

II. Method

The analysis for this report was conducted using National Incident-Based Reporting System (NIBRS) data for Vermont for 1993-2000. The Vermont Center for Justice Research maintains a NIBRS database derived from submissions to the Federal Bureau of Investigation (FBI) by the Vermont Department of Public Safety. The NIBRS database contains most incidents reportable to the FBI, although not all Vermont law enforcement agencies are on the Department of Public Safety law enforcement records management system known as the Vermont Incident-Based Reporting System, or VIBRS.²

Examination of the relationship between race and various characteristics of crime, including the likelihood of subsequent arrest, is complicated by the fact that Vermont has historically had a relatively small non-white population. U.S. census figures for 1990 indicate that, of Vermont's 562,758 population, only 7,670 were non-white, or 1.4%. Figures from the 2000 U.S. census indicate that Vermont experienced about a 30% growth in the non-white population with 12,284 of 608,827 residents, or 2.0% falling into this group. Moreover, the nature of the minority population shifted in Vermont during the 1990s such that by 2000 Asians comprised the largest single minority (5,358 residents) compared to Blacks (3,063 residents), which were largest single minority in 1990 (see U.S. Census Bureau, 2002).

The impact of this racial distribution on an analysis of crime incidents is that a relatively large number of incidents is required to obtain a sufficient number of incidents involving non-white offenders. In the present study, a total of 101,851 incidents were used, from which 33,119

²The NIBRS database maintained by the Vermont Center for Justice Research has been compiled in several stages. Data from 1993 through October 1996 were added to the database from Department of Public Safety monthly submissions to the FBI NIBRS program. Data for November and December 1996 were not available during the transition to a new reporting system at the Department of Public Safety. Since 1997 annual data have been obtained from the FBI and processed by the VCJR for analysis.

Additionally, the VIBRS network began operation in late 1992 with the Vermont State Police and a limited number of municipal agencies. It has expanded to include additional municipal and sheriff's departments, most notably Burlington, Vermont's largest agency, joined the network in 1998. Approximately 94% of the state's population is covered by an agency that reports crime through the VIBRS network. The Vermont Department of Public Safety has issued requirements that all departments not on the VIBRS network submit NIBRS format data beginning January 1, 2004.

offenders were identified from the Segment 5³ record for use in the analysis. Also captured by the database and related to the 33,119 offenders were 30,401 Segment 2 offenses, 19,480 Segment 3 property/drug records, 28,535 Segment 4 victims, and 20,005 Segment 6 arrestees.⁴

The methods employed for this study were relatively straightforward and derived from the major analytical objectives of the research and structure of the NIBRS data. Specifically, the two objectives were to examine descriptive differences in the characteristics of crime reported to NIBRS by race and, through non-linear regression analysis determine the degree to which race contributes to the likelihood of arrest. The analysis is offender-based in that it seeks to examine differences in offending and arrest rates for specific offenders. Therefore, the Segment 5 offender file was used as the base file, with individual offenders the unit of analysis in the descriptive and logistical regression analyses presented below.

Five major stages of data manipulation were required to construct the base file used in the analysis. Each of these steps is briefly described below.

Step 1: Build Segment Files for 1993-2000 from NIBRS Data. NIBRS data were obtained for the most recent year available at the time of the analysis (2000) and incorporated into the existing NIBRS files for earlier years. When combining data from different years with a NIBRS annual datafile, caution must be exercised to eliminate duplicate records created from update submissions to records in years subsequent to the original analysis. Only the most recent set of segment records should be used for the analysis in compliance with the FBI record update procedures.⁵

³This report references the segment structure and data format of the NIBRS system. Readers unfamiliar with the NIBRS system and layout should consult the Federal Bureau of Investigation description of the system and data structure contained in Volume 1 of the NIBRS system guidelines (See FBI, 1998 and footnote 4).

⁴NIBRS data are collected in a series of segment files, each with a different unit of analysis, included are: Segment 1 (incident characteristics), Segment 2 (offenses), Segment 3 (property/drugs), Segment 4 (victims), Segment 5 (offenders), Segment 6 (people arrested) and Segment 7 (arrests for minor offenses). For a description of the NIBRS program, data structure, coding and field specifications, see the Federal Bureau of Investigation publication Uniform Crime Reporting, National Incident-Based Reporting Program, Volume I: Data Collection Guidelines (February 1998).

⁵Duplicate records account for about 0.5% of records on an annual data tape and depend in part on when the data file was created. For example, a burglary may have occurred in November and have an original data submission for Segments 1, 2, 3 and 4; if no offender was known and an arrest did not result, then there would be no Segment 5 or 6 submissions.

Step 2: Create a basefile for the analysis. The Segment 5 offender record was used as the flat basefile from which the analysis was conducted, given that the unit of analysis was the offender and what happened to him or her. The Segment 5 record contains basic information about the offender (age, gender and race) central to the analysis. Additional variables from the other segment files were prepared and appended to the Segment 5 basefile.

Step 3: Extract relevant records from Segments 1-4 and 6. Several variables from other segments required for the analysis were extracted and appended to the offender segment. Several decisions regarding which variables were extracted and how they were appended are described below.

Step 4: Construct Summary Variables. A series of summary variables were constructed from Segments 2-6 for use in the logistical regression analysis. Most important, the number of records in each segment was computed through the aggregation procedure in SPSS in order to identify the number of offenses, property/drug records, victims, offenders and arrests that were tied to each incident. Additional summary variables were constructed and are described in more detail below.

Step 5: Match and Append Segment Records to Basefile. The final step was to append the relevant segment variables and construct variables to the Segment 5 basefile. Matching was done primarily using incident number and ORI (the unique law enforcement agency identifier assigned by the FBI) to create a unique record for each offender which captured the major characteristics of the offense in which they were involved, including whether or not they were arrested.

The resulting flat basefile structure is schematically depicted in [Figure 1](#). The figure does not represent every variable that was created for data manipulation and may have ended up in the file. It does, however, provide a visual representation of how a flat file was built from segment files using the offender segment as the basefile. A number of issues were encountered during the file building process, some of which are briefly discussed here, beginning with Segment 5.

However, an offender may have been identified and arrested in April or later of the following calendar year. A NIBRS datafile constructed early in the following year from the previous year's data would not have Segment 5 or 6 records, or any updates to the property record in Segment 3 to reflect recovered property. However, a calendar year datafile for the year in which the offender was identified and arrested would include the complete set of updated Segment 1 through 6 records. Duplicate records of the incident will result if the two separate calendar year files are simply merged.

Figure 1: Basefile Structure for Offender-Based File

Segment 5 Elements 1 & 2	Segment 5 Elements 36-39	Count Segment 5 Records	Segment 2 Elements 6-13 (most serious offense)	Count Segment 2 Records	Segment 3 Elements 14-22	Count of Segment 3 Records	Total Property Value
Segment 3 Drug Type	Segment 4 Elements 23-35	Count of Segment 4	Number Victims Under 18 Years Old	Number Minority Victims	Number Hispanic Victims	Segment 6 Elements 40-52 (matched to offender)	Count of Segment 6

Segment 5. Perhaps the most important issue surrounding construction of the base file was to identify all offenders that were part of the same incident. Each offender served as a unit of analysis and thus created a unique record based on ORI, incident number and offender number. The number of offenders in each incident were counted and this total was attached to each offender record for later use in the analysis.

Segment 2. The most important decision regarding Segment 2 data centered around which offense to use in cases where more than one offense occurred in an incident. For purposes of this analysis an offense seriousness ranking was created and each Segment 2 record was assigned the relevant seriousness score. The offense with the highest seriousness ranking was appended to the offender records for that incident. A count of the total number of offenses was also computed and attached to each record in the basefile for use in regression analysis. Multiple offenses occurred in only 5% of the incidents, which suggests that multiple offending is relatively rare and was not a significant problem for this analysis. Table 1 below presents the offense seriousness rankings used in the study.

Segment 3. Segment 3 presented a series of unique problems in that it combines both drug (seized) and property characteristics. In most instances, only one Segment 3 record existed and this was appended to the basefile. However, in instances of multiple segment 3 records the first record was appended to the basefile and additional variables were computed from the remaining records. Specifically, the total property value across all records was summed and appended to the basefile as a variable. In addition, the first three drug types were identified and appended to the file as separate variables.

Segment 4. The victim segment provides important information for an analysis of arrest, especially if victim characteristics are uniquely predictive of arrest. In most instances a single victim existed and could be linked to all offenders in the incident; single victim incidents comprised about 86% of the incidents used in this study. A more complicated procedure was required for multiple victim incidents in which the victim record with the most serious victimization offense, as determined by the offense seriousness ranking, was used as the main victim record in the analysis. The procedure involved assignment of a seriousness ranking score to each victim offense variable (data element 24) and evaluation of the scores across the ten possible offenses related to that victim. The record with the most serious victimization in the offense was appended to the base file and was matched along with the specific offender number related to the victim, found in element 34.

Additional victim variables were constructed to count the total number of victims, number of victims under 18 years of age, number of racial minority victims, and number of Hispanic victims. The variables related to number of minority and Hispanic victims were limited in usefulness by the small number of victims that fell into these categories.