJS have generated standardized documentation that details changes to NCS-NCVS data over time and decisions made to allow for consistent trend analysis over time. These efforts include but are not limited to imputing missing or corrupted data, imposing consistency on variables for which measurement changed over time, accounting for historical changes in the definition of series crimes, and adjusting for changes in the definition and measurement of rape or sexual assault over time.

Providing this documentation and refined data in one location assists both internal and external users in conducting analyses with the historical and contemporary victimization survey data that BJS has vetted. The NHT project was designed to provide centralized access to information from historical and contemporary survey years through a crosswalk file and technical documentation to guide the application of these products in generating historical victimization trend analyses.

The NHT data files were also developed to include all variables for which meaningful trends could be generated over time. For instance, demographic variables such as sex, race, and age can be generated for the entire history of the NCS and NCVS. In contrast, police response variables, such as whether an arrest was made, are available beginning in 1986. The length of any given trend can be different based on the variable under consideration. Continuity over time was the overriding criteria for NHT variable selection. In addition, the NHT project included variables that were—

- descriptively and analytically important
- consistent/replicable over time
- present within the data file beginning no later than 1986.

In addition, the structure and organization of the NHT data allow for easily appending additional years of NCVS data to the existing data files.

1.2 Components of the NHT Project

As part of the NHT project, RTI and BJS developed a series of products that enable analysis of historical trends in victimization using NCS and NCVS data. These products include—

1. a crosswalk describing key variables of interest and changes in their measurement over time
2. master data files and codebooks linking the crosswalk variables as measured from 1973 to 2014
3. extract files and codebooks to support analysis of specific historical trends research questions
4. program code that allows users to append additional years of data
5. an initial statistical figures report with accompanying estimates describing key trends in victimization survey data from 1973 to 2014.

The primary focus of this technical report is describing the linking methodology and considerations that went into integrating the NCS and NCVS for the NHT master data files, the extract files, and the initial statistical tables. The end of this report briefly describes the other components of the project and information on where these products can be found.

SECTION 2. INTEGRATION OF NCS AND NCVS

The Bureau of Justice Statistics (BJS) has collected data on personal and household victimization since 1973 through the NCVS and its predecessor, the National Crime Survey (NCS). Each year approximately 100,000 individuals from 50,000 households are sampled to allow estimates of criminal victimization (Kindermann, Lynch, & Cantor, 1997). The guiding principles for the design of the NCS and NCVS are to—

1. develop detailed information about the victims and consequences of crime
2. estimate the numbers and types of crimes not reported to the police
3. provide uniform measures of selected types of crimes
4. permit comparisons over time and across subgroups.

The NCVS and NCS have undergone a number of survey protocol and questionnaire changes since the NCS was first fielded in 1973. Design changes made over the course of the survey must be examined and adjusted for prior to generating historical victimization trend estimates. The NHT crosswalk (see Section 5.3) details measurement and protocol changes that should be considered when examining NCVS historical trends, and key (or major) changes are summarized below.

2.1 General Changes in Measurement over Time

In 1976, the National Academy of Sciences (NAS) released findings from its evaluation of the NCS program. NAS provided the following recommendations to the program—

- eliminate the survey's commercial and central cities components
- revise the crime screening questions to improve prompting of respondents' memories
- add questions that allow examination of ecological factors and lifestyle activities associated with crime victimization
- add questions about crime preventive or protective measures that respondents have taken (Eidson Penick & Owens, 1976).

BJS responded to these recommendations with revisions to the survey questionnaire and protocols over the next several years. BJS eliminated the survey's commercial and central cities components in 1977 because of concerns about viability over time and overlap with existing

information reported through the Federal Bureau of Investigation's Uniform Crime Reporting Program.

Changes judged not to affect rates were implemented beginning in 1986 and included revisions to the incident report for the victim to describe the characteristics and circumstances of the crime incident. Other measurement changes are described below as they relate to household-, person-, and incident-level variables.

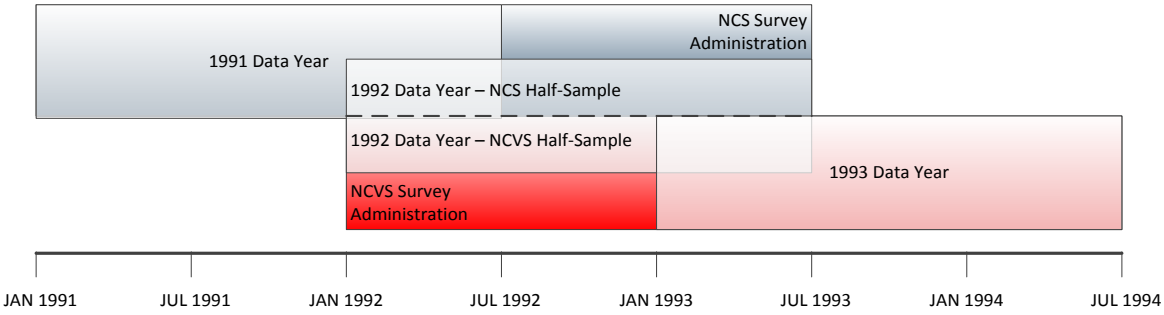
2.1.1 Screener Changes

Changes judged to affect victimization rates were implemented in 1992 and 1993 to allow the study of their impact on NCVS estimates and to develop adjustments for crime types and subpopulations significantly affected by the screener changes. The screener changes were designed to help respondents define and report incidents that may not have been identified or recalled under the previous NCS screener: victimizations by non-strangers, attempted victimizations, and those not reported to the police. In particular, the revised screener—

- asked specific questions about rape or sexual assault
- emphasized that assault includes grabbing, punching, and choking
- reduced the number of ambiguous terms so that respondents better understand what the NCVS is trying to measure
- asked lifestyle questions to better stimulate recall
- reminded respondents of the reference period to reduce telescoping.

The new screener questions were implemented in half the survey's sample for 18 months beginning in January 1992 as shown in **Figure 2-1**. The remaining half sample was enumerated using the old methodology, thereby enabling an evaluation of the impact of the new screening questions.

Figure 2-1. Implementation of the Split-Sample Design, January 1992–June 1993



The redesigned screening questions had a differential impact on each measured crime type. For robbery, burglary, and motor vehicle theft, the estimates from the old and new methodologies showed little or no difference. The rates of every other measured offense increased in the redesigned survey, and increases were greatest for difficult-to-measure offenses, such as rape or sexual assault and domestic violence.

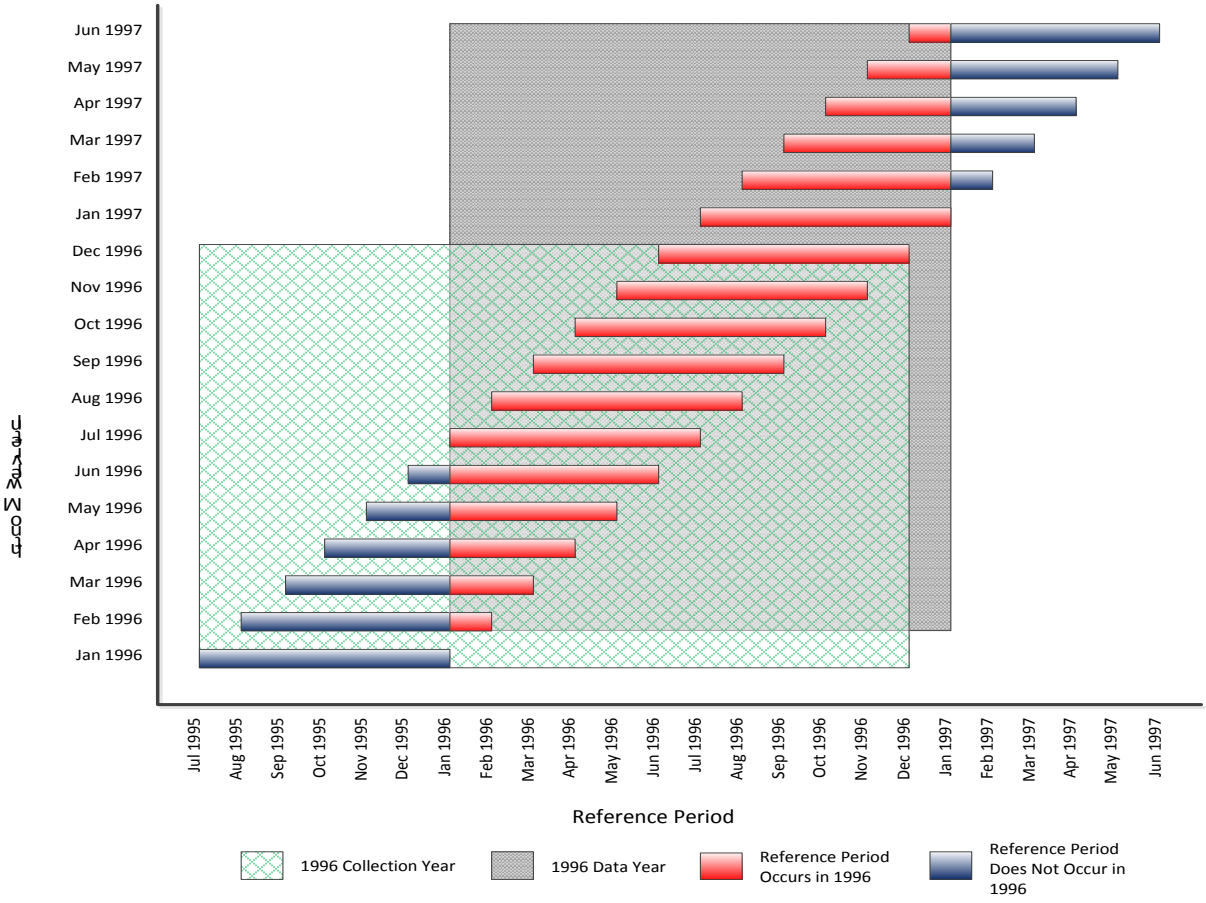
Analysis of respondent subgroups revealed a general pattern of increased recounting for traditionally low-victimization groups, compared to traditionally high-victimization groups. Ratio adjustments were replicated and updated as part of the NHT project to account for the increased reporting due to the revised screener (see Section 3).

2.1.2 Transition from Data Year to Collection Year

Prior to 1996, BJS published estimates based on the data-year format so that crimes were counted in the year in which they occurred. Because the survey uses a 6-month reference period that ends on the last day of the month before the interview month, the data-year format requires 18 months of interviews to generate annual victimization estimates. For example, to generate 1996 data-year estimates, interviews would be conducted from January 1996 to June 1997 and reported incidents would be weighted based on the proportion of the reference period occurring in 1996. As shown in **Figure 2-2**, the reference period for an interview in March 1996 would cover 2 months of the 1996 data-year (January 1996 and February 1996). For any incidents reported during this interview, the weight assigned to the incident would be adjusted to reflect the portion of the reference period that occurred in 1996. To allow more timely reporting of victimization estimates, BJS transitioned to the collection-year format for the annual criminal victimization bulletin and statistical tables beginning in 1996 (Ringel, 1997). Under the

collection-year method, victimizations are counted in the year in which they are reported so that estimates may be generated using only 12 months of data (i.e., January to December). Using this approach, if a crime was reported during an interview that took place in 1996, that incident would be included in the 1996 collection-year estimates regardless of whether the incident actually occurred in 1995 or 1996.

Figure 2-2. Relationship Between Collection-Year and Data-Year, July 1995–June 1997



2.1.3 Series Victimization

In 1979, the NCS incident report was revised to collect additional information on high-frequency repeat victimizations (known as series victimizations)—incidents occurring multiple times during the reference period that were too similar to be distinguished from one another. Prior to 1979, victims of a series crime estimated the number of incidents in the series by choosing one of three categories: (1) 3 or 4, (2) 5 to 10, or (3) 11 or more. Beginning in 1979,

victims were asked to report the exact number of incidents in the series and assign them to specific calendar quarters to permit additional study of series crimes and to determine the feasibility of combining them with regular crimes for tabulation. The NCS defined a series crime as three or more incidents that the respondent could not remember as discrete events.

In the 1992 redesign, the threshold for a series victimization was raised from three to six incidents, and the respondent had to report these incidents as essentially the same type of crime. The NCVS also included more information about series crimes, including whether they occurred at the same place, were committed by the same person, and characteristics of the series as a whole in addition to the latest incident. Historically, BJS excluded series victimizations from counts and rates of criminal victimization because victims only provided details on the last incident in the series. However, with the release of the 2011 NCVS estimates, BJS began including series crimes up to 10 incidents in annual estimates of criminal victimization. **Figures 2-3** and **2-4** present the items from the Crime Incident Report interviewers use to identify series victimizations in the NCS and NCVS, respectively.

Figure 2-3. Identification of Series Crimes in NCS, NCS-2 Crime Incident Report, 1973–1991

CHECK ITEM B	Is this incident report for a series of crimes? (Note: Series must have three or more similar incidents which respondent cannot recall separately.)	Yes – Ask 3a No – **SKIP** to 4b
3a. Altogether, how many times did this happen during the last 6 months?		_____ Number of incidents

Figure 2-4. Identification of Series Crimes in NCVS, NCVS-2 Crime Incident Report, 1992–2014

4. Altogether, how many times did this type of incident happen during the last 6 months?		_____ Number of incidents
CHECK ITEM B	**5a. How many incidents? (Refer to 4.)**	1-5 incidents (not a "series") – SKIP to 6 6 or more incidents – Ask 5b
CHECK ITEM C	**5b. Are these incidents similar to each other in detail or are they different types of crimes?**	Similar – Ask 5c Different (not a "series") – SKIP to 6

CHECK ITEM D	**5c. Can you recall enough details of each incident to distinguish them from each other?**	Yes (not a "series") No (is a "series")

2.2 Household-Level Measurement Changes

Household variables include household identifying information, characteristics of the housing unit, characteristics of the respondents from that housing unit, the number of respondents and crime incidents reported, and variables to link the housing information to person-level respondents and incident-level data.

Household characteristics include household income, urbanicity of household, whether the dwelling was owned or rented, and the type of structure (e.g., single-family home, multiple-family dwelling, mobile home, motel). Household income has been measured variously over the life of the victimization survey, and an increasing proportion of income data has been missing over time. In addition, a household composition variable was added in 1993; no directly comparable measure existed before 1993. Due to a data processing issue, urbanicity is unavailable in the historical trends data for 1977, 1978, and 1979.

2.3 Person-Level Measurement Changes

Person-level variables include identifying information, responses to victimization screener questions, and sociodemographic variables. The main changes to the person-level variables stem from the significant changes to the screener form implemented with the 1992 redesign. Sex, race, ethnicity, age, Metropolitan Statistical Area status, marital status for adults, and family status for juveniles are generally available throughout the 1973–2014 study period with some limited exceptions as noted here.

Questions about unemployment and respondents' attempts to find work were discontinued in 1986. The race/ethnicity measures also changed over time. Prior to 2003, respondents could only select one race category, but beginning in 2003, could identify as multiple race categories. Prior to 1986, many ethnicity categories were available, but only "Hispanic" and "Non-Hispanic" are available ethnicity options since that time.

BJS also introduced new measures in 1986 designed to capture lifestyle variables theorized to be related to victimization risk. Prior to this change, the NCS only included sex, race/ethnicity, age, marital status, relationship to other household members, occupation, membership in the Armed Forces, education, frequency of residential mobility, and family income. The expanded NCVS person-level variables included—

- occupational responsibilities, hours, locations, and contacts; commuting patterns
- for students, perceptions of safety in and around schools
- evening, shopping, and leisure activities
- neighborhood characteristics
- perceptions of safety at home and other places where respondents spend time
- precautions respondents took to protect themselves
- for respondents who experienced personal victimization, what they were doing at the time of the incident.

2.4 Incident-Level Measurement Changes

As described above, the 1992 NCVS redesign included screener changes designed, in part, to better measure rape or sexual assault. The incident form was revised to add violent crime questions that allow reporting of victimizations that involved sexual contact and those that did not. For victims reporting sexual assault from 1973 to 1991, the crime may have been classified as a rape, robbery, aggravated assault, simple assault, or not reported at all to an interviewer. The revised incident form also included measures of vandalism—not as discrete incidents, but to collect information on the type of property damaged and the costs of vandalism during the reference period.

The NCS defined larceny as household or personal, and then further divided personal larceny as "with contact" or "without contact" crimes. The distinction between household and personal larceny was based solely on the location of the property that was taken. The 1992 NCVS redesign removed this distinction and classified almost all larcenies as thefts, which fall under the property crime classification. (Contact thefts, such as purse snatching and pocket-picking were the exception.)

Otherwise, crime classification categories have been defined largely in the same manner since the NCS was launched in 1973. Violent victimizations include rape, robbery, aggravated assault, and simple assault. Property crimes include burglary, motor vehicle theft, and other theft. Simple assault includes attempted or completed attacks without a weapon, and aggravated assault includes attempted or completed attacks with a weapon and completed attacks with serious injury. Robbery includes attempted or completed thefts by force or threat of force. The definition of rape was further refined and expanded with the 1992 redesign to include sexual assault and is defined as attempted or completed attacks involving unwanted sexual contact, verbal threats, or forced penetration.

Additional measures for time and place of occurrence, medical treatment, property loss, and reporting to police were added to the NCS incident report in 1979. Other measures were introduced in 1986 that provided information on victims' perception of drug and alcohol use by violent offenders, protective actions that victims and bystanders took, actions of the police in the investigation of reported crimes, and contacts between the victim and other persons or organizations in the criminal justice system (Whitaker, 1989).

Other new incident-level measures included the type of location where the crime occurred, whether the offender was believed to be a gang member, the victim's activity at the time of the incident, and additional response categories for several questions.

Measures of weapon presence are an example of changing levels of detail over time. The redesigned NCVS instrument implemented in 1986 asked whether the offender had a weapon, with a "yes" response eliciting additional questions about the type. Prior to these questionnaire revisions, the presence of a weapon was measured through a single variable that included values for several weapon types (firearm, knife, other type of weapon, and type of weapon unknown, or no weapon present and unknown whether a weapon was present).

In 1986 this variable was replaced with questions asking whether more individual weapon types were present, including various types of firearms, sharp objects, and blunt weapons. Some variables map exactly throughout the study period, while others are available only from 1986.

2.5 BJS's Previous Approach to Integration of NCS and NCVS

The previous method developed by BJS to integrate NCS and NCVS data serves as a foundation for the current work, and the NHT builds on this previous approach. Although each change described in the preceding sections must be considered during such an undertaking, perhaps the greatest challenge is the series of methodological improvements to the NCS that culminated in 1992 with a redesigned and renamed survey (Kindermann et al., 1997).

The primary change implemented during the 1992 redesign was the phase-in of a newly designed questionnaire designed to elicit more reports of crime through the use of new cues and additional prompts. The methodological changes implemented in 1992 had the intended consequence of generally increasing rates of crime. NCVS estimates were higher than NCS estimates for most crime types, including violent crime (up 49%), rape (up 157%), aggravated assault (up 23%), simple assault (up 75%), property crime (up 23%), household burglary (up 20%), and household theft (up 27%). For these crimes, NCS estimates were adjusted upward to bring them level with NCVS rate estimates.

To determine the most appropriate method for adjusting NCS estimates, BJS evaluated two potential methods using data from the 1992 split-sample design (Rand, Lynch, & Cantor, 1997). The first method used an overall crime-specific ratio adjustment based on the rates of criminal victimization reported in each design during the 1992 overlap period. The second method identified characteristics of both the incident and the respondent that interacted with the survey design, and then used these variables to generate subgroup-specific ratio adjustments for each major type of crime. Although crime rates for certain subgroups were affected differentially by the redesign, the overall adjusted crime rates computed using the subgroup-specific method did not differ significantly from crime rates computed using the simpler crime-specific ratio adjustment (Rand et al., 1997).

Based on these findings, an overall crime-specific ratio adjustment method was selected to adjust NCS estimates prior to 1992 and make them comparable to estimates from the redesigned NCVS. For each major crime type, the adjustment was calculated as the ratio of the NCVS data-year estimate from the split-sample overlap period to the NCS data-year estimate from the same time period. **Table 2-1** presents the rates of criminal victimization reported in

both the NCS and NCVS during the 1992 overlap period and the crime-specific ratio adjustments based on the methodology BJS developed (Kindermann et al., 1997).

Table 2-1. Comparison of Estimated NCS and NCVS Victimization Rates, 1992

Crime Type	NCS Crime Rate[1]	NCVS Crime Rate[1]	NCVS/NCS Ratio
Personal Crimes	34.4	49.6	1.44‡
Violent Crimes	32.1	47.8	1.49‡
Rape	0.7	1.8	2.57‡
Robbery	5.9	6.1	1.03
Assault	25.5	40.0	1.57‡
Aggravated	9.0	11.1	1.23‡
Simple	16.5	28.9	1.75‡
Personal Theft	2.4	1.8	0.75
Property Crimes	264.5	325.3	1.23‡
Household Burglary	48.9	58.6	1.20‡
Household Theft	195.5	248.2	1.27‡
Motor Vehicle Theft	20.1	18.5	0.92

‡ The ratio of the NCVS to NCS estimates was statistically significant from 1 at the 90% confidence level.
[1]Rate of victimization per 1,000 persons or households.
Note: Adapted from Kindermann, C., Lynch, J. P., & Cantor, D. (1997). *Effects of the redesign on victimization estimates* (NCJ 164381). Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics.

Based on these results, BJS adjusted the following major crime types during the 1992 redesign overlap period: rape,[1] aggravated assault, simple assault, household burglary, and household theft. Significant differences were not observed in the rates of victimization reported in the NCS and NCVS for robbery, personal theft, and motor vehicle theft; thus, NCS estimates were not adjusted for these crime types. For adjusted NCS estimates of composite measures (e.g., violent crimes), the ratio-adjusted crime rates for each specific component are calculated and aggregated to obtain the combined measure. Ratio-adjusted NCS crime rates for rape, aggravated assault, simple assault, household burglary, and household theft were calculated as follows:

$$AR_j^c = K^c R_j^c$$

[1]Respondents were not specifically asked about sexual assaults other than rape until the 1992 survey, so these incidents were excluded from the analysis presented in Table 2-1.

where

AR_j^c equals the Ratio-Adjusted rate for crime type c in year j (j=1973–1991),
K^c equals the NCVS to NCS ratio adjustment shown in Table 2.1 for crime type c, and
R_j^c equals the unadjusted crime rate for crime type c in year j (j=1973–1991).

SECTION 3. APPROACH FOR INTEGRATING IN NHT

Although the previous approach used to integrate the historical NCS data with the contemporary NCVS is a good framework for measuring trends in criminal victimization from 1973 to 2014, some possible limitations of the analysis described by Rand et al. (1997) include—

- Given crime is, generally, a rare event, using the split-sample data from a single survey year restricted the analyses to certain subgroups and crime types.
- It was not possible to evaluate all possible respondent and incident characteristics to determine if a certain subgroup-specific adjustment would produce different results than the simpler overall crime-specific adjustment.
- The overall crime-specific adjustment assumes that the differential effect of the redesign on subgroups did not affect the overall trends in crime.
- The adjustment method assumes the long-term impact of the redesign was accurately reflected in the split-sample design implemented in 1992.

In addition to these limitations, changes in the enumeration of series victimizations, the measurement of rape and sexual assault, and the reporting of estimates on a collection-year basis (described in Sections 2.1 and 2.3) were not accounted for in the previous analysis to determine the most appropriate adjustment method for integrating data from the NCS with data from the NCVS (described in Section 2.5).

3.1 Adjustment Approaches Considered in the NHT Project

One common theme underlying the limitations discussed above is that the ratio-adjustment method is applied at the crime-type level and does not account for the differential impact of the redesign on population subgroups. This limitation was imposed because of the relatively small sample sizes the split sample provided and because certain crimes can be rare events in the population of interest. To address this shortcoming as part of the NHT project, incorporating additional years of data beyond the overlap period were examined as a possible solution. With the larger sample sizes afforded by using additional years of data, the following options are alternatives to the crime-type specific ratio adjustments.

1. **Subgroup-Specific Ratio Adjustments**. If the addition of data beyond the overlap period does not significantly impact the overall crime-type ratio adjustments (i.e.,

changes in survey design do not appear to be confounded with temporal changes in criminal victimization), more precise adjustments could be generated on the basis of characteristics of the victim and/or incident.

2. **Constrained Exponential Raking Model**: RTI developed this model-based method and used it on the National Survey of Drug Use and Health to account for a new data collection instrument in the 1994 survey (U.S. Department of Health and Human Services, 1997). This particular method is desirable because it allows the effect to be defined at the crime type and person level, and the modeling procedure allows more explanatory variables in the adjustment compared to a ratio correction factor. With this method, model parameters are estimated by solving generalized raking equations that can be conducted in SUDAAN's WTADJUST procedure.

As part of the NHT project, subgroup-specific adjustments were examined using one additional year of data (i.e., the 1991 NCS data were combined with the NCS portion of the 1992 split-sample and the 1993 NCVS data were combined with the NCVS portion of the 1992 split-sample data) and two additional years of data. Although the redesigned questionnaire had a differential impact on some subgroups, this analysis failed to consistently demonstrate that a more complex adjustment was needed to integrate the NCS and NCVS data. This observation, along with the fact that adding additional years of data could potentially confound changes in survey design with temporal changes in crime rates, ultimately led analysts to refrain from using data beyond the 1992 split-sample overlap period to adjust estimates of criminal victimization in the NCS. Appendix A presents the results of these analyses.

Once the decision was made not to use additional years of data to develop adjustments for NCS estimates, the constrained exponential raking model was not evaluated further. This model-based method is best used when sample sizes and prevalence measures are sufficient to support many explanatory variables in the model. Although it could have been used for composite crimes (e.g., total violent crime) or more prevalent crime types (e.g., simple assault), this method would likely not have been feasible for all crime types of interest—and the use of multiple adjustment methods was seen as less desirable for data users.

Upon determining that subgroup-specific adjustments were not needed and more complex adjustment methods were not likely to be feasible with the data from the split-sample, the project team evaluated the overall crime-specific ratio adjustment more closely. This assessment focused on changes in measurement and reporting that were not incorporated into the initial adjustment

factors BJS developed. Namely, the impact of changes in the enumeration of series victimizations, measurement of rape and sexual assault, and reporting of estimates on a collection-year basis on the crime-specific ratio adjustments were examined.

3.2 Adjustment for NHT Project

To accommodate the changes in enumerating series victimizations, measuring rape and sexual assault, and reporting of estimates on a collection-year basis, a series of steps were undertaken to update the approach used to integrate historical NCS and contemporary NCVS data. Important considerations for the most appropriate adjustment method included—

- ease of use by stakeholders
- ability to be easily updated to include future iterations of the survey
- maintaining consistency with current reporting guidelines for estimates of criminal victimization
- statistical validity.

The first step of the analysis was to reproduce the crime-specific ratio adjustments BJS made during the 1992 redesign as presented in **Table 2-1**. The ratio adjustment estimates showed minor differences between the original work BJS did and the work completed as part of the NHT project for the following crime types: rape (2.57 compared to 2.62), robbery (1.03 and 1.02), and aggravated assault (1.23 and 1.24). These minor differences are most likely attributed to whether estimates are rounded before or after the calculation of ratio adjustment factors. Once the original adjustment factors had been satisfactorily reproduced, the evaluation of the most appropriate adjustment for the NHT consisted of three steps:

1. assess impact of producing adjustment based on collection-year data
2. assess adjustments that incorporate series victimizations
3. assess the impact of adjusting for sexual assault.

3.2.1 Assess Impact of Producing Adjustment Based on Collection-Year Data

One objective of the NHT is to provide seamless estimates of the NCS and contemporary NCVS data. Thus, analysts had to determine if calculating adjustments based on the collection-year method would result in substantively different adjustments from what was done when

estimates were calculated using the data-year format. Collection-year estimates were calculated for both the 1992 NCVS and 1992 NCS, and the ratios of rates were compared to the corresponding data-year estimates. For this analysis, series crimes were excluded. Because the goal of this exercise was to isolate the impact of switching to the collection-year format, the inclusion of series crimes may have confounded findings because the definition of a series crime is not consistent across the NCS and NCVS (see Section 2.1).

This analysis showed only minor differences between the data-year and collection-year ratio adjustments. The most notable change in switching from data-year estimates to collection-year estimates occurred for motor vehicle theft. Under the data-year method, the ratio adjustment for motor vehicle theft (0.92) was not significantly different from 1, indicating no significant difference between the NCS and NCVS as it relates to the measurement of this type of crime. However, under the collection-year method, this ratio adjustment (0.84) was significantly less than 1 at the 95% confidence level, indicating that the rate of motor vehicle theft as measured in the NCVS was significantly lower than the rate as measured in the NCS. For all other crime types, only minor differences were observed between the ratio adjustments calculated using the data-year method and collection-year method with no other changes noted in the significance level of the ratio. For complete details on the NCVS/NCS ratio adjustments calculated using data-year and collection-year estimates, see **Table B-1** in Appendix B.

3.2.2 Assess Impact of Including Series Victimizations in Adjustment

Another change implemented since the previous analysis to integrate NCS and NCVS data is the inclusion of series crimes into ongoing NCVS estimates. The impact of including series crimes on the ratio adjustments was examined in more detail by calculating collection-year estimates from the 1992 NCS and 1992 NCVS with series crimes included and comparing the subsequent ratio adjustments to those that excluded series crimes. A significant difference in these ratios would suggest that the differential counting rules should be considered when adjusting historical NCS estimates, particularly for crimes with a high occurrence of repeat victimizations such as sexual assault.

To conduct the assessment of including series crimes in the adjustment ratio, analysts needed an approach for handling missing data for the number of incidents in the series. For some

incidents identified as series crimes, the total number of incidents in the series is missing on the data files. Current NCVS counting procedures replace missing values for the total number of incidents with a value of 6, which is the minimum number of similar incidents required for a victimization to be considered a series crime under the NCVS. That is, an incident file record in the NCVS that is classified as a series crime represents a minimum of six unweighted incidents during the prior 6 months. Under the NCS, however, a minimum of three incidents (victimizations) was required for an incident to be classified as a series crime.

To account for this difference, two approaches for handling missing data in the variable for total number of incidents in the NCS were examined as part of the NHT project. For the first method, missing values in the total number of incidents variable in the NCS were set to a value of 3, which is the minimum number of incidents required for a crime to be considered a series under the NCS methodology (i.e., an incident-level record identified as a series crime represents a minimum of three unweighted incidents during the prior 6 months in the NCS). Under the second method, missing values were replaced with a value of 6 to match what is currently done in the NCVS and to simplify file creation and analyses. Comparing these two methods (i.e., missing=3 and missing=6) failed to show any significant differences in the ratio adjustments (see **Table B-2** in Appendix B). This result was largely expected because about 5% to 10% of personal victimizations in the 1992 NCS were classified as series crimes and about 10% to 20% of series crimes had a missing value for the total number of incidents. When combined, the total number of incidents that were both classified as a series crime and had a missing value for the total number of incidents made up less than 1% of all victimizations for each crime type in the 1992 NCS data.

Given these results, the decision was made to set missing values for the total number of incidents to a value of 6 in the NCS portion of the 1992 split-sample and to evaluate the overall impact of including series crimes on ratio adjustment factors. Collection-year estimates from the 1992 split-sample were calculated with series crimes excluded and series crime included along with the resulting NCVS/NCS ratio adjustments (see **Table B-3** in Appendix B).

The impact of including series crimes was greatest for the crime types of rape and robbery. During the 1992 split-sample, a larger number of rape incidents classified as series

crimes were present in the NCS than in the NCVS. The NCS rate of rape with series crimes excluded was 0.7 victimizations per 1,000 persons age 12 or older, compared to 1.6 per 1,000 observed for NCS rape estimates that included series crimes. The differential impact of series crime was associated with a drop in the ratio adjustment for rape from 2.59 for estimates excluding series crimes to 1.53 for estimates including series crimes.

A similar pattern was observed for robbery where a larger number of robbery incidents were classified as series crimes in the NCS than in the NCVS. The NCS rate of robbery with series crimes excluded was 5.8 victimizations per 1,000, compared to 7.9 per 1,000 for NCS robbery estimates that included series crimes. This disparity in the impact of series crimes between the NCS and NCVS was associated with a drop in the ratio adjustment for robbery from 1.08 for estimates excluding series crimes to 0.86 for estimates including series crimes. Although the ratio calculated with series crimes excluded was not significantly different from 1, the ratio adjustment with series crimes included was significantly less than 1. Given these results, series crimes were included in the adjustments.

3.2.3 Assess Impact of Adjusting for Sexual Assault

Previous adjustments to NCS data excluded sexual assaults from adjustment factors used to make data comparable over time. To assess the impact of adjusting for sexual assault, the first step was to examine the relationship between rape and sexual assault in the NCVS. From1992 to 2014, the percentage of rapes and sexual assaults classified as rapes averaged approximately 59.9% over this period, with a low of 47.1% in 2014 and a high of 76.9% in 2008 (see **Table B-4** in Appendix B).

In the NCVS portion of the 1992 split-sample, rape accounted for approximately 58.8% of all crimes in the rape and sexual assault category. Although the percentage of rapes and sexual assaults defined as rape fluctuated from year to year under the NCVS, the percentage of rapes and sexual assaults classified as rape in 1992 was representative of the percentages observed for the remaining years in which the NCVS was implemented (mean=59.9%, median=59.6%). Once it was determined that the 1992 data were not anomalous with respect to the relationship between rape and sexual assault, the impact of including sexual assault in the NCVS portion of the ratio

adjustment was analyzed by calculating the NCVS/NCS adjustment ratios with and without sexual assault included (see **Table B-5** in Appendix B).

In the NCVS portion of the 1992 split-sample data, using collection-year estimates and including series crimes, the rate of rape-only was 2.39 per 1,000, compared to 4.07 per 1,000 when sexual assault was included. Consequently, the resulting adjustment ratio that included sexual assaults, 2.61, was much higher than the adjustment ratio in which sexual assaults were excluded from NCVS estimates, 1.53.

3.3 Final NHT Adjustment Factors

To better align the reporting of historical NCS estimates prior to 1992 with the current reporting guidelines, the following changes were implemented during the creation of crime-specific ratio adjustments as part of the NHT project.

- Collection-year estimates from the 1992 overlap period were used to create NCVS/NCS ratio adjustments.
- Series crimes were included in 1992 estimates to create the ratio adjustments; the number of incidents were capped at 10 victimizations, and missing values for the number of incidents in both the NCS and NCVS data were set to 6.
- Sexual assaults from the NCVS portion of the 1992 split-sample were included in the calculation of ratio adjustments so that NCS estimates of rape prior to 1992 were comparable to NCVS estimates of rape and sexual assault.

After incorporating these changes, NCVS estimates were observed to be higher than NCS estimates for most crime types including rape and sexual assault (up 161%), aggravated assault (up 27%), simple assault (up 86%), household burglary (up 15%), and household theft (up 26%).

In contrast, NCS estimates were higher than NCVS estimates among some crime types including robbery (up 17%), personal theft (up 71%), and motor vehicle theft (up 21%). Because previous analyses examining the impact of the redesign showed that the new NCVS questionnaire consistently elicited more reports of criminal victimization from survey respondents than the NCS, RTI and BJS undertook analyses to determine why the NHT data showed NCS survey procedures captured more robbery, personal theft, and motor vehicle theft than did NCVS procedures during the 1992 overlap period. These analyses revealed that

applying contemporary counting procedures for series victimizations to NCS data differentially impacted crime counts in the NCS. As a result, BJS determined that no adjustments would be made to NCS estimates of robbery, personal theft, or motor vehicle theft. BJS was guided in this decision by two related factors.

First, the phase-in of the new NCVS methods was designed to equally balance the number of households and persons receiving the old NCS and new NCVS such that effective comparisons could be made between the two designs and ascertain the impact of new methods. However, in 1992, constraints related to the implementation of computer-assisted telephone interviewing and the necessity of allocating equal but efficient workloads for NCS and NCVS interviewers across geographic areas resulted in an imbalanced sample. About 10% more households and persons were administered the NCS interview than the NCVS; the additional NCS sample was concentrated in larger, urban areas. The distribution of the split-sample by characteristics of the household and person is provided in **Table B-6** of Appendix B. This sample allocation imbalance did not impede comparisons of NCS and NCVS estimates at the time of the redesign.

Second, populated urban areas can be associated with certain crime types such as robbery, personal theft, and motor vehicle theft. To determine if the imbalanced sample could be responsible for the differential crime rates observed between the NCS and NCVS, the NCS weights were post-stratified[2] to the NCVS control totals and the NCS rates of criminal victimization were recalculated. The resulting adjustment ratios were then compared to the unadjusted NCVS/NCS ratios. This exercise revealed no significant differences between the post-stratified adjustment ratios and those that were unadjusted. For more details, see **Table B-7** in Appendix B. In addition, prior to 1992 the threshold for counting an incident as a series victimization was three or more times, compared to six or more times in the NCVS. This change

[2]The NCS household level weights were post-stratified to NCVS control totals by location of residence, population size, household income, and household tenure. The person level weights were post-stratified to NCVS control totals by sex, race, and age.

in counting and classifying series victimizations means that substantially more victimizations were added to NCS counts relative to the NCVS.

Ultimately RTI and BJS determined that larger estimates of robbery, personal theft, and motor vehicle theft observed in the NCS data when applying NHT methods for adjustments was an artifact of changed counting procedures for series crimes and the use of collection year data and did not indicate better data quality in the NCS for these crimes.

Using the 1992 split-sample data to identify measurable differences between the two designs, taking into consideration the use of collection year data, new methods for counting series crimes, and the inclusion of sexual assault, to adjust historical NCS data to ensure comparability with contemporary NCVS data resulted in the final crime-specific ratio adjustments shown in **Table 3-1**.

Table 3-1. Final Adjustment Ratios for NCS Estimates Used in NHT Project, 1973–1991

Crime Type	NCS Crime Rate[1]	NCVS Crime Rate[1]	NCVS/NCS Ratio
Rape/Sexual Assault	1.56	4.07	2.61
Robbery	7.94	6.79	—
Aggravated Assault	12.74	16.14	1.27
Simple Assault	28.21	52.36	1.86
Personal Theft	2.68	1.57	—
Household Burglary	57.31	66.01	1.15
Household Theft	227.38	286.81	1.26
Motor Vehicle Theft	22.53	18.63	—

— No adjustment applied to estimates pre-1992.
[1]Rate of victimization per 1,000 persons or households in the 1992 split-sample data using the collection year method and including series crimes.

3.4 Applying Adjustment Factors in the NHT Project

In contrast to the previous adjustment approach developed by BJS in which final estimates of victimization from the NCS were multiplied by the ratio-adjustment factors to obtain adjusted NCS estimates, in the NHT project, the final ratio adjustment factors presented in **Table 3-1** were applied to the incident weights in the three NCS periods to create comparable estimates between the NCS and NCVS. Using these ratio adjustments, new weights were

calculated by multiplying the NCS incident weights for a given year by the 1992 ratio of NCVS/NCS crime counts. This modified incident weight took the form of—

If $NCS_{92} \leq NCVS_{92}$ then $NCS_{awy} = NCS_{wy} \times (NCVS_{92}/NCS_{92})$
else if $NCS_{92} > NCVS_{92}$then $NCS_{awy} = NCS_{wy} \times 1$

where

NCS_{awy} = adjusted NCS incident weight for a given year *y* (1973–1991)
NCS_{wy} = original NCS incident weight for a given crime in year *y* (1973–1991)
$NCVS_{92}$ = rate of NCVS victimizations for a given crime type in 1992 half-sample
NCS_{92} = rate of NCS victimizations for a given crime type in 1992 half-sample.

The new incident weight NCS_{awy} was applied to the NCS data to generate adjusted NCS victimization estimates. For example, an adjusted victimization rate for crime type *t* and year *y* in the NCS would be computed as—

$$NCS_{aty} = \frac{\sum(NCS_{awy} \times I_t \times W_{series})}{N_y} \times 1{,}000$$

where

NCS_{aty} = adjusted NCS victimization rate per 1,000 persons or households for a given crime *t* in year *y* (1973–1991)
$I_t = \begin{cases} 1 \; if \; incident \; is \; of \; crime \; type \; t \\ 0 \; otherwise \end{cases}$
$W_{series} = \begin{cases} 1 \; if \; incident \; is \; not \; a \; series \; crime \\ number \; of \; incidents \; in \; series \; (\max of \; 10) \; for \; series \; crime \end{cases}$
N_y = total population or persons or households in year *y* (1973–1991).

These adjustments were applied to crimes for which NCVS estimates were significantly higher than NCS estimates, including violent crime, serious violent crime, rape and sexual assault, aggravated assault, simple assault, property crime, household burglary, and household theft. NCS estimates for robbery, personal theft, or motor vehicle theft were assigned an adjustment factor of 1 so that no adjustments were made to the NCS estimates prior to 1992 for these crime types. These adjusted NCS victimization counts were then used to generate rates and percentages following standard BJS estimation procedures.

All final adjusted weights and adjustment factors are included for each record the historical trends data file. In addition, the household- and person-level weights from 1973 to 2014 and the incident weights from 1992 to 2014, on the NHT master data files and extract files have not been altered from the annual survey data files. For more information on the creation of household, person, and incident weights in the NCVS see *National Crime Victimization Survey Technical Documentation* (U.S. Department of Justice, 2014).

3.5 Impact on Standard Errors

As described in the previous section, the NHT project accounts for the NCS redesign by modifying the victimization weight instead of applying ratio adjustment to weighted estimates. A second difference between NHT adjustments of historical NCS data and previous treatment of NCS data is that the generalized variance function (GVF) parameters used to generate NHT estimates were originally designed by the Census Bureau for use with non-series data using data-year estimates from the NCS. However, the NHT data file includes series crimes for up to 10 incidents unlike previously published estimates of NCS data by BJS that excluded series crimes or counted series crimes as equal to one.

The effect of these two differences—using a modified incident weight and GVFs designed for non-series estimates, does not affect rate, percentage, or count estimates of victimization. However, the effect of these two differences does result in an underestimation of standard errors prior to 1992. RTI and BJS conducted analyses and determined that the implications of these underestimations were minor given fluctuations of the trend data over a long term and the NHT emphasis on long-term change rather than year-to-year level changes. However, caution is warranted when assuming year-to-year comparisons not explicitly discussed in analyses. An example of the impact on standard errors of applying the adjustment factors to the incident weight rather than applying the ratio adjustment to weighted estimates is shown in Appendix D.

SECTION 4. VARIANCE ESTIMATES

4.1 Estimating Generalized Variance Functions (GVFs)

GVFs are models of the relationship among estimates, their characteristics (i.e., crime type, year, and subpopulation size), and their variances. In practice, they are estimate type-specific formulas that users can implement with simple software (e.g., Excel) and without knowledge of the National Crime Victimization Survey (NCVS) design. Couzens, Berzofsky, and Krebs (2014) provide guidance about use of GVFs in the NCVS.

For the NHT because the design parameters (i.e., pseudo strata and half sample) do not exist in the National Crime Survey (NCS) study periods, GVFs must be used to compute proper standard errors. Even though direct variance estimation is possible in the NCVS, to be consistent across all periods of analysis, GVFs are recommended for use with NHT data (Shook-Sa, Lee, & Berzofsky, 2015).

4.2 Two and Three Parameter GVF Models

As detailed by Couzens et al. (2014), GVFs are based on regression models. The number of parameters in the model depend on the model fit. In the NCS the GVF was based on a two-parameter model. As discussed by Ash et al. (2008), a three-parameter model accounts for a quadratic curve in the regression model, so NCVS moved to a three-parameter model.

4.3 Using GVFs to Compute Standard Errors of Victimization Estimates

As detailed by Couzens et al. (2014), separate GVFs exist for victimization totals, rates, and percentages. In this section, the NCVS and NCS GVF formulas are presented for each estimate type. **Appendix C** presents the constants produced by the regression models. These constants are provided annually by the U.S. Census Bureau.

4.3.1 Totals

For NCS years 1973–1991 the GVF formula for calculating the standard error for a total is—

$$SE(x) = \sqrt{ax^2 + bx}$$

where *a* and *b* are constants provided through the GVF regression models and *x* is the estimated number of personal or household crime victimizations or incidents.

For the NCVS (1992 to present) the GVF formula for calculating the standard error for a total is—

$$SE(x) = \sqrt{ax^2 + bx + cx\sqrt{x}}$$

Where *a*, *b*, and *c* are the parameters provided through the regression models and *x* is the estimated number of personal or household crime victimizations or incidents

4.3.2 Rates

For the NCS years the GVF formula for calculating the standard error for a rate per 1000 persons or households is

$$SE(r) = \sqrt{\frac{b}{y} r(1000 - r)}$$

where *b* is a constant from the regression models, *y* is the estimated base population of persons or households, and *r* is the victimization rate per 1000 persons or households.

For the NCVS years the GVF formula for calculating the standard error for a rate per 1000 persons or households is—

$$SE(r) = \sqrt{\frac{b}{y} r(1000 - r) + \frac{c}{\sqrt{y}} r\left(\sqrt{1000r} - r\right)}$$

where *b* and *c* are constants from the regression models, *y* is the estimated base population of persons or households, and *r* is the victimization rate per 1000 persons or households.

4.3.3 Percentages

For the NCS years the GVF formula for calculating the standard error for a percentage is—

$$SE(p) = \sqrt{\frac{b}{y}p(100-p)}$$

where *b* is a constant from the regression models, *y* is the estimated base population of persons or households, and *p* is the percentage of interest.

For the NCVS years the GVF formula for calculating the standard error for a percentage is—

$$SE(r) = \sqrt{\frac{b}{y}r(100-p) + \frac{c}{\sqrt{y}}r\left(\sqrt{100p} - p\right)}$$

where *b* and *c* are constants from the regression models, *y* is the estimated base population of persons or households, and *p* is the percentage of interest.

SECTION 5. ADDITIONAL PRODUCTS OF THE NHT PROJECT

As part of the NHT project, RTI and BJS developed a series of products that enable analysis of historical trends in victimization using NCS and NCVS data. These products include master data files, extract data files, and a crosswalk to link key variables of interest across survey administrations from 1973 to 2014. These products are discussed in the following sections.

5.1 NHT Master Data Files

Annual NCS and NCVS data files consist of a household-level file, a person-level file, and an incident-level file. The household-level file contains pertinent information about the household (e.g., household income), characteristics of the surrounding area, and information about the principal and reference persons within the household. The household screening interview determines whether the household experienced any property victimizations during the previous 6 months.

The person-level file contains information on each member of the household age 12 years and older (e.g., age and race). Data are obtained from personal screening interviews that determine whether an individual experienced a personal victimization during the prior 6 months. For any eligible events identified in the screener, a detailed crime incident report is completed for each reported incident with information stored in the incident-level file. The incident file may then be linked to the person- and household-level files through identification variables to allow analysis of criminal victimization across many subdomains defined by the respondent and household characteristics.

When necessary, a unique identifier was created to merge household, person, and incident files together in each survey year. For example, the variable "IDHH – NCVS ID for Households" was not available on data files prior to 1992. To create a unique household identifier on the NCS data files, the household identification number (e.g., V1001 1973–1978) was concatenated with the survey year and quarter (e.g., V1002 1973–1978) to allow the household-level information to be merged with the person- and incident-level information as shown below for the years 1973–1978.

$$IDHH_{1973-1978} = V1001 \,||' - '|| \, V1002$$

Once data files were developed for each survey year, individual years were aggregated into four study periods. The data from these study periods were collected on the core survey instruments including the control card, screener, and detailed crime incident report with study periods chosen so that minimal design and questionnaire changes occurred within each period. **Table 5-1** presents the survey years included in each period. As shown, the NCS was split into three periods while the NCVS consists of a single period.

Table 5-1. Study Periods Used in the NHT, 1973–2014

Period	Survey	Survey Year	ICPSR Study Number
1	NCS	1973–1978	7635
2	NCS	1979–1986	8608
3	NCS	1987–1991	8864
4	NCVS	1992–2014	36143

To create the NHT master data files and simplify analysis, a parent-child relationship across the study periods was used to map exact or similar measures of a construct over time. The parent-child relationship identified one period as the parent period and the other three periods as child periods. For the NHT, the current period (Period 4) was identified as the parent period. Given this, the following taxonomy was developed for each variable on the master file datasets:

- Root variable name: the root variable name corresponds to Period 4 variable in the ICPSR public use files.
- Period extensions: Using the root variable, each period had a period extension added to the variable name to indicate the period it represents. The period extensions were defined as _X where X=1, 2, 3, or 4 for the four periods shown in Table 5-1.

On the NHT master data files, the variable levels within a period were not altered. In other words, even though each variable was mapped to a single parent variable, the definition for an item across each child was left unchanged (i.e., the number of response options or level definitions across child variables may be different). For example, the levels available for household income were not consistent across the four survey periods and, while the household income variable from each period was mapped to the same parent variable, the levels included on each survey period master file could vary.

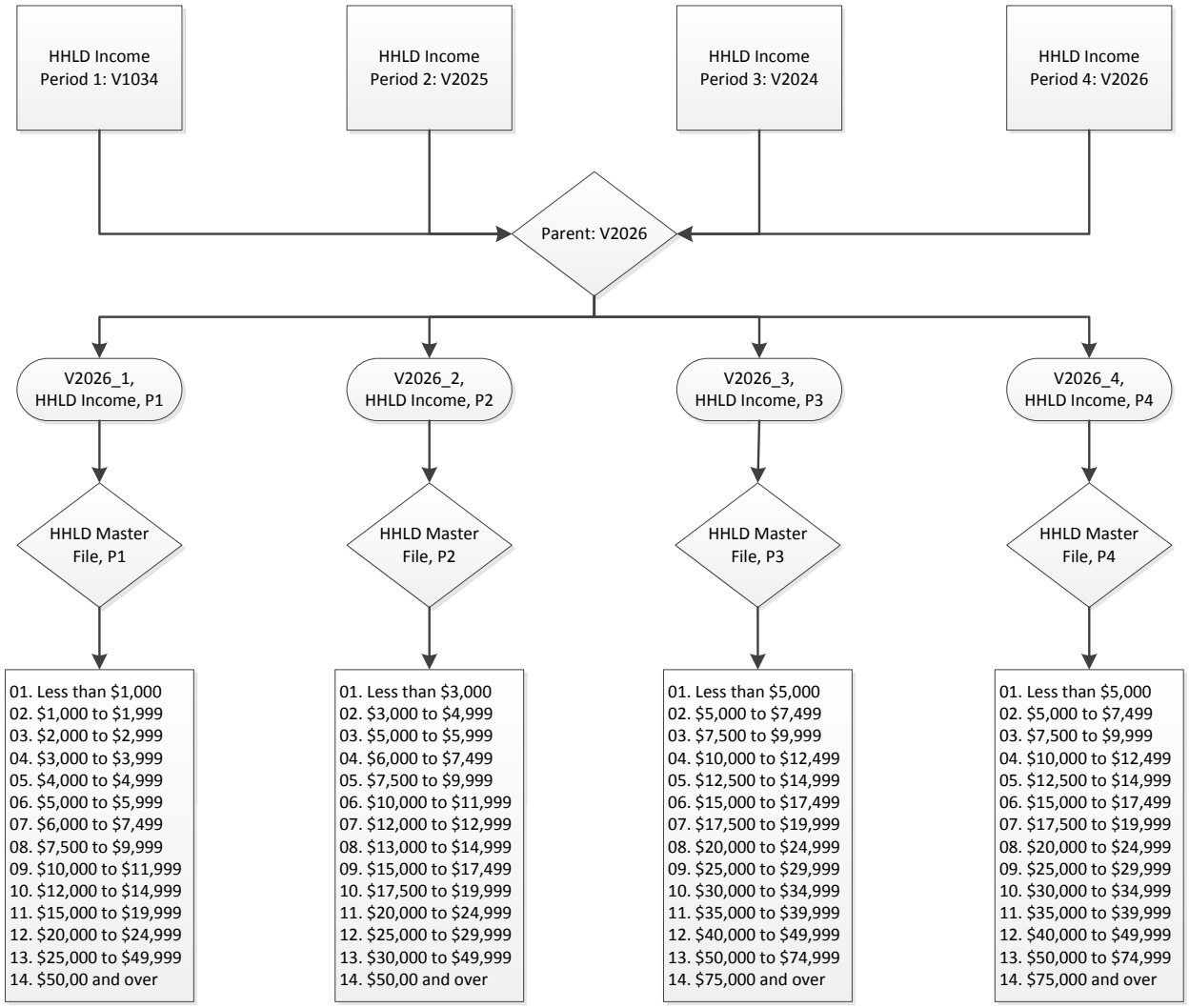
As shown in **Figure 5-1**, the household income variable from each period (V1034 in Period 1; V2025 in Period 2; V2024 in Period 3; and V2026 in Period 4) was mapped to the same parent variable with suffixes added to identify the period as prescribed by the taxonomy discussed previously. Once the variables were renamed (V2026_1, V2026_2, V2026_3, V2026_4), each variable was added to its respective household master file with the original level definitions intact. As an example, in Periods 1 and 2, Level 14 corresponds to a household income of "$50,000 and over" while in Periods 3 and 4, Level 14 corresponds to a household income of "$75,000 and over."

As with the yearly NCVS data files, the NHT master date files are maintained separately at the household, person, and incident levels. The primary differences between annual NCS/NCVS files and the NHT master data files are: (1) variable names were changed to reflect the parent-child taxonomy, (2) individual survey years were aggregated into four study periods, and (3) variables that did not meet the criteria described in section 1.1 were excluded from the files. The four study periods combined with the three file types (household, person, and incident) resulted in 12 master data files maintained for the NHT project.

For each master file, an accompanying codebook provides a guide to the unique identifiers used to link household, person, and incident files across years, the source and measurement of all "child" variables, or variables of interest as they were measured in each survey administration, and the construction of the "parent" variables that support trend analysis over time. The master files and codebook are available for public download and use through the Inter-University Consortium for Political and Social Research at www.icpsr.umich.edu.

From the master file, extract files were generated to facilitate accessibility and analytical use of the data as discussed in the following section.

Figure 5-1. Mapping of Household Income to a Common Parent Variable on the NHT Master Data Files

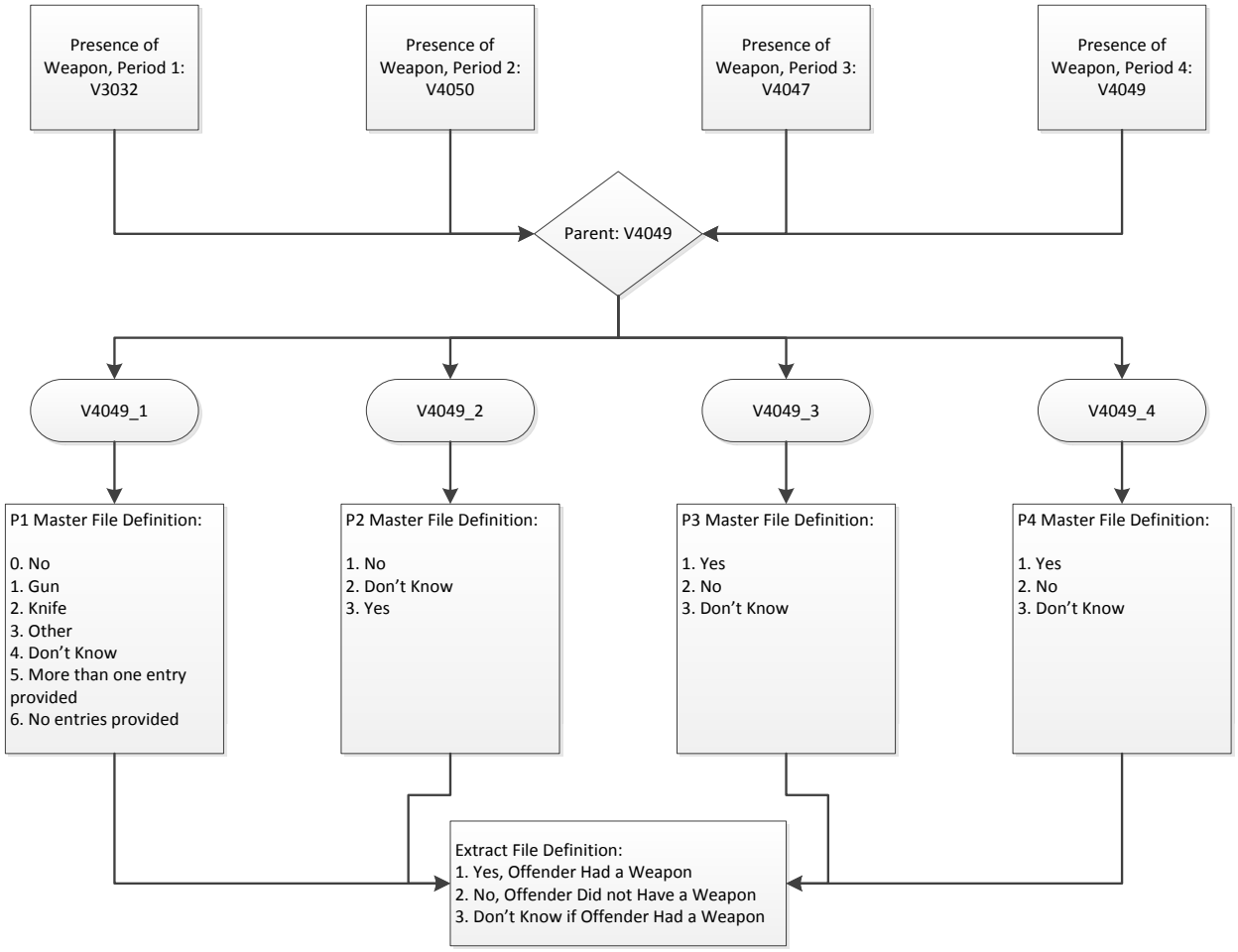


5.2 NHT Extract Data Files

Once the NHT master files were created, data extracts were also created that mapped the NHT child variables to a single set of response options that is common across the four study periods. This allowed for the comparison of estimates across the four periods and simplified the process of generating estimates from 1973 to 2014. For example, the presence of a weapon during an incident has varied response options and/or response ordering across the four study periods. These response options are left unchanged on the NHT master files. However, to

streamline analysis, a common set of response options, maintained within a single variable, is created on the NHT incident extract file as shown in **Figure 5-2**.

Figure 5-2. Mapping of Presence of a Weapon to a Common Set of Response Options on the NHT Incident Extract File



Once a common definition was created for each construct of interest, accounting for the major changes in the NCVS design as described in Chapter 2, estimates of victimization and characteristics of victimizations could be produced across the entire survey period, 1973–2014, using the NHT extract files. As with the NHT master files, separate extract files are created at the household, person, and incident levels with appropriate variables that can be used to link the three files. The variables required to link files, along with the common constructs, are detailed in an accompanying codebook. The extract files and codebook are available for public download

and use through the Inter-University Consortium for Political and Social Research at www.icpsr.umich.edu.

5.3 NHT Crosswalk

A core element of the NHT project is the crosswalk. The NHT crosswalk identifies and links key variables of interest across survey administrations from 1973 to 2014 and describes changes in measurement across that time period. The selected crosswalk variables include variables that allow analysts to evaluate variables over time.

The NHT crosswalk was designed to guide the development of the NHT master and extract files, and also serves as a reference document for ongoing NHT research and analysis. The crosswalk provides detailed variable information at the household, person, and incident levels. Within each level, the information is further organized by variable group, variable, values for those variables that can be matched over time, and alternative measurement and values for variables that cannot be matched over time.

The crosswalk's primary point of reference is the most recent measurement as found in the 2014 NCVS. Once the variables of interest were identified based on theoretical interest and general availability from 1973 to 2014, each variable from the 2014 administration was matched to exact or similar variables from previous administrations. Variable matches are defined as exact, near, or potential.

Exact matches are either measured the same over time, or variables can be constructed to create exact measurement matches from the available information. Near matches are measured in a similar manner, with generally small wording changes in the interview question or response categories. Near-match variables may involve a one-time change in the level of detail available over the study period, and may include exact matches for a subset of the time period the measures are available.

Potential matches often involve multiple changes in the level of detail available over time. Multiple variables may be matched to a single variable, and more than one method may be used to construct a matched variable.

The crosswalk provides variable-specific and more general information on measurement changes over time, when those changes occurred, corresponding measures information for exact variable matches over time, and recommended or potential corresponding measures or constructed variables for near or potential variable matches. Where exact or near-variable matches are available, the crosswalk indicates the study period that the trend is available (if less than the entire 1973–2014 period) and how the measurement has changed over that period, if at all. The notations used in the NHT crosswalk file are detailed in **Table 5-2**. Please see the crosswalk for specific examples of the notation used for the historical trends project. The crosswalk is available for public download and use from the BJS website at www.bjs.gov.

Table 5-2. Notations Used in NHT Crosswalk File

Crosswalk Notations	
Three crosswalk worksheets are used: household level, person level, and incident level. Each worksheet is outlined so that level— 1 denotes variable groups 2 denotes variables within each group 3 denotes values within each variable 4 is not often used, but denotes variables that offer an alternative measurement to the level 3 variable. The conclusion of each worksheet has notes related to variable changes over time, when those changes occurred within a study period. In general, changes in measurement are noted from study to study according to the coding scheme below, particularly in the incident-level worksheet.	
Symbols used in crosswalk worksheet	
/	Variable or value not found.
*	Exact match to value found.
^VALUE or VAR	One to many match. Whenever this symbol appears, the same value or variable is found in multiple places in the column.
VAR^VAR	"Or"—meaning that either value/ variable combination results in the specified measurement. When multiple VALUES apply to a specified measurement, those values are separated by commas.
#	Value is integer that may be coded into categories.
~	Fuzzy match; unsure if correct variable or value matched or even possible.
+	Question or measurement began in the relevant study or in the year noted at the bottom of the column if the measurement change occurred within the study period.

Continued on next page

Table 5-2. Notations Used in NHT Crosswalk File

	Symbols used in crosswalk worksheet (continued)
-	Question or measurement ended with the relevant study or in the year noted at the bottom of the column if the measurement change occurred within the study period.
=	Measurement not explicitly found, but may be computed from the variable indicated.
+-	Variable changed during study period; differences in measurement noted by semicolon.
Cells with any symbol other than * or / are generally commented to indicate why an exact match was not found.	
	Crosswalk filters for comparability over time
Match quality applies to the years when the trend is available.	
	Exact: The exact same measure was found, with direct value comparability over time.
	Similar: Same variable was found, but there may be measurement changes at the value level that are not directly comparable to the specified values.
	Potential: Variable is not directly comparable due to measurement changes at the variable level. Not sure if the correct variable was matched OR there are multiple options presented for measurement.
	Trend available from
	Indicates the earliest year the variable is available.
	? Indicates a break in the trend—variable is not found in the years following the ?
	'stopped in YYYY' indicates that the variable was only available until that year.
Detail available over time:	
	= Same level of detail available across all years.
	< More detail available over time.
	> Less detail available over time.
	>< Variable detail over time.
	NOTE: The level of detail only applies to the years when the trend is available.
The same filters are applied to all values of a variable—even if the specific value level of detail or match quality over time is different than what is noted at the variable level.	

SECTION 6. QUALITY CONTROL PROCEDURES

Throughout the creation of the NHT master data files and extract files, quality control procedures were implemented to reduce errors and identify any issues with data quality. The first check compared the distribution of each variable by survey year in each of the three NCS periods shown in **Table 5-1**. This check helped to identify coding issues and/or additional or missing levels for certain variables. For example, in 1977 and 1978, all incidents in which a weapon was present were also flagged as having a gun present. The cross-check of distributions identified this issue so that it could be addressed when creating the master files. The second check determined the levels present in the data for each variable by year in each of the three NCS periods. This check helped to identify changes in level definitions that were not always evident by the codebooks. The third check ensured that no errors were made when creating the parent-child taxonomy by comparing the distribution of each variable from the original annual data files with the distribution from the NHT master data files.

SECTION 7. REFERENCES

Ash, S., Hornick, D., McCune, J., Rottach, R., Sanduckchi, M., & Warner, T. (2008). *Variance methods of the Demographic Statistical Methods Division*. New York, NY: U.S. Census Bureau.

Couzens, G. L., Berzofsky, M. E., & Krebs, C. P. (2014). Analyzing potential mode effects in the National Crime Victimization Survey. In *Proceedings of the Joint Statistical Meetings, Survey Research Methods Section* (pp. 2915–2923). Alexandria, VA: American Statistical Association.

Eidson Penick, B. K., & Owens, M. E. B., III (Eds.). (1976). *Surveying Crime*. Washington, DC: National Academy of Sciences.

Kindermann, C., Lynch, J. P., & Cantor, D. (1997). *Effects of the redesign on victimization estimates* (NCJ 164381). Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics. Retrieved from https://www.bjs.gov/content/pub/pdf/ERVE.PDF

Rand, M. R., Lynch, J. P., & Cantor, D. (1997). *Criminal Victimization, 1973–95* (NCJ 163069). Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics. Retrieved from https://www.bjs.gov/content/pub/pdf/Cv73_95.pdf

Ringel, C. (1997). *Criminal victimization 1996: Changes 1995-96 with trends 1993-96* (NCJ 165812). Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics. Retrieved from https://bjs.gov/content/pub/pdf/cv96.pdf

Shook-Sa, B., Lee, P., & Berzofsky, M. (2015). *Assessing the coverage and reliability of subnational geographic identifiers in the NCVS public-use file* (NCJ 249467). Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics. Retrieved from https://www.bjs.gov/content/pub/pdf/acrsgincvspuf.pdf

U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration. (1997). *National household survey on drug abuse: Main findings 1995*. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service. Retrieved from https://catalog.hathitrust.org/Record/100569392

U.S. Department of Justice, Bureau of Justice Statistics. (2014). *National Crime Victimization Survey: Technical documentation* (NCJ 247252). Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics. Retrieved from https://www.bjs.gov/content/pub/pdf/ncvstd13.pdf

Whitaker, C. J. (1989). *The redesigned National Crime Survey: Selected new data* (NCJ 114746). Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics. Retrieved from https://www.ncjrs.gov/pdffiles1/Digitization/114746NCJRS.pdf

APPENDIX A: ANALYSIS OF SUBGROUP-SPECIFIC ADJUSTMENT AND USING ADDITIONAL YEARS OF DATA

In creating the initial adjustment, the Bureau of Justice Statistics (BJS) investigated the use of subgroup-specific ratio adjustments to more accurately integrate the historical National Crime Survey (NCS) and contemporary National Crime Victimization Survey (NCVS) data. Although crime rates for certain subgroups were found to be differentially affected by the 1992 redesign, the overall adjusted crime rates did not differ significantly from those created using the simpler crime-specific ratio adjustments.[3] As part of the NHT project, the feasibility and need for creating subgroup-specific ratio adjustments was again investigated by RTI and BJS. The use of subgroup-specific adjustment ratios could be beneficial to the goals of the NHT by ensuring that estimates from the NCS were comparable with estimates from the NCVS for more domains of interest rather than only at the crime-type level. In contrast to the previous approach, the approach implemented as part of the NHT involved the use of additional years of data beyond the split-sample overlap period. This method was motivated, in part, by the small sample sizes often observed for certain crime types (e.g., rape) and subgroups. To investigate this further, researchers undertook the following two steps:

1. Assess the impact on the crime-specific ratio adjustments of including additional years of data to address small sample sizes.
2. If crime-specific ratio adjustments are not significantly affected by adding additional years of data, evaluate the need for subgroup-specific ratio adjustments using the 1992 split-sample data augmented with additional years of data.

Address adjustments based on small sample sizes. Because crime victimization is a relatively rare event, the 1992 split sample used by BJS to generate adjustment factors can be characterized by small sample sizes that can affect the stability of some adjustment factors. To

[3]James Lynch and David Cantor, *Models for adjusting the NCS trends to account for design difference between the NCS and NCVS*, memorandum to the NCVS Sub-committee of the American Statistical Association Committee on Law and Justice Statistics, May 15, 1996.

address this issue in the NHT, the split-sample data from 1992 were augmented with additional years of data to increase the sample size and, potentially, to allow for more detailed adjustments.

The NHT examined adding an additional one year and 2 years of NCS and NCVS data to the data from the 1992 split-sample overlap period to increase sample size. For instance, the addition of a single year of data means that 1991 NCS data were combined with 1992 NCS data and 1993 NCVS data were combined with 1992 NCVS data, as shown in the formula below. In this way, the split sample data from 1992 were augmented with additional years of data, thereby increasing the sample size and potentially allowing for more detailed adjustments. A statistically significant difference in the ratios of the split-sample data ratio and the augmented split-sample data ratio would suggest that improvements could be made by incorporating additional years of data into adjustment factors.

$$ADJ_{1,c} = \frac{Rate_{c,1992-1993\ NCVS}}{Rate_{c,1991-1992\ NCS}}$$

where

ADJ1,c is the ratio adjustment factor with 1 additional year of data added for crime type *c*,

$Rate_{c,\ 1992–1993\ NCVS}$ is the rate of victimization for crime type *c* with the combined data from the 1992 and 1993 NCVS surveys, and

$Rate_{c,\ 1991–1992\ NCS}$ is the rate of victimization for crime type *c* with the combined data from the 1991 and 1992 NCS surveys.

Ratios calculated using an additional year of data (**Table A-1**) did not reveal statistically significant differences between the ratios from the augmented data and the ratios obtained using only the split-sample data for the detailed crime types (rape, robbery, aggravated assault, simple assault, burglary, theft, and motor vehicle theft). A small difference was detected for overall property crime between the ratio using only the split-sample data and the ratio using one additional year of data.

Table A-1. Difference Between Ratios Calculated From 1992 Split-Sample Data Only and Ratios Obtained From Adding One Additional Year of Data

Crime Type	Augmented Data Ratio (1992–1993 NCVS) (1991–1992 NCS)	Split-Sample Data Ratio (1992 NCVS) (1992 NCS)	Difference in Ratios
Violent Crimes	1.53	1.49	0.04
Rape	2.13	2.62	(0.49)
Robbery	1.03	1.02	0.01
Aggravated Assault	1.37	1.24	0.13
Simple Assault	1.76	1.75	0.00
Personal Theft	0.88	0.75	0.13
Property Crimes	1.18	1.23	(0.05)‡
Household Burglary	1.15	1.20	(0.05)
Household Theft	1.21	1.27	(0.06)
Motor Vehicle Theft	0.90	0.92	(0.02)

Note: Numbers in parentheses represent negative values.
‡ The difference in ratios between the augmented data and the split-sample only data was statistically significant at the 90% confidence level.

Ratios calculated using 2 years of additional data were calculated next by combining data from the 1992, 1993, and 1994 NCVS and the 1990, 1991, and 1992 NCS, as illustrated below. Once calculated, these ratios were compared to the ratios using only the split-sample data, as shown in **Table A-2**.

$$ADJ_{2,c} = \frac{Rate_{c,1992-1994\ NCVS}}{Rate_{c,1990-1992\ NCS}}$$

where

$ADJ_{2,c}$ is the ratio adjustment factor with 2 additional years of data added for crime type *c*;

$Rate_{c,\ 1992–1994\ NCVS}$ is the rate of victimization for crime type *c* with the combined data from the 1992, 1993, and 1994 NCVS surveys; and

$Rate_{c,\ 1990–1992\ NCS}$ is the rate of victimization for crime type *c* with the combined data from the 1990, 1991, and 1992 NCS surveys.

For the majority of crime types, the ratios calculated using the augmented data approach with two additional years of data did not differ significantly from the ratios based on the split-sample data only (**Table A-2**). Small differences were detected for overall property crime and household theft.

Table A-2. Difference Between Ratios Calculated From 1992 Split-Sample Data Only and Ratios Obtained From Adding Two Additional Years of Data

Crime Type	Augmented Data Ratio (1992–1994 NCVS) (1990–1992 NCS)	Split-Sample Data Ratio (1992 NCVS) (1992 NCS)	Difference in Ratios
Violent Crimes	1.58	1.49	0.09
Rape	2.18	2.62	(0.44)
Robbery	1.04	1.02	0.02
Aggravated Assault	1.40	1.24	0.16
Simple Assault	1.84	1.75	0.09
Personal Theft	0.82	0.75	0.07
Property Crimes	1.15	1.23	(0.08)†
Household Burglary	1.10	1.20	(0.09)
Household Theft	1.18	1.27	(0.09)†
Motor Vehicle Theft	0.89	0.92	(0.03)

Note: Numbers in parentheses represent negative values.
† The difference in ratios between the augmented data and the split-sample only data was statistically significant at the 95% confidence level.

As shown in **Tables A-1** and **A-2**, only minor differences were detected between the ratios based on the split-sample data only and the ratios calculated using additional years of data beyond the overlap period. This suggests that the temporal changes in crime rates observed during this time frame (Lauritsen, Heimer, & Lynch, 2009) are not significantly affecting the ability to measure changes caused by the redesigned survey. Although the potential for these two factors to be confounded is recognized, the decision was made to continue with the next step to evaluate the feasibility of producing subgroup-specific ratio adjustments.

Assess adjustments at the subgroup level. Once it had been established that the temporal changes in crime rates that could be introduced from including additional years of data beyond the split-sample overlap period did not significantly alter the relationship observed between crime-specific estimates as measured before and after the redesign, the next step was to

determine whether the data could support subgroup-specific ratio adjustments and whether such adjustments were necessary.

To examine this issue in more detail, the first step was to use the 1992 split-sample data only to calculate NCVS-NCS ratios for the major personal crime types by respondent characteristics, including sex (male, female); race (non-Hispanic white, non-Hispanic black, Hispanic, and non-Hispanic other); and age (12–17, 18–24, 25–34, 35–44, 45–64, and 65 or older). A significant difference in the ratios would suggest that subgroup-specific adjustments might improve upon the overall crime-specific adjustment factors by aligning NCS estimates before 1992 with those from the NCVS for more domains of interest by accounting for the differential impact of the redesign on certain groups.

Table A-3 presents the NCVS/NCS ratio for males and females from the 1992 split-sample data and the difference in observed ratios. Significant differences in the ratios for males and females were observed for the crime types of rape and simple assault at the 95% confidence level.

Table A-3. NCVS/NCS Subgroup-Specific Ratio Adjustments and Difference in Ratios for Personal Crime Types by Sex Using 1992 Split-Sample Data

Crime Type	Male Ratio	Female Ratio	Difference
Violent Crimes	1.53	1.44	0.09
Rape	0.26	4.28	(4.02)†
Robbery	1.00	1.06	(0.07)
Aggravated Assault	1.30	1.12	0.18
Simple Assault	1.95	1.52	0.43†
Personal Theft	0.84	0.69	0.15

Note: Numbers in parentheses represent negative values.
† The difference in ratios is statistically significant at the 95% confidence level.

Table A-4 presents the subgroup-specific ratio adjustments calculated from the 1992 split-sample data by race/ethnicity. The differences in ratios and significance of those differences are presented in **Table A-5**. Most of the differences in ratios were found to be not significant. However, differences were detected between non-Hispanic whites and non-Hispanic blacks ($p < 0.05$) and between non-Hispanic blacks and non-Hispanic others ($p < 0.10$) for overall violent crimes.

Table A-4. NCVS/NCS Subgroup-Specific Ratio Adjustments for Personal Crime Types by Race Using 1992 Split-Sample Data

Crime Type	NH White	NH Black	Hispanic	NH Other
Violent Crimes	1.60	1.08	1.44	2.16
Rape	2.67	2.04	2.52	N/A
Robbery	1.07	0.90	1.07	1.79
Aggravated Assault	1.34	0.83	1.44	2.13
Simple Assault	1.79	1.50	1.64	2.18
Personal Theft	0.88	0.59	0.51	1.05

Table A-5. Differences in Subgroup-Specific Ratio Adjustments for Personal Crime Types by Race Using 1992 Split-Sample Data

Crime Type	NH Black	Hispanic	NH Other
Violent Crimes			
NH White	0.51†	0.15	(0.56)
NH Black	--	(0.36)	(1.08)‡
Hispanic	--	--	(0.72)
Rape			
NH White	0.62	0.15	N/A
NH Black	--	(0.48)	N/A
Hispanic	--	--	N/A
Robbery			
NH White	0.18	0.01	(0.72)
NH Black	--	(0.17)	(0.89)
Hispanic	--	--	(0.72)
Aggravated Assault			
NH White	0.51	(0.10)	(0.79)
NH Black	--	(0.61)	(1.30)
Hispanic	--	--	(0.69)
Simple Assault			
NH White	0.29	0.15	(0.39)
NH Black	--	(0.14)	(0.68)
Hispanic	--	--	(0.54)
Personal Theft			
NH White	0.29	0.37	(0.17)
NH Black	--	0.08	(0.46)
Hispanic	--	--	(0.54)

Note: Numbers in parentheses represent negative values. NH, non-Hispanic.
† The difference in ratios is statistically significant at the 95% confidence level.
‡ The difference in ratios is statistically significant at the 90% confidence level.

Table A-6 presents the subgroup-specific ratio adjustments calculated from the 1992 split-sample data by age group. The differences in ratios and significance of those differences are presented in **Table A-7**. As with the other comparisons, most of the differences in ratios were found to be not significant. Significant differences were observed between the following age groups for overall violent crimes: 18–24 and 35–44 ($p < 0.05$); and 35–44 and 65 or older ($p < 0.05$). Differences in the ratios for simple assault were also observed between the following age categories: 12–17 and 18–24 ($p < 0.05$); 12–17 and 35–44 ($p < 0.10$); 12–17 and 65 or older ($p < 0.10$); 18–24 and 35–44 ($p < 0.05$); 25–34 and 35–44 ($p < 0.05$); 35–44 and 45–64 ($p < 0.05$); and 35–44 and 65 or older ($p < 0.05$). No significant differences in ratios between age groups were observed for the personal crime types of rape, robbery, aggravated assault, and personal theft.

Table A-6. NCVS/NCS Subgroup-Specific Ratio Adjustment for Personal Crime Types by Age Group Using 1992 Split-Sample Data

Crime Type	12–17	18–24	25–34	35–44	45–64	65+
Violent Crimes	1.54	1.26	1.51	1.81	1.48	1.06
Rape	1.81	2.79	3.34	3.32	2.20	0.54
Robbery	1.24	0.93	1.02	0.77	1.03	1.24
Aggravated Assault	1.07	1.12	1.36	1.17	1.96	1.09
Simple Assault	1.84	1.36	1.71	2.65	1.48	0.95
Personal Theft	1.21	0.65	0.41	0.58	1.03	0.81

Table A-7. Differences in Subgroup-Specific Ratio Adjustments for Personal Crime Types by Age Using 1992 Split-Sample Data

Crime Type	18–24	25–34	35–44	45–64	65+
Violent Crimes					
12–17	0.28	0.03	(0.27)	0.06	0.48
18–24	--	(0.25)	(0.56)†	(0.23)	0.19
25–34	--	--	(0.31)	0.02	0.44
35–44	--	--	--	0.33	0.75†
45–64	--	--	--	--	0.42

(continued)

Table A-7. Differences in Subgroup-Specific Ratio Adjustments for Personal Crime Types by Age Using 1992 Split-Sample Data (continued)

Crime Type	18–24	25–34	35–44	45–64	65+
Rape					
12–17	(0.98)	(1.53)	(1.51)	(0.39)	1.27
18–24	--	(0.55)	(0.53)	0.59	2.25
25–34	--	--	0.01	1.14	2.80
35–44	--	--	--	1.12	2.79
45–64	--	--	--	--	1.67
Robbery					
12–17	0.30	0.22	0.46	0.20	(0.01)
18–24	--	(0.09)	0.16	(0.10)	(0.31)
25–34	--	--	0.24	(0.02)	(0.23)
35–44	--	--	--	(0.26)	(0.47)
45–64	--	--	--	--	(0.21)
Aggravated Assault					
12–17	(0.05)	(0.30)	(0.11)	(0.90)	(0.03)
18–24	--	(0.24)	(0.05)	(0.84)	0.03
25–34	--	--	0.19	(0.60)	0.27
35–44	--	--	--	(0.79)	0.08
45–64	--	--	--	--	0.87
Simple Assault					
12–17	0.48‡	0.12	(0.81)‡	0.36	(0.89)‡
18–24	--	(0.35)	(1.29)†	(0.12)	0.41
25–34	--	--	(0.93)†	0.24	0.76
35–44	--	--	--	1.17†	1.69†
45–64	--	--	--	--	0.52
Personal Theft					
12–17	0.56	0.81	0.64	0.19	0.40
18–24	--	0.25	0.08	(0.37)	(0.16)
25–34	--	--	(0.17)	(0.62)	(0.41)
35–44	--	--	--	(0.45)	(0.24)
45–64	--	--	--	--	0.21

Note: Numbers in parentheses represent negative values.
† The difference in ratios is statistically significant at the 95% confidence level.
‡ The difference in ratios is statistically significant at the 90% confidence level.

The second step of the analysis was to compare subgroup-specific adjustments using the augmented data approach with one additional year of data (i.e., 1992–1992 NCVS / 1991–1992 NCS). For this analysis, the differences in ratios for personal crimes between subgroups defined by sex (male, female) and race (non-Hispanic white, non-Hispanic black, Hispanic, and non-Hispanic other) were analyzed.

Table A-8 presents the NCVS/NCS ratio for males and females using the 1992/1993 NCVS and 1991/1992 NCS and the difference in observed ratios. The only significant difference in ratios between males and females occurred for the crime type of rape at the 95% confidence level. No other differences between males and females were observed.

Table A-8. NCVS/NCS Subgroup-Specific Ratio Adjustments and Difference in Ratios for Personal Crime Types by Sex Using Augmented Data With One Additional Year Included, (1992 + 1993 NCVS) / (1991 + 1992 NCS)

Crime Type	Male Ratio	Female Ratio	Difference
Violent Crimes	1.48	1.59	(0.11)
Rape	0.45	2.70	(2.25)†
Robbery	1.01	1.07	(0.06)
Aggravated Assault	1.34	1.41	(0.07)
Simple Assault	1.78	1.72	0.07
Personal Theft	0.97	0.82	0.15

Note: Numbers in parentheses represent negative values.
† The difference in ratios is statistically significant at the 95% confidence level.

Table A-9 presents the subgroup-specific ratio adjustments calculated from the 1992/1992 NCVS and 1991/1992 NCS by race/ethnicity. The differences in ratios and significance of those differences are presented in **Table A-10**. Despite the larger sample sizes from adding additional years of data, the only significant difference in ratios between subgroups defined by race/ethnicity was found between non-Hispanic whites and non-Hispanic blacks for overall violent crimes.

Table A-9. NCVS/NCS Subgroup-Specific Ratio Adjustments for Personal Crime Types by Race Using Augmented Data with One Additional Year Included, (1992 + 1993 NCVS) / (1991 + 1992 NCS)

Crime Type	NH White	NH Black	Hispanic	NH Other
Violent Crimes	1.60	1.26	1.50	1.68
Rape	1.99	2.72	1.68	11.53
Robbery	1.08	0.92	1.03	1.32
Aggravated Assault	1.42	1.14	1.44	1.49
Simple Assault	1.78	1.55	1.86	1.86
Personal Theft	0.97	0.87	0.63	0.80

NH=non-Hispanic.

Table A-10. Differences in Subgroup-Specific Ratio Adjustments for Personal Crime Types by Race Using Augmented Data with One Additional Year Included, (1992 + 1993 NCVS) / (1991 + 1992 NCS)

Crime Type	NH Black	Hispanic	NH Other
Violent Crimes			
NH White	0.34†	0.10	(0.08)
NH Black	--	(0.24)	(0.42)
Hispanic	--	--	(0.18)
Rape			
NH White	(0.73)	0.31	(9.54)
NH Black	--	1.04	(8.81)
Hispanic	--	--	(9.85)
Robbery			
NH White	0.16	0.05	(0.25)
NH Black	--	(0.11)	(0.41)
Hispanic	--	--	(0.29)
Aggravated Assault			
NH White	0.27	(0.02)	(0.07)
NH Black	--	(0.29)	(0.35)
Hispanic	--	--	(0.06)
Simple Assault			
NH White	0.22	(0.09)	(0.08)
NH Black	--	(0.31)	(0.30)
Hispanic	--	--	0.01
Personal Theft			
NH White	0.10	0.34	0.17
NH Black	--	0.24	0.07
Hispanic	--	--	(0.17)

Note: Numbers in parentheses represent negative values. NH=non-Hispanic.
† The difference in ratios is statistically significant at the 95% confidence level.

Given the lack of differences in ratios observed between subgroups, and the added complexity of adjusting NCS estimates at the subgroup level, the decision was made to use an overall crime-specific ratio adjustment as described in Chapter 3.

APPENDIX B: SUPPORTING TABLES FOR NHT APPROACH TO INTEGRATING NATIONAL CRIME SURVEY (NCS) AND NATIONAL CRIME VICTIMIZATION (NCVS) ESTIMATES

Table B-1. Comparison of Data-Year and Collection-Year Estimates and Ratios, 1992

	Data-Year Estimates			Collection-Year Estimates		
Crime Type	NCS Crime Rate[1]	NCVS Crime Rate[1]	NCVS/ NCS Ratio	NCS Crime Rate[1]	NCVS Crime Rate[1]	NCVS/ NCS Ratio
Violent Crimes	32.1	47.8	1.49†	31.8	50.0	1.58†
Rape[2]	0.7	1.8	2.62†	0.7	1.9	2.59†
Robbery	5.9	6.1	1.02	5.8	6.2	1.08
Aggravated Assault	9.0	11.1	1.24†	8.7	11.1	1.28†
Simple Assault	16.5	28.9	1.75†	16.6	30.8	1.86†
Personal Theft	2.3	1.8	0.75‡	2.5	1.6	0.62†
Property Crimes	264.5	325.3	1.23†	269.7	335.4	1.24†
Household Burg	48.9	58.6	1.20†	48.3	61.0	1.26†
Household Theft	195.5	248.2	1.27†	199.8	256.2	1.28†
Motor Vehicle Theft	20.1	18.5	0.92	21.5	18.1	0.84†

[1]Rate of victimization per 1,000 persons or households.
[2]Estimates of rape exclude sexual assaults.
† The ratio of the NCVS to NCS estimates was statistically different from 1 at the 95% confidence level.
‡ The ratio of the NCVS to NCS estimates was statistically different from 1 at the 90% confidence level.

Table B-2. Comparison of Including Series Crimes with NCS Missing Values Set to 3 or 6, Data-Year Estimates, 1992

Crime Type	NCVS/NCS Ratio: NCS Missing Set to 3	NCVS/NCS Ratio: NCS Missing Set to 6	Difference in Ratios
Violent Crimes	1.544	1.511	0.032
Rape[1]	1.857	1.857	0.000
Robbery	0.935	0.915	0.019
Aggravated Assault	1.274	1.257	0.017
Simple Assault	1.828	1.782	0.046
Personal Theft	0.707	0.707	0.000
Property Crimes	1.189	1.179	0.010
Household Burg	1.124	1.115	0.009
Household Theft	1.234	1.223	0.011
Motor Vehicle Theft	0.894	0.889	0.005

[1]Estimates of rape exclude sexual assaults.

Table B-3. Impact of Including Series Crimes on Ratio Adjustments Using Collection Year Data, 1992

Crime Type	Series Excluded			Series Included		
	NCS Crime Rate[1]	NCVS Crime Rate[1]	NCVS/ NCS Ratio	NCS Crime Rate[1]	NCVS Crime Rate[1]	NCVS/ NCS Ratio
Violent Crimes	31.8	50.0	1.58†	50.4	77.7	1.54†
Rape[2]	0.7	1.9	2.59†	1.6	2.4	1.53‡
Robbery	5.8	6.2	1.08	7.9	6.8	0.86‡
Aggravated Assault	8.7	11.1	1.28†	12.7	16.1	1.27†
Simple Assault	16.6	30.8	1.86†	28.2	52.4	1.86†
Personal Theft	2.5	1.6	0.62†	2.7	1.6	0.59†
Property Crimes	269.7	335.4	1.24†	307.2	371.5	1.21†
Household Burg	48.3	61.0	1.26†	57.3	66.0	1.15†
Household Theft	199.8	256.2	1.28†	227.4	286.8	1.26†
Motor Vehicle Theft	21.5	18.1	0.84†	22.5	18.6	0.83†

[1]Rate of victimization per 1,000 persons or households.
[2]Estimates of rape exclude sexual assaults.
† The ratio of the NCVS to NCS estimates was statistically different from 1 at the 95% confidence level.
‡ The ratio of the NCVS to NCS estimates was statistically different from 1 at the 90% confidence level.

Table B-4. Relationship Between the Rates of Rape and Sexual Assault in the NCVS, 1992–2014

Year	Rape		Sexual Assault		Rape/Sexual Assault
	Rate[1]	% of RSA	Rate[1]	% of RSA	Rate[1]
1992	2.39	58.77	1.68	41.23	4.07
1993	2.69	63.26	1.56	36.74	4.26
1994	2.12	66.93	1.05	33.07	3.16
1995	1.70	64.80	0.92	35.20	2.62
1996	1.06	52.49	0.96	47.51	2.01
1997	1.51	59.84	1.01	40.16	2.52
1998	0.98	55.42	0.79	44.58	1.76
1999	1.61	61.20	1.02	38.80	2.63
2000	0.89	54.87	0.73	45.13	1.62
2001	0.99	47.77	1.09	52.23	2.08
2002	0.83	54.74	0.68	45.26	1.51
2003	0.81	59.62	0.55	40.38	1.36
2004	0.61	57.43	0.45	42.57	1.06
2005	0.60	70.50	0.25	29.50	0.85
2006	1.08	57.57	0.80	42.43	1.88
2007	0.56	56.64	0.43	43.36	0.99
2008	1.07	76.89	0.32	23.11	1.39
2009	0.68	56.86	0.52	43.14	1.20
2010	0.65	62.32	0.40	37.68	1.05
2011	0.62	65.00	0.33	35.00	0.95
2012	0.80	60.13	0.53	39.87	1.32
2013	0.76	66.86	0.38	33.14	1.14
2014	0.50	47.07	0.56	52.93	1.07

[1]Rate of victimization per 1,000 persons.

Table B-5. Impact of Including Sexual Assault on Ratio Adjustments Using Collection Year Data, 1992

Crime Type	Sexual Assault Excluded			Sexual Assault Included		
	NCS Crime Rate[1,2]	NCVS Crime Rate[1]	NCVS/ NCS Ratio	NCS Crime Rate[1,2]	NCVS Crime Rate[1]	NCVS/ NCS Ratio
Rape/Sexual Assault	1.6	2.4	1.53‡	1.6	4.1	2.61†

[1]Rate of victimization per 1,000 persons.
[2]The NCS portion of the 1992 split-sample did not include questions about sexual assault. Consequently, the NCS estimates for "sexual assault included" are the same as those for "sexual assault excluded."
† The ratio of the NCVS to NCS estimates was statistically different from 1 at the 95% confidence level.

Table B-6. Distribution of 1992 Split-Sample by Household and Person Characteristics, Collection-Year Data

Geographic/Person Characteristic	NCS			NCVS		
	Number of Respondents	% Unweighted	% Weighted	Number of Respondents	% Unweighted	% Weighted
All Persons	107,073	100	100	94,985	100	100
All Households	53,767	100	100	48,882	100	100
Location of Residence						
Urban	18,679	34.7	33.9	13,744	28.1	29.2†
Suburban	23,240	43.2	43.5	23,775	48.6	48.4†
Rural	11,848	22.0	22.6	11,363	23.3	22.4
Population Size						
Not in a Place	15,819	29.4	29.9	14,678	30.0	30.3
Under 1,000	882	1.6	1.7	744	1.5	1.5
1,000–9,999	6,678	12.4	12.7	6,910	14.1	14.1†
10,000–24,999	6,455	12.0	12.1	5,911	12.1	11.8
25,000–49,999	4,772	8.9	8.9	5,211	10.7	10.1†
50,000–99,999	3,821	7.1	6.8	5,080	10.4	9.8†
100,000–249,999	4,371	8.1	7.7	4,024	8.2	8.9†
250,000–499,999	3,403	6.3	6.4	2,180	4.5	4.9†
500,000–999,999	2,250	4.2	4.3	2,587	5.3	5.2†
1,000,000 or more	5,316	9.9	9.5	1,557	3.2	3.6†
Household Income						
Less than $25,000	23,277	43.3	43.9	21,082	43.1	43.6
$25,000–$49,999	15,148	28.2	28.1	14,104	28.9	28.6
$50,000–$74,999	5,227	9.7	9.5	5,162	10.6	10.4†
$75,000 or greater	3,221	6.0	5.8	3,200	6.6	6.4
Unknown	6,894	12.8	12.6	5,334	10.9	11.0†
Household Tenure						
Own	34,916	64.9	64.0	31,904	65.3	64.1
Rent/No Cash Rent	18,851	35.1	36.0	16,978	34.7	35.9
Sex						
Male	49,712	46.4	48.1	43,900	46.2	48.4
Female	57,361	53.6	51.9	51,085	53.8	51.6
Age						
12–17	10,750	10.0	10.1	9,346	9.8	10.2
18–24	11,447	10.7	12.2	10,398	11.0	12.5
25–34	21,355	19.9	20.3	19,058	20.1	20.6
35–49	28,790	26.9	26.6	25,441	26.8	26.6
50–64	17,858	16.7	16.0	15,879	16.7	15.6
65 or older	16,873	15.8	14.8	14,863	15.7	14.5

(continued)

Table B-6. Distribution of 1992 Split-Sample by Household and Person Characteristics, Collection-Year Data (continued)

Geographic/Person Characteristic	NCS Number of Respondents	NCS % Unweighted	NCS % Weighted	NCVS Number of Respondents	NCVS % Unweighted	NCVS % Weighted
Race/Ethnicity						
NH White	83,725	78.2	76.9	75,603	79.6	77.7
NH Black	10,739	10.0	11.4	9,768	10.3	11.7
Hispanic	9,108	8.5	8.5	6,723	7.1	7.5†
NH Asian	3,137	2.9	2.9	2,283	2.4	2.5†
NH Other	364	0.3	0.3	608	0.6	0.6†
Education						
Less than High School	27,811	26.2	26.1	23,884	25.4	25.6
High School/GED	36,674	34.5	34.7	32,153	34.3	34.2
Some College	20,375	19.2	19.4	18,615	19.8	20.1
Bachelor's or Higher	21,320	20.1	19.8	19,216	20.5	20.1
Marital Status						
Never Married	30,239	28.3	30.0	26,527	28.0	30.2
Married	58,603	54.8	53.6	52,077	54.9	53.3
Widowed	7,421	6.9	6.6	6,532	6.9	6.4
Divorced	8,188	7.7	7.6	7,476	7.9	7.8
Separated	2,420	2.3	2.3	2,218	2.3	2.4

NH=Non-Hispanic.

† The difference in weighted percentages between the NCS and NCVS is statistically significant at the 95% confidence level.

Table B-7. Impact of Post-Stratifying NCS Estimates to NCVS Control Totals on Adjustment Ratios, Collection Year Estimates, Series Crimes Included, 1992

	Unadjusted			Post-Stratified		
Crime Type	NCS Crime Rate[1]	NCVS Crime Rate[1]	NCVS/ NCS Ratio	NCS Crime Rate[1]	NCVS Crime Rate[1]	NCVS/ NCS Ratio
Violent Crimes	50.4	77.7	1.54	51.3	77.7	1.51
Rape[2]	1.6	2.4	1.53	1.6	2.4	1.51
Robbery	7.9	6.8	0.86	8.0	6.8	0.85
Aggravated Assault	12.7	16.1	1.27	12.9	16.1	1.25
Simple Assault	28.2	52.4	1.86	28.8	52.4	1.82
Personal Theft	2.7	1.6	0.59	2.7	1.6	0.59
Property Crimes	307.2	371.5	1.21	306.3	371.5	1.21
Household Burg	57.3	66.0	1.15	57.2	66.0	1.15
Household Theft	227.4	286.8	1.26	228.0	286.8	1.26
Motor Vehicle Theft	22.5	18.6	0.83	21.1	18.6	0.88

[1]Rate of victimization per 1,000 persons or households.
[2]Estimates of rape exclude sexual assaults.
Note: The NCS weights from the split-sample were post-stratified to match NCVS control totals from the split-sample. Consequently, the unadjusted and post-stratified NCVS estimates are the same.

APPENDIX C: GENERALIZED VARIANCE FUNCTIONS (GVF) PARAMETERS

Table C-1. GVF Parameters for Personal Crimes

	Overall Estimate Parameters			Domain Estimate Parameters		
Survey Year	**a**	**b**	**c**	**a**	**b**	**c**
1973	-0.000012104	1,821	N/A	-0.000012104	1,821	N/A
1974	-0.000011016	1,821	N/A	-0.000011016	1,821	N/A
1975	-0.000010847	1,821	N/A	-0.000010847	1,821	N/A
1976	-0.000010689	1,821	N/A	-0.000010689	1,821	N/A
1977	-0.000010406	1,821	N/A	-0.000010406	1,821	N/A
1978	-0.000014845	2,616	N/A	-0.000014845	2,616	N/A
1979	-0.000014197	2,509	N/A	-0.000014197	2,509	N/A
1980	-0.000013179	2,355	N/A	-0.000013179	2,355	N/A
1981	-0.000012748	2,355	N/A	-0.000012748	2,355	N/A
1982	-0.000012567	2,355	N/A	-0.000012567	2,355	N/A
1983	-0.00001236	2,355	N/A	-0.00001236	2,355	N/A
1984	-0.000015024	2,884	N/A	-0.000015024	2,884	N/A
1985	-0.000015534	3,015	N/A	-0.000015534	3,015	N/A
1986	-0.000015814	3,102	N/A	-0.000015814	3,102	N/A
1987	-0.00001835	3,629	N/A	-0.00001835	3,629	N/A
1988	-0.000015952	3,181	N/A	-0.000015952	3,181	N/A
1989	-0.00001752	3,528	N/A	-0.00001752	3,528	N/A
1990	-0.000018325	3,725	N/A	-0.000018325	3,725	N/A
1991	-0.00002309	4,714	N/A	-0.00002309	4,714	N/A
1992	-0.0000858	4,157	3.337	-0.00012979	4,716	3.735
1993	-0.00042893	6,085	5.818	-0.00073642	3,923	9.934
1994	-0.00040854	2,864	5.768	-0.00051953	2,838	7.021
1995	-0.00037079	2,826	5.245	-0.00058524	2,084	7.992
1996	-0.00054329	1,224	7.924	-0.00071706	2,141	9.853
1997	-0.00060876	3,973	8.758	-0.000844	3,084	11.606
1998	-0.00069365	716	10.284	-0.00083876	1,593	11.674
1999	-0.00062265	2,974	9.132	-0.00085331	2,881	11.837
2000	-0.00063445	1,650	9.445	-0.00091124	2,138	12.745
2001	-0.00052825	1,928	7.87	-0.00069129	2,774	9.633
2002	-0.00085677	2,087	12.901	-0.00108486	1,822	15.331
2003	-0.00051622	3,646	7.75	-0.00085819	2,593	12.28
2004	-0.00035852	4,395	5.291	-0.00060837	3,104	8.659
2005	-0.00060311	2,805	9.251	-0.00081369	3,213	11.713
2006	-0.00056113	5,459	8.476	-0.00073569	3,217	10.626
2007	-0.00059713	1,795	9.335	-0.00107591	264	15.939
2008	-0.00063195	5,004	9.722	-0.00076651	5,025	11.857
2009	-0.00054111	2,410	8.475	-0.0010674	3,529	16.794
2010	-0.00086145	1,384	13.696	-0.0010876	1,960	17.278
2011	-0.00060211	2,439	9.511	-0.00081383	2,309	12.916
2012	-0.00025343	6,439	3.704	-0.00054187	3,893	8.53
2013	-0.00082347	2,059	13.264	-0.001118	1,958	18.059
2014	-0.00050049	4,230	7.914	-0.00077074	3,234	12.388

Table C-2. GVF Parameters for Property Crimes

Survey Year	Overall Estimate Parameters			Domain Estimate Parameters		
	a	b	c	a	b	c
1973	-0.000026454	1,821	N/A	-0.000026454	1,821	N/A
1974	-0.000025809	1,821	N/A	-0.000025809	1,821	N/A
1975	-0.000025318	1,821	N/A	-0.000025318	1,821	N/A
1976	-0.000010689	1,821	N/A	-0.000010689	1,821	N/A
1977	-0.000010405	1,821	N/A	-0.000010405	1,821	N/A
1978	-0.000014845	2,616	N/A	-0.000014845	2,616	N/A
1979	-0.000014197	2,509	N/A	-0.000014197	2,509	N/A
1980	-0.000013179	2,355	N/A	-0.000013179	2,355	N/A
1981	-0.000012748	2,355	N/A	-0.000012748	2,355	N/A
1982	-0.000012567	2,355	N/A	-0.000012567	2,355	N/A
1983	-0.00001236	2,355	N/A	-0.00001236	2,355	N/A
1984	-0.000015024	2,884	N/A	-0.000015024	2,884	N/A
1985	-0.000015534	3,015	N/A	-0.000015534	3,015	N/A
1986	-0.000015814	3,102	N/A	-0.000015814	3,102	N/A
1987	-0.00001835	3,629	N/A	-0.00001835	3,629	N/A
1988	-0.000015952	3,181	N/A	-0.000015952	3,181	N/A
1989	-0.00001752	3,528	N/A	-0.00001752	3,528	N/A
1990	-0.000018325	3,725	N/A	-0.000018325	3,725	N/A
1991	-0.00002309	4,717	N/A	-0.00002309	4,717	N/A
1992	-0.00018414	3,108	2.642	-0.00010957	4,524	2.387
1993	-0.00034948	838	3.406	-0.00033613	2,011	3.019
1994	-0.0003808	319	3.787	-0.00033372	2,044	3.004
1995	-0.00035466	332	3.54	-0.00033252	2,002	3.004
1996	-0.00022338	2,041	2.062	-0.00025538	2,528	2.216
1997	-0.00050902	-526	5.242	-0.00042269	1,609	3.955
1998	-0.00045078	1,647	4.466	-0.00039879	2,640	3.639
1999	-0.00041482	960	4.201	-0.000368	2,455	3.383
2000	-0.00073678	-1,921	7.854	-0.00061236	827	5.981
2001	-0.00059233	-793	6.276	-0.00052248	1,263	5.07
2002	-0.0004039	1,053	4.142	-0.00037459	2,209	3.507
2003	-0.00048584	-197	5.209	-0.00047387	1,566	4.621
2004	-0.00033261	2,145	3.379	-0.00031991	2,552	2.995
2005	-0.00036754	2,039	3.789	-0.00036501	2,698	3.46
2006	-0.00025057	3,500	2.398	-0.00023644	3,800	2.052
2007	-0.0002774	2,364	2.816	-0.00028978	2,565	2.732
2008	-0.00035885	1,887	3.778	-0.00032519	3,348	3.275
2009	-0.00030998	2,160	3.233	-0.00026607	3,098	2.663
2010	-0.00020353	2,450	2.035	-0.00017899	3,297	1.687
2011	-0.00028723	4,182	2.809	-0.00037148	2,981	3.852
2012	-0.00020722	2,345	2.116	-0.00020565	2,810	2.057
2013	-0.00023036	2,084	2.418	-0.00021207	2,786	2.149
2014	-0.00019138	2,809	1.931	-0.00019976	3,071	2.003

APPENDIX D: IMPACT OF NHT ADJUSTMENT METHOD ON STANDARD ERRORS

Table D-1. Standard Errors of Estimates Obtained When Applying Adjustment Factors to Generalized Variance Functions Variance Estimates Compared With Applying Adjustment Factors to Weights

	Standard Error																		
	1973	**1974**	**1975**	**1976**	**1977**	**1978**	**1979**	**1980**	**1981**	**1982**	**1983**	**1984**	**1985**	**1986**	**1987**	**1988**	**1989**	**1990**	**1991**
Adjustment Factors Applied after GVF																			
Violent Victimization																			
Male	2.0	1.9	1.9	1.9	1.9	2.4	2.2	2.1	2.1	2.1	2.0	2.2	2.0	2.0	2.2	2.1	2.2	2.2	2.6
Female	1.4	1.3	1.3	1.4	1.4	1.7	1.6	1.4	1.5	1.5	1.5	1.6	1.7	1.7	1.8	1.7	1.8	1.7	2.0
Serious Violent Victimization																			
Male	1.0	1.0	1.0	0.9	0.9	1.1	1.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.1	1.0	1.1	1.1	1.3
Female	0.7	0.7	0.6	0.7	0.6	0.8	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.8	0.8	0.8	0.8	0.7	0.9
Simple Assault																			
Male	1.7	1.6	1.7	1.7	1.7	2.1	1.9	1.8	1.8	1.8	1.8	2.0	1.8	1.8	1.9	1.9	1.9	1.9	2.3
Female	1.2	1.1	1.1	1.2	1.2	1.5	1.4	1.3	1.3	1.4	1.3	1.4	1.5	1.5	1.6	1.5	1.6	1.6	1.8
Adjustment Factors Applied to Weights																			
Violent Victimization																			
Male	1.6	1.5	1.5	1.5	1.5	1.8	1.7	1.6	1.6	1.6	1.6	1.7	1.6	1.6	1.7	1.6	1.7	1.8	2.1
Female	1.1	1.0	1.0	1.1	1.1	1.3	1.3	1.1	1.2	1.2	1.1	1.2	1.3	1.3	1.4	1.3	1.4	1.3	1.6
Serious Violent Victimization																			
Male	1.0	0.9	0.9	0.9	0.8	1.0	1.0	0.9	1.0	0.9	0.9	1.0	0.9	0.9	1.0	0.9	1.0	1.0	1.2
Female	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.6	0.7	0.7	0.6	0.7	0.7	0.7	0.7	0.7	0.8	0.7	0.8
Simple Assault																			
Male	1.2	1.1	1.2	1.2	1.2	1.5	1.4	1.3	1.3	1.3	1.3	1.4	1.3	1.3	1.4	1.3	1.4	1.4	1.6
Female	0.9	0.8	0.8	0.9	0.9	1.1	1.1	0.9	0.9	1.0	1.0	1.0	1.1	1.1	1.2	1.1	1.1	1.2	1.3