

# APPENDIX B

## Accuracy of Sample Data

### CONTENTS

Confidentiality of the Data .....	B-1
Editing of Unacceptable Data .....	B-1
Sources of Error .....	B-2

### CONFIDENTIALITY OF THE DATA

Every effort has been made to protect the confidentiality of the 1997 U.S. Virgin Islands Survey data, and to make sure that published data do not disclose information about specific individuals, households, or housing units. All full-time or part-time employees of the University who handle the questionnaires are required to swear to the same oath of confidentiality that is administered by the Bureau of the Census in the collection of census data.

### EDITING OF UNACCEPTABLE DATA

The objective of the processing operation was to produce a set of data that describes the population as clearly and accurately as possible. The questionnaire contained three sections: population and housing, general expenditure and daily diary. The general expenditure and daily diary sections sought to obtain data on the spending habits of respondents on items for all categories of goods and services purchased in the Virgin Islands. The general expenditure covered the one year period and the daily diary a two or one week period before the reference date of the survey.

Questionnaires were reviewed and edited during field data collection operations by field supervisors for consistency, completeness, and acceptability. Questionnaires were also reviewed by edit clerks in the Survey office for omissions, certain inconsistencies, and improper sample selection. For example, write-in entries such as “Don’t know” or “NA” were considered

unacceptable in certain quantities and/or in conjunction with other data omissions.

As a result of this review operation, a follow-up telephone or personal visit was made to obtain missing information. Potential coverage errors were included in the follow-up, as well as questionnaires with omissions or inconsistencies beyond the completeness and quality tolerances specified in the review procedures.

Subsequent to field operations, remaining incomplete or inconsistent information on the questionnaires was assigned using imputation procedures during the final automated edit of the collected data. Allocations—computer assignments of acceptable data in place of unacceptable entries or blanks—were needed most often when an entry for a given item was lacking, or when the information reported for person, housing unit or household expenditure was inconsistent with other information for that same person or housing unit. As in previous surveys, the general procedure for changing unacceptable entries was to assign an entry for a person or housing unit that was consistent with entries for persons or housing units with similar characteristics. The assignment of acceptable data in place of blanks or unacceptable entries enhanced the usefulness of the data.

Another way to make corrections during the computer editing process is substitution. Substitution is the assignment of a full set of characteristics for a person or housing unit. Because of the detailed field operations, substitution was not needed for the 1995 and 1997 surveys.

## **SOURCES OF ERRORS**

In any large-scale statistical operation, such as the 1997 U.S. Virgin Islands Consumer Expenditure Survey, human- and machine-related errors do occur. These errors are commonly referred to as nonsampling errors. Such errors include not enumerating every household or every person in the population, not obtaining all required information from the respondents, obtaining incorrect or inconsistent information, and recording information incorrectly. In addition,

errors can occur during the field review of the enumerators' work, during clerical handling of the questionnaires, or during the electronic processing of the questionnaires.

To reduce various types of nonsampling errors, a number of techniques were implemented during the planning, data collection, and data processing activities. Quality assurance methods were used throughout the data collection and processing phases of the survey to improve the quality of the data.