

# UNITED STATES CENSUSES OF POPULATION AND HOUSING, 1960: PROCESSING THE DATA

## CHAPTER I. INTRODUCTION

At 10-year intervals, the Government of the United States, through the Bureau of the Census, takes censuses of its population and housing. These are large-scale undertakings. The 1960 censuses involved visits to 60 million homes to get information on more than 180 million people and their dwellings.

Beginning on April 1, 1960, each housing unit was visited by a census enumerator. At each occupied unit, the enumerator obtained information on 5 characteristics for each member of the household--age, sex, marital status, color, and relationship to the head of the household--and on 10 characteristics of the housing unit--type of housing unit, access to unit, whether it contained a kitchen or cooking equipment (and whether it was for shared or exclusive use), condition of housing unit, occupancy, number of rooms, running water, flush toilet, bathtub or shower, and tenure. For vacant housing units, the same housing items were obtained except that information on vacancy status--whether for rent or for sale or "other"--was obtained instead of information on tenure. In addition, in those cities for which data were to be published separately for each city block, data were obtained on the value or rent of every housing unit. All of these items were collected on a 100-percent basis, i.e., for every person and housing unit in the area. Each enumerator recorded these data for every person and housing unit in his enumeration district (ED) in an enumeration book which contained schedules carrying only these "100-percent" questions. If the household had filled in an Advance Census Report form which it had received in the mail before the census, the enumerator transcribed the information from that form to the enumeration book; if not, he obtained the responses by questioning the householder and entered the information directly in the enumeration book. A separate book (or books) was used for each ED, and no enumeration book, therefore, contained information from more than one ED.

In addition to the collection of these "100-percent" data, additional information was obtained from every fourth housing unit. A feature of the 1960 Censuses of Population and Housing was the separation of the collection of data obtained on a 100-percent basis from the data obtained from the 25-percent sample, so that the data were collected in two stages--stage I for the 100-percent data and stage II for the sample data--in the more densely settled areas which contained about 80 percent of the population and housing units of the Nation. The stage-I enumerator left at the sample household a household questionnaire con-

taining the sample questions, and requested that it be filled out and mailed to the local Census District Office. When the sample questionnaires were received by mail in the District Office, they were sorted by ED. Each stage-II enumerator was assigned to about three ED's. He transcribed the answers from the sample household questionnaires to a separate enumeration book provided for the sample data for each ED. He obtained missing information by followup interview, in person or by telephone, and entered it directly in the enumeration book.

In the remaining, sparsely settled areas of the country, the enumerator entered the data for every fourth housing unit in the sample enumeration book at the time of his original visit when he obtained information on the 100-percent items.

Thus, for each ED, there were one or more enumeration books containing the data collected on a 100-percent basis and one or more enumeration books containing the sample data.

A sample Survey of Components of Change and Residential Finance (SCARF) was made in connection with the 1960 Census of Housing, (a) to provide measures of changes in the housing inventory--new construction, demolition, conversion, and merger--since the 1950 census and since the 1956 National Housing Inventory, (b) to provide information on the characteristics of the housing units involved in each type of change and of the units in which there was no change, and (c) to provide information about mortgaged and mortgage-free properties, characteristics of mortgage financing, and selected characteristics of property owners. The data for the first phase of SCARF were collected before the regular censuses of population and housing, and the first stages in processing the SCARF data took place before the massive operations on the main census data began. A second phase took place after the regular censuses.

With the separation of the 100-percent items from the sample, the collection of much of the data (including all of the data that needed manual coding) on a 25-percent sample, the use of high-speed electronic equipment, and a system of operations which expedited the flow of work, the results of the 1960 censuses have appeared significantly earlier than the corresponding publications of earlier censuses. The following comparisons provide evidence of the advance in publication dates that was accomplished for the 100-percent population data:

Report series	Publication date of first State report			Publication date of last State report		
	1950 census	1960 census	Timing gain	1950 census	1960 census	Timing gain
Advance reports on number of inhabitants (final counts)	Feb. 1951	Aug. 1960	6 mos.	Oct. 1951	Dec. 1960	10 mos.
Final detailed reports on number of inhabitants.	July 1951	Sept. 1960	10 mos.	May 1952	May 1961	12 mos.
Final detailed reports on general characteristics of the population.	Mar. 1952	Oct. 1960	16 mos.	Feb. 1953	July 1961	22 mos.

Although the content of the reports on general characteristics of the population in 1960 was less inclusive than the corresponding reports for the 1950 censuses (all items which would require a separate clerical coding operation having been omitted from the 100-percent enumeration in 1960, and the information collected instead on a sample basis) the figures given above correctly reflect gains of 6 to 22 months in the dates of public issuance of key population figures from the 1960 population census. There were similar gains in the issuance of statistics from the 1960 housing census, and comparable gains were expected in the issuance of the U.S. summary reports terminating each series of State reports, and in the later series based on the sample, as well as in the issuance of the bound volumes that bring the series together.

## BASIC APPROACH TO THE DATA PROCESSING

### General Plan of Tabulation Procedures

The main purpose of census data-processing operations is the production of statistical tables containing the results of the censuses. In the interest of maintaining comparability from one census to the next, it was agreed that the arrangement of the data and the table formats would not be radically different from those of the 1950 censuses. It was also decided at a quite early date that population statistics and housing data would generally be published in separate volumes, the only major exception occurring in the case of reports for census tracts.<sup>1</sup>

The law requires that the Bureau of the Census report the final population count of each State to the President of the United States by December 1 of the census year, for subsequent transmittal to the Congress for purposes of reapportionment of the House of Representatives. Of necessity, these figures had to be obtained from the data collected on a 100-percent basis. In selecting questions to be included on the 100-percent schedules, it was decided

<sup>1</sup> Census tracts are small geographic areas delineated for statistical purposes within selected cities and standard metropolitan statistical areas.

to include only those which could be answered by direct and simple entries that would require no editing or coding prior to processing on the electronic equipment.

These 100-percent data, as a result, could be processed rapidly, to permit early publication of total population and housing counts. It was found possible to obtain on a single run of the tapes through the electronic computers not only the basic information required for the total counts of population and housing but also statistics for the population and housing characteristics collected on a 100-percent basis. This meant that the publications containing total counts by area could be rapidly followed by publications on population and housing characteristics.

Because the sample population questions included some whose answers could not be precoded by the enumerators--e.g., occupation and industry, place of work, place of birth, etc.--the enumeration books containing the sample data had to be processed through a manual editing and coding operation. The sample data therefore took longer to process.

As in previous census publications, the larger the area, the greater the level of tabulation detail provided. For the smallest areas, such as minor civil divisions, only so-called inventory-type data, e.g., counts of persons by age and sex, are published; whereas detailed cross-tabulations are tabulated and published for the United States as a whole, for States, for large standard metropolitan statistical areas, and for large cities.

### Geographic Area Work Units

The ED's served as basic building blocks in the data processing. Data for single ED's or small groups of ED's constituted the statistics for the smallest reported geographic places such as towns, tracts, wards, or townships. Data for these ED's were then combined to provide statistics for progressively larger geographic entities such as cities, counties, standard metropolitan statistical areas, States, regions, and, finally, U.S. totals.

There were 239,000 ED's established originally for the 1960 censuses in the 50 States and the District of Columbia, ranging in size from a part of a city block to hundreds of square miles. However, in both the field operation and the processing operations, it was sometimes necessary to separate, or split, an ED into two or more parts. This was done during enumeration because of annexations and other changes in place boundaries, or to reduce the enumerators' workloads, or for various other reasons. It was done during processing because the computers were programmed to handle only ED's having populations of 4,000 or less, and consequently an ED having a population of more than 4,000 had to be arbitrarily split. The total number of ED's, including splits, was 272,600 (including 19,198 ED's with no population).

For geographic identification and control purposes, the ED's were grouped by "prefix areas." These usually consisted of either a complete county, a city of 50,000 or more inhabitants, or that portion of a county which lay outside a city of 50,000 or more inhabitants. There were 3,628 prefix areas in the 50 States and the District of Columbia.

Processing for both the 100-percent and sample data was conducted on a State-by-State basis, i.e., each major phase of the operation was performed for all parts of a State before that State was released for the next major phase.

### Steps in the Data Processing

The processing of the data from the 1960 Censuses of Population and Housing included the following principal steps:

1. Receiving, from the census field offices, enumeration books in which the enumerators had recorded most of the responses to the questions by filling in the appropriate circles on schedules especially designed for processing on electronic equipment (See example in appendix I.)

2. Checking enumeration books against control records

3. Coding on the sample schedules of those items of data to which there could be so many different responses that the enumerators had been instructed to write the answers in longhand on the schedules (These were in all cases items of population data which had been collected on a sample basis.)

4. Microfilming the schedules--first those containing data collected on a 100-percent basis, then those containing the sample data

5. Putting the microfilm through a system of electronic equipment known as FOSDIC--Film Optical Sensing Device for Input to Computers--which converted the marks on the microfilmed schedules into magnetic signals on computer tape

6. Feeding these tapes, together with instruction tapes known as computer programs, into an electronic computer which edited and tabulated the data

7. Running the tapes containing the tabulated data through the computer with tapes containing the requisite historical data and with instructions to the computer. The computer arranged the data in publication table format. The tables were then printed on a high-speed electronic printer, for offset reproduction

The data-processing job might be thought of as a series of successive condensations. Information collected by enumerators on a tremendous mass of schedules, which occupied 9 miles of steel shelves when they were received from the field offices, was transferred to microfilm which occupied approximately 575 cubic feet of files, then to compact files of computer tape, and finally summarized in the computers and issued as a set of tables in published census volumes.

## THE DATA-PROCESSING ORGANIZATION

### Decennial Operations Division

In July 1958, the Bureau established a new division, the Decennial Operations Division, to handle the data-processing operations for the 1960 Censuses of Population and Housing. The Decennial Operations Division was charged with responsibility for the planning, coordination, control, and accomplishment of all processing operations for these censuses and related surveys. Its duties included formulating detailed processing procedures, programming for computer operations, and developing integrated time schedules. It worked in close cooperation with the Population, Housing, Field, Geography, Electronic Systems, Machine Tabulation, Statistical Methods, and Statistical Research Divisions as well as with Administrative Service, Personnel, and Budget and Management Divisions.

The organization of the Decennial Operations Division consisted of a headquarters staff in Washington, a processing office in Jeffersonville, Ind., and another in San Juan, P.R., and processing units at computer installations in Chicago, Ill., Chapel Hill, N.C., and, for a short time, Rome, N.Y., which supplemented the tabulation work done by the electronic computers at the Bureau of the Census headquarters. Organization charts appear in appendix C.

Washington headquarters.--The Washington headquarters staff of the Decennial Operations Division included professional, technical, and administrative personnel engaged in developing clerical procedures; evaluating processing methods and equipment; preparing quality control procedures; preparing, testing, and monitoring computer programs; and performing administrative and managerial functions. This staff was supplemented by a small clerical force assigned to reviewing the results of the computer and high-speed-printer production and performing various computer processing activities.

Jeffersonville Decennial Operations Branch.--The Decennial Operations Branch in Jeffersonville, Ind., was responsible for clerical processing of the data. The operations performed in Jeffersonville included receiving and checking in the enumeration books, coding specified population items entered in the enumeration books, preparing documents for microfilming, microfilming enumeration schedules and related documents, compiling, checking, and verifying statistical tables, and performing a variety of related clerical processing operations.

Puerto Rico Decennial Operations Office.--The clerical and machine processing of the data collected in the 1960 Censuses of Population and Housing for the Commonwealth of Puerto Rico was carried out in the Decennial Operations Office established for this purpose in San Juan, P.R. Clerical operations were similar to those performed in Jeffersonville, but punched cards and conventional card tabulating equipment were used rather than the microfilm and electronic system. The Puerto Rico data-processing operations are described in chapter VI.

Computer installations.--To augment its electronic computer facilities in Washington, the Bureau made advance arrangements with the University of North Carolina (Chapel Hill, N.C.) and the Armour Research Foundation (Chicago, Ill.) for the installation of Univac 1105 computers, compatible with those of the Bureau. In addition, for a short time an 1105 computer of the Air Materiel Command at Griffiss Air Force Base near Rome, N.Y., was used on a contract basis. The Decennial Operations Division's processing units at these installations consisted of from two to six employees who were responsible for directing and controlling the operations required to schedule the census data through the computers.

### Other Divisions of the Bureau

The data processing of the 1960 censuses was a major activity and concern of a number of other divisions of the Bureau in addition to the Decennial Operations Division.

The Population Division and the Housing Division provided the subject specifications for the tabulations and for the coding of the items on the schedules, reviewed the final tables before publication, and in general provided those services which required an expert understanding of the subject content of the censuses. They also detailed members of their staffs to work in the Decennial Operations Division for long periods of time,

both for training and supervising employees engaged in coding and editing particular items on the schedules and to provide subject guidance during the preparation of the computer programs.

The Geography Division provided continuous assistance in resolving problems of definition and boundaries of geographic areas--problems crucial to the data processing since each person and housing unit in the censuses was included in just one ED and then was included in the tallies for all the successively larger areas of which that ED was recognized to be a part.

The Field Division, which had directed the census enumeration, assisted in resolving problems which had to be referred back to the Census District Offices.

The Electronic Systems Division<sup>2</sup> was responsible for the actual maintenance and operation of the electronic equipment--FOSDIC, computers, and high-speed printers. In addition, this division provided general programing materials, trained programers, and rendered substantial programing assistance to the Decennial Operations Division.

The Machine Tabulation Division<sup>2</sup> conducted the processing on punchcard equipment of some special areas and groups, notably the outlying areas of American Samoa, the Canal Zone, Guam, and the Virgin Islands, and the overseas population, and also provided instructions for the data-processing operation conducted in Puerto Rico. This division's most notable contribution to the data

processing, however, probably was the construction by its Machine Development Laboratory of the FOSDIC equipment.

The Statistical Methods Division developed and provided technical guidance for the quality control program and for those portions of the data processing which involved problems in sampling methodology.

The Statistical Research Division and the Office of the Assistant Director for Research and Development developed the general census procedures and data-processing techniques throughout the 10-year interval between the 1950 and 1960 decennial censuses. Their work included research on data-processing equipment and methodology as well as on the effect on the data processing of innovations in enumeration procedures, sample design, and other phases of the census.

In addition, the service divisions of the Bureau--Administrative Service Division, Budget and Management Division, and Personnel Division--provided assistance with regard to (1) procurement of space, equipment, and supplies; (2) fiscal controls and production standards; and (3) recruitment and classification of personnel.

The administrative staff of the Jeffersonville Census Operations Office provided services of recruitment, procurement, and facilities for the Jeffersonville Branch of the Decennial Operations Division as well as for other divisions which maintained branches or work units in Jeffersonville.

The administrative relationships of the divisions are shown in the organization chart of the Bureau, in appendix C.

<sup>2</sup>In September 1961, the Electronic Systems Division and the Machine Tabulation Division were combined to form the Data Processing System Division.