

CHAPTER VI

EDITING, CODING, AND TABULATING

The Census data, which had been collected by enumerators throughout the country, were edited, coded, punched, and tabulated in Washington or Philadelphia on a mass production basis. Thousands of new employees were hired by the Bureau in the summer and fall of 1950 to help with this work.

Briefly, "editing" involves inspection of the schedule entries to make sure that they are complete and consistent. It prevents improper information from entering the tabulation process, where it can be buried in a mass of detail and become hard to trace. A considerable amount of editing is done by machine after the information has been transferred from the schedules to punch cards (see "mechanical editing" below).

"Coding" involves preparation of schedule entries for transfer to a punch card. Only the numerals, 0 to 9, and the letters, X and V, can be punched on that card (see Appendix B). Consequently, all data to be punched on the cards and run through the tabulating machines must be converted to those symbols. For example, the enumerator entered "M" on the Population and Housing Schedule for male and "F" for female. These entries were translated into "1" and "2", respectively, before they were punched.

Many codes, such as the one for sex, are simple enough so that the punch card operator can punch them directly from the schedule entries. This is "punch coding." Other codes, however, are more complicated or they require more checking for their determination; so a special coder enters them on the schedule. Some codes, or recodes, are determined after the cards have been punched (see "mechanical coding" below).

"Punching" involves transferring the codes for the schedule entries to the punch cards. A card is usually prepared for each person, each farm, or each dwelling unit. Each column (or group of columns) on the card represents a characteristic of that person, farm, or dwelling unit, and only one entry is made in each column (except when codes are used for letters of the alphabet). Thus, column 26 on the population P card was for the sex of the person enumerated. On the card for a male, a hole was punched in the space allotted to "1", the code for male. When the card was tabulated, one male was counted.

"Mechanical editing" involves inspection of the punch cards by machine to detect errors in punching, coding, or enumerating. The machines make three types of checks: (1) checks for impossible codes, (2) checks for inconsistent code combinations, and (3) checks of quantitative data for relationships, magnitudes, and arithmetic.

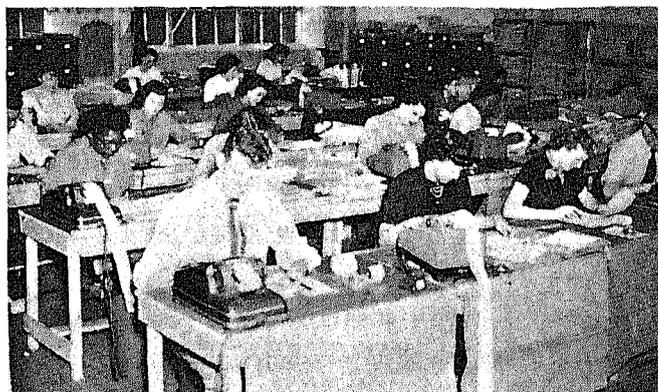
An impossible code is a number which has no meaning attached to it. For example, a perforation of "4" in column 26 (sex) of the population P card was "impossible" because only "1" and "2" had meaning ("1" for male and "2" for female).

Inconsistent code combinations result when the code punched in one column or group of columns is inconsistent with that in another. For example, a code in one column indicating that a dwelling unit was vacant would not be consistent with a code in another column indicating that five persons lived there.

Relationship checks are made when two figures are closely related. For example, farm acreage irrigated could not be greater than the total acreage of the farm. Magnitude checks are used when any figure above or below a certain amount would be unreasonable. Thus, a report for a farm indicating a yield of 100 bushels of wheat per acre is unusual and probably wrong. Arithmetic checks are made when an arithmetic relationship exists among figures punched on the card. For example, total cattle on hand should equal the sum of the numbers of cows, calves, heifers, bulls, and steers.

"Mechanical coding" or "recoding" involves the development of a code mechanically from entries already on the punch card. A machine sorts the cards into groups according to specified combinations of entries. The code for each group is then punched in another field of the cards. Finally, the code punched in each card is mechanically verified. For example, if a column were needed to give the combination, male white, male nonwhite,

female white, and female nonwhite, the cards could be sorted into those groups according to columns 25 and 26 on the population P card. Then, all cards in the male white group could be punched "1" in another column, all cards in the male nonwhite group could be punched "2" in that same column, etc. Finally each card recoded "1" would be checked to be sure that it was punched "1" (for white) in column 25 and "1" (for male) in column 26, and so on. Most of the recoding for the 1950 Censuses was done on the Census Multi-Column Sorter (which sorted the cards), the Census Gang Punch Machine (which punched them), and either the Census Unit Count or the IBM Electronic Statistical Machine (which verified them). The Census Recode Machine, which became available late in the Census period, can sort, punch, and verify in one operation.



Section of office staff checking results.

"Tabulating" is summarizing the information from the punch cards. The machine reads the entries on the cards, adds or counts the number in each classification, and prints the results. For example, the number of males and females in each 5-year age group could be determined by wiring the machine to relate the entries in column 26 (sex) of the population P card to those in columns 27 and 28 (age). When so wired, the machine would determine the number of males under 5 years of age by counting the cards with "1" (for male) punched in column 26, with "0" punched in column 27, and with "0", "1", "2", "3", or "4" punched in column 28. The numbers in each age and sex group would be printed on the "tabulation sheet."

Plans for editing, coding, and tabulating were started in the summer of 1949. After the questions to be asked in the Censuses had been determined, subject-matter specialists began to draft the table outlines and to plan the editing and coding. Machine technicians worked with them to insure optimum use of the mechanical equipment. The plans included not only the data to be published, but also those data to be used for checking and for more intensive research work.

Interested persons, including the appropriate advisory committees, reviewed the proposed table outlines. The plans were then adjusted and coordinated. By March 1950, the subject-matter specialists had prepared working copies of table outlines, and machine technicians had drafted tabulation specifications for the initial operations. On the basis of these specifications, the machine technicians formulated the tabulation procedures, prepared instruction materials, trained personnel, and scheduled the machine operations.

In addition to planning the tables, the subject-matter divisions handled the manual editing and coding operations. The Machine Tabulation Division punched the cards (punch coding many entries), made mechanical edits, tabulated the data, and sent the tabulation sheets to the subject-matter divisions, which reviewed the tabulations and prepared the material for publication.

This chapter is divided into parts corresponding to the subject matter of the data handled--population, housing, Survey of Residential Financing, agriculture, irrigation, drainage, and Post-Enumeration Survey. A final section describes some of the mechanical equipment used.

This chapter describes the general procedures followed in handling the schedules. Details on how the various subject items were treated are given in Chapters VIII, IX, and X.

Population

Editing

The editing of the population and housing questionnaires involved: (1) a screening operation in which all portfolios were inspected and those that had been poorly enumerated were rejected; and (2) an editing operation in which the rejected portfolios were "repaired." The initial screening operation included an inspection of the housing items as well as the population items.

The screening determined whether a portfolio was ready to go into the general coding operation or whether the schedule entries would raise so many questions in that process that it could not be handled routinely. If considerable editing was needed, it was less costly to separate the faulty portfolios, correct them centrally, and then return them to the regular processing channels than to have them interrupt the smooth flow of subsequent coding and punching operations.

The portfolios that had been poorly enumerated were identified with a minimum amount of time and effort by inspecting a sample of schedules. Every fourth schedule in the portfolio was examined for types of errors which the coder was not expected to correct. These errors were of three kinds: (1) errors in items which were not to be inspected by the coder but were to be punched directly from the schedule (housing items, for example); (2) errors which could be detected only by comparing related entries on the schedule (inconsistencies between entries for migration, for example); and (3) errors which required considerable time to correct or which could be corrected only under very complex rules (incorrect entries of serial numbers, for example).

The screener made the following checks of the sample schedules:

- A. General population entries:
 1. The portfolio was to contain only the schedules for the enumeration district described on the Portfolio Control Label.
 2. Entries in the basic population items (relationship, race, sex, age, marital status, birthplace, and citizenship) were to be codable.
 3. Entries on the sample lines were to be consistent with entries on the 100-percent lines.
 4. Migration entries were to be consistent and complete.
 5. Education entries were to be acceptable.
- B. Economic data entries:
 1. Employment status entries were to be consistent.
 2. Employment status entries were to be reasonable in terms of the entries for occupation and industry.
 3. Entries for weeks looking for work and for weeks worked were to be acceptable.
 4. Entries for individual income were to be consistent with entries for weeks worked, for class of worker, and for family income.
- C. Housing entries:
 1. The serial number of the dwelling unit and the number of persons in the household were to be the same on the population and housing sides of the schedule.
 2. Entries for housing characteristics (items 3, 4, 5, 16, and 24) were to be acceptable.

The screener also checked all housing lines on all schedules for cities of 50,000 or more population to see if the block numbers had been entered correctly. The housing edit was limited at this stage, because many of the items were later edited mechanically.

The screener tallied each error in a portfolio on a Screening Reject Slip, and when he had entered a specified number of tally marks for some check item, he rejected the portfolio. (The number of permissible errors varied with the items checked and the number of schedules in the portfolio.) After the screener rejected the portfolio, he discontinued tallying errors for the check which caused the rejection, but he continued tallying errors for other checks. Some portfolios were rejected for more than one reason. The screener rejected portfolios without tallying the errors if they had erroneous or missing block numbers when such numbers were required, inconsistencies in migration entries, or any defect of major importance. About half of the portfolios were rejected in the screening operation.

The portfolios accepted by the screeners were expected to be free of major defects. A sample of them was checked, however, to be sure that the screeners had done their work properly.

Editors were organized into groups to "repair" the rejected portfolios. One group edited portfolios rejected for errors in general population entries, another repaired those with errors in the economic data, and another corrected the housing items. Although the screeners had inspected only a sample of the schedules of the portfolio, the editors examined every schedule. They confined their corrections, however, to the items for which the folio was rejected.

After the editors had corrected the schedules, they returned the portfolio to the bin which held the other portfolios in that work unit. The entire bin was then moved to a coding section.

The screening and special editing operations were started the latter part of June 1950 and were completed the first week of January 1951. During that period, portfolios were screened and repaired at a rate of approximately 8,800 lines (or persons enumerated) per clerk day. The entire operation required more than 17,000 clerk days.

Coding

Much of the coding was done by the punch card operator as he punched the cards. Some coding (or recoding) was done mechanically on equipment which assigned codes according to the combinations of entries in different items. Some items, however, were best coded by a clerk who examined the entries; he then wrote the codes near them on the schedule. Special code columns had been printed on the Population and Housing (P1) Schedule for a few of these items--column A for relationship, column B for birthplace and citizenship, column C for occupation and industry, etc.

Types of items which required manual coding were:

- a. Items which had so many categories that it was necessary to refer to complex code schemes and reference books (for example, occupation and industry).
- b. Items for which some entries could be punch-coded but other entries required reference to instructions. In such cases, the code clerk examined every entry, but he entered the codes only in certain cases. For example, the code clerks did not code relationship entries of "head," "wife," "son," or "daughter" if they were correct, but did code other entries such as mother-in-law, cousin, maid, lodger, and inmate.
- c. Items which required inspection of the information for more than one person. (The punch card operator could examine entries for only one person at a time, and the mechanical coding equipment inspected only one card at a time.) For example, "married, spouse present" could be entered only by checking the line for the spouse to see if he was enumerated in the same household.
- d. Items which required some operations before coding. For example, blanks in income items were changed to entries of "None" if other information made such an entry reasonable.

Coders were organized into specialized units. Each section had a General Coding Unit and an Occupation-Industry Coding Unit. The general coders usually worked on the portfolios first and then made them available to the occupation and industry coders.

To process the schedules with maximum speed and efficiency, the coders usually worked on one item at a time, coding vertically down the schedule and disregarding entries in other items except when special editing was needed. They coded each sheet completely before proceeding to the next sheet. In a further attempt to speed up the process, the coder did not stop to get an answer to an isolated problem unless it would affect the coding for a large part of the portfolio. In that case, he referred it immediately to one of the technical assistants assigned to the unit. Otherwise, he entered a brief description of the difficulty on a Problem Referral Slip and turned in the slip with the completed portfolio. The problem was reviewed later by a technical assistant, and, if he could not solve it, by a subject-matter specialist.

General Coding.--The general coders entered all required population codes except those for occupation and industry. They used a manual "General Coding Instructions", which gave the codes in full detail, and a summary "Code Card" which provided, in convenient form for quick inspection, the codes used most frequently. To determine the birthplace or ancestry of a person from his surname, the coders referred to lists of common Spanish surnames, of typical French-Canadian surnames, and of surnames common to Austria, Hungary, Czechoslovakia, and Yugoslavia.

General coding was checked by a system of quality control. The work of each coder was completely verified at first, and

records were kept for him. Coders who reached a certain level of accuracy were designated "qualified coders"; thereafter, only 8½ percent of their work was verified. Verification of the coding was also verified. One method used was to note coding errors before the portfolio went to the verifier; after the verifier completed his work, the errors he listed were compared with those noted.

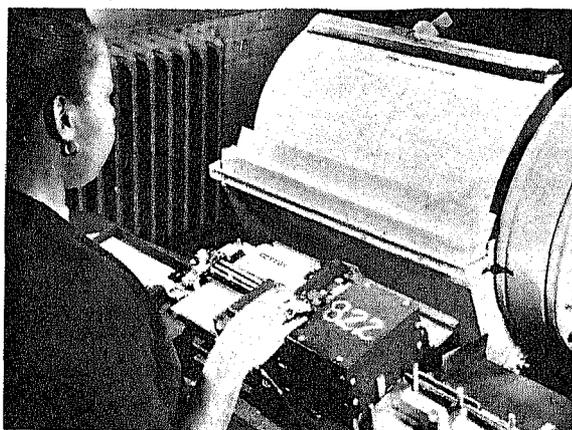
General population coding began in June 1950 and lasted about 10 months. The coders processed the schedules at an average rate of 1,800 lines (persons enumerated) per clerk day.

Occupation and Industry Coding.--The occupation and industry coders used the manual, "Alphabetical Index of Occupations and Industries," to determine the codes for the schedule entries. Occupations and industries not listed in the "Index" were coded by experienced persons until the regular coders became skilled in interpreting the entries. The verification procedure was the same as that for general coders except that 10 percent of the work of the qualified occupation and industry coders was verified.

Occupation and industry coding started in July 1950 and proceeded at an average rate of 1,665 lines (persons enumerated) per clerk day. This operation took 10 months to complete.

Punching

After the schedules were coded, information for each person was transferred to the basic population P punch card. One card was punched for each person giving his residence (farm or nonfarm), race, sex, age, marital status, citizenship, birthplace, employment status, occupation and industry, migration status, parents' birthplaces, education, income, and veteran status. The numbers of the sheet and line on which the person was enumerated were punched on the card. Area identification codes for all persons in each enumeration district were automatically duplicated from a prepunched master deck of punch cards.



Punching the population information using the Richards copyholder.

A special copy holder was designed to hold the Population and Housing (P1) Schedule for this operation. After the 100-percent questions were punched for the sample line, the machine spaced itself automatically to the sample questions, then returned to the next line.

The punch card operators began work on the P cards in August 1950 and completed the job by July 1951. About 112,000,000 of the 151,000,000 cards for the continental United States were punched at the Philadelphia Decennial Tabulation Office; the other 39,000,000 were punched in Washington. About 114,000 man days were used on this operation.

A variation of the P card was used for each of the Territories and larger possessions. About 2,200,000 cards were punched for residents of Puerto Rico, 500,000 for Hawaii, 128,000 for Alaska, 27,000 for the Virgin Islands, 53,000 for the Canal Zone, 59,000 for Guam, and 19,000 for Samoa. The punching operation for the Territories and possessions took 7,300 man days; it was started in March 1951 and completed the following November. All cards for the Territories and possessions were punched in Washington.

The punching was verified by a system of quality control. Initially, each punch card operator's work was completely verified. When he met certain minimum efficiency standards, he became a "qualified puncher," and only 5 percent of his work was verified. Punchers also worked under an incentive pay plan whereby their pay was adjusted if they exceeded certain rates of production and maintained certain standards of quality.

Mechanical Recode

After the P card was punched, the employment status classification was determined for each person 14 years old or over, and the recode was punched on the card. This classification was based on the perforations in 10 columns of the punch card: column 24, relationship; columns 27-28, age; column 33, main activity; column 34, work at all; column 35, looking for work; column 36, have a job; columns 37-38, hours worked; and column 45, class of worker.¹

The cards for the continental United States were put in the Census Multi-Column Sorter, which sorted them into appropriate groups on the basis of the perforations in the 10 columns. Cards for all persons working at a civilian job went into one group, cards for persons with a civilian job but not at work fell in another, etc. The code for each group of cards was then "gang-punched" in column 78 of the card on the Census Gang Punch Machine. Thus, all cards for persons who were at work on a civilian job were punched "1", those for persons with a civilian job but not at work were punched "2", etc. Finally, the codes were mechanically verified on the Census Unit Count Machine.

The mechanical recodes for the Territories and possessions were prepared on the Census Recode Machine.

Tabulating

The population data collected by the enumerators reached the tabulating process in the form of punch cards--at least one for each person enumerated. These cards were in county groups. Within each county, they were divided first into cities of 50,000 or more and the balance of the county; then, within those groups, they were arranged by minor civil divisions and enumeration districts.

The information, which was scattered through the 150,000,000 P cards, had to be funneled into meaningful groups. These groups had been generally determined by the Census staff in consultation with users of the data. The subject-matter specialists put the general requirements in the form of tables, which were translated by the machine specialists into tabulation specifications. Those specifications indicated how the cards should be grouped and how the machines should be wired to provide the data for the proposed Bureau publications.

The tabulations of the basic population P cards provided data for three series of State bulletins, for census tract bulletins, and for special studies. The first series of reports (Series P-A) gave the number of inhabitants in all political subdivisions of the State. Series P-B gave the general characteristics--such as sex, age, race--of the people in small areas of the State. Series P-C provided detailed characteristics--such as age of employed persons in each occupation--of people in relatively large areas of the State. Series P-D provided general characteristics by census tracts in tracted areas.

Other tabulations of the P cards were needed to get more detailed data on some subjects or on particular groups of people. A 3 1/3 percent sample of the cards was drawn for Series D tabulations, which provided detailed cross-classifications of selected characteristics for the United States and, in a few cases, for regions. Series W tabulations provided data for six special reports: (1) institutional population, (2) nativity and parentage of foreign-white stock, (3) State of birth, (4) Puerto Ricans in the continental United States, (5) white persons with Spanish surnames, and (6) nonwhite population by race. Series Z tabulations provided data on the mobility of the population.

The tabulations for each series resulted in a series of "counts," which were the data-producing steps. Each count was planned so that it would use as much of the capacity of the tabulating machine as possible. Details of these counts are given in Appendix B.

The tabulation specifications for each count showed exactly how the cards were to be handled. In count 2 of Series B, for example, the cards for each county were grouped according to census tracts, untraced urban places of 2,500 or more, remainder of urban fringe, rural nonfarm balance of the county, and rural farm balance of the county. These groups were run through the machine separately so that data could be obtained for each. The cards for each area were sorted by sex and color, and each of the sex-color groups (male white, etc.) was classified by age, veteran status, school completed, type of household, race and nativity, marital status, and employment status.

Wherever possible in the tabulation process, techniques were introduced to save time and reduce costs. One of these techniques was the preparation of summary cards. When totals

¹See Figure 18, Employment Status Recode, in Chapter VIII.

were obtained for a specified group of people, summary cards giving these totals were punched automatically. The summary cards were then run through other machines to get the data for larger groups. This summarization process reduced the number of cards needed for tabulating the Series B data for final publication from 150,000,000 to 1,500,000.

Cards needed for special purposes were duplicated to expedite the work. Machines selected and reproduced these cards. The duplicate cards were then used for the special tabulations without interrupting the main operations. For example, a duplicate 20-percent deck was established for the tabulation of sample data.

Family and Fertility Data

Family and fertility data for persons on selected lines of the Population and Housing (P1) Schedule were transcribed to special sheets.¹ From these sheets, the family F card and the fertility C card were punched.

A family card was punched for each person in the sample who was head of a primary or secondary family or of a subfamily, who was a primary or secondary individual (that is, was living alone or with persons not related to him), or who was a male who had been married but was not head of a family unit. For heads of households (primary family heads and primary individuals), data for some columns were reproduced from the PH cards (the card with both population and housing data for heads of households).

A fertility card was punched for each ever-married woman in the sample. Most of the data for this card were taken from the transcription sheet, but some columns were reproduced from the population P card or, in a few cases, were punched from additional data transcribed from the Population and Housing Schedule.

Housing

The housing entries, which were on one side of the Population and Housing (P1) Schedule, were precoded on the questionnaire and mechanically edited. Consequently, almost no manual operations were required for these questions.

Manual Editing

Editing of the housing entries was similar to that of the population entries; and it was done at the same time. In the screening operation, portfolios were rejected if they had errors in block numbers (in cities with 50,000 or more inhabitants) or in serial numbers, or if the entries were unacceptable in the items for type of living quarters, type of structure, number of dwelling units in structure, occupancy, or cost of utilities. In the serial number check, item 8 (number of persons in the household) was examined and corrected if in error. The rejected portfolios were edited and repaired by housing statistics specialists.

Manual Coding

The housing entries were precoded. That is, the code for the answer (if it was not a numerical answer which could be punched directly on the punch card) was printed opposite it. For example, a "1" was printed beside the "not dilapidated" check box in item 7. When that box was checked, the punch card operator punched a "1" in the appropriate column on the punch card.

Only one housing item was coded manually and that only in certain instances. That item was "type of living quarters." When the general population coder found a lodging house with 5 to 9 lodgers, he entered a code of "9" in that item and canceled the entry previously made.

Punching the H Card

The basic housing card was the H card. It carried information on the type of living quarters, type of structure, number of dwelling units in the structure, number of persons living in the unit, occupancy, condition of the unit, number of rooms, number of persons per room, and type of facilities. Most of the data were punched manually, but a number of columns were used for recodes or for reproduction of population data from the PH card.

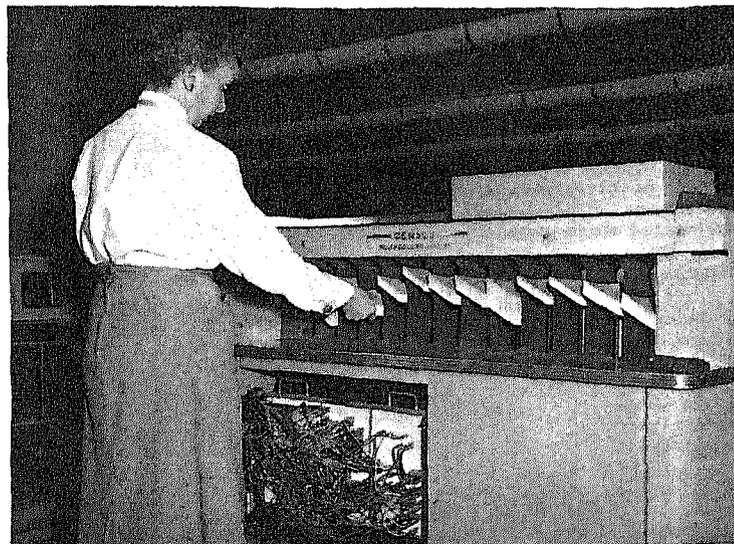
An H card was punched for each of the 46,000,000 dwelling units in the continental United States. Approximately 12,000,000 of these cards were punched in Washington, and the remaining 34,000,000 were punched in the Decennial Tabulating Office in

Philadelphia. Another 620,000 cards were punched for dwelling units in the Territories and possessions. This included 458,000 for Puerto Rico, 119,000 for Hawaii, 34,000 for Alaska, and 8,500 for Virgin Islands. Punching cards for the Territories and possessions required 1,000 man days. The entire operation required about 44,000 man days.

Mechanical Editing of H Card

Most of the editing of the housing schedule was done mechanically. The H card was subjected to six mechanical edits to detect and eliminate inconsistencies in the entries for a dwelling unit and, in some cases, to supply missing information. These six checks were concerned with occupancy, amounts paid by renters, and plumbing facilities.

In the occupancy edit, the entry in item 16 (occupancy) was made to conform to that in item 8 (number of persons in household). Amounts paid by renters were edited before the gross-rent recode was determined; cards were examined when the amount paid for utilities was unusually large. Four edits were concerned with plumbing facilities. Entries of flush toilet for exclusive use were related to the number of dwelling units in the structure,



The Census multi column sorter, developed for the 1950 censuses.

the type of living quarters, and the availability of a bathtub or shower. Entries of shared bathtub or shower were checked against type of living quarters, type of toilet, and number of dwelling units in the structure. Cards for dwelling units without running water inside were examined to see if logical entries were made for type of toilet and bathtub or shower. Entries of shared kitchen sink were made consistent with entries for number of dwelling units in the structure and type of living quarters.

Mechanical Recodes on the H Card

When combinations of entries on the punch card are used frequently, or when several different combinations have the same meaning, it is economical to put the combinations in the form of a "recode" and punch that recode in one of the columns of the punch card. Tabulations can then be made from that one column instead of from several columns. For example, the following 25 combinations of housing items 8 and 9 were needed to identify the dwelling units in which there were 0.5 persons or less per room:

Number of rooms (item 9)	Number of persons (item 8)	Number of combinations
Total		25
2	1	1
3	1	1
4	1, 2	2
5	1, 2	2
6	1, 2, 3	3
7	1, 2, 3	3
8	1, 2, 3, 4	4
9	1, 2, 3, 4	4
10+	1, 2, 3, 4, 5	5

¹For more detailed discussion of procedures and for definitions of different types of families and individuals, see sections on "Families" and "Fertility" in Chapter VIII.

All these combinations were coded '1' which was punched in another column of the card to identify the dwelling units with 0.5 persons or less per room.

Four recodes were used for the housing H card (see Chapter IX). The type-of-structure recode combined information on type of living quarters, number of dwelling units in structure, type of structure, and business unit in structure. The persons-per-room recode (described above) used data on number of rooms and number of persons. The condition-and-plumbing-facilities recode related information on water supply, toilet, bath, and condition of dwelling unit. In the gross-monthly-rent recode, gross rent was computed by adding amounts paid for utilities to the amount paid for rent on an unfurnished basis.

The gross-monthly-rent recode was computed and punched on the cards by the IBM Electronic Calculating Punch Machine. For the other recodes, the Census Multi-Column Sorter, the IBM Reproducing Punch Machine, and the IBM Electronic Statistical Machine were used.

Reproducing Population and Housing Information

The PH card required both population and housing data for occupied dwelling units. The initial entries for these cards were taken from selected P cards for heads of households. The area and household identification, as well as population information, were reproduced mechanically from the P cards to the PH cards.

The PH cards were then matched to the H cards for the same households through the common entries of geographic area and serial number. While the cards were in this relationship, two operations were performed: (1) the housing data from the H card were mechanically reproduced on the PH card, and (2) certain population data from the PH card were punched on the H card.

The population data needed for the H card related to type of family (available only if the household head was on a population sample line), farm residence, and color of occupants. The H cards were punched in two groups. The first group contained cards for all farm households, all nonwhite households, and all households for which the head fell on a sample line; the second group contained all other cards--those for nonfarm white households whose head did not fall on a sample line. For the first group--which was 40 percent of the total--data on type of family, farm residence, and color of occupants were mechanically reproduced, card by card, from the PH cards to the matching H cards. Because the H cards in the second group were all for nonfarm white households whose head did not fall on a sample line, they could be gangpunched as a group.

Vacant, as well as occupied, dwelling units had to be classified as farm or nonfarm on the H cards. This information for vacancies had not been punched on any card so it had to be obtained from other sources. In urban areas, all vacant units were assumed to be nonfarm, and in some other cases nonfarm residence was assumed. In rural areas, however, information on the farm or nonfarm status of many vacant units had to be taken directly from the Population and Housing Schedule.

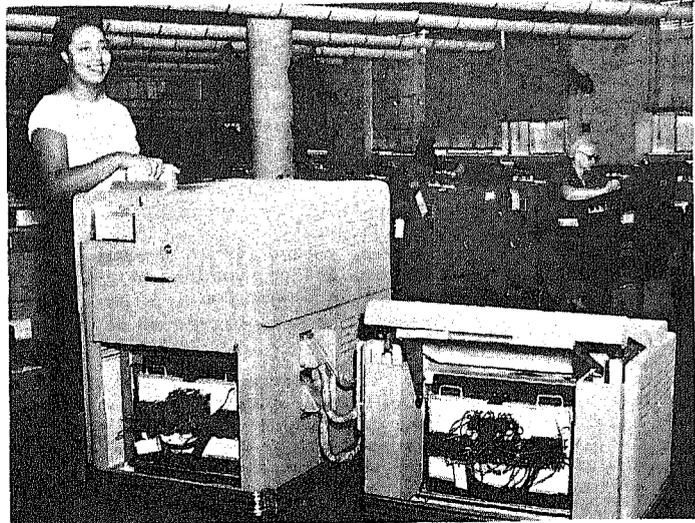
Tabulations

The H and PH housing cards were tabulated in three series. Series A tabulations resulted in State bulletins presenting basic data on virtually all subjects for which housing information was collected. Series B tabulations were limited to the housing characteristics of nonfarm households, and Series C tabulations to farm households.

The data for most of the subjects in Series A were based on the 46,000,000 H cards. For some subjects, however, a 20 percent sample of the dwelling units were used. The tabulations were run in five basic counts (see Appendix B). Data for the larger geographic areas--such as counties, standard metropolitan areas, urbanized areas, and States--were obtained from summary cards for smaller geographic areas. These summary cards had been punched automatically when data for the smaller areas were tabulated.

The PH cards were used to get the data for Series B. These cards were transcribed to magnetic tape which was run through the Universal Automatic Computer (Univac). Five tabulations (two for owners and three for renters) were made for each city and standard metropolitan area of 100,000 or more population. The PH cards were matched to the F (family) cards on the Univac magnetic tape to provide housing data for nonfarm households in the United States and in its nine geographic divisions.

In the Series C tabulations, the farm PH cards were tabulated to provide data for 119 economic subregions of the United States. These subregions are groups of two or more counties which are relatively homogeneous with regard to agriculture.



The Census recode machine, developed for the 1950 censuses.

Data for nonwhite households were tabulated in 35 of these subregions, and data for households of each race were tabulated for the United States.

Housing data were tabulated from the H cards for Alaska, Hawaii, Puerto Rico, and the Virgin Islands. The tabulations for Alaska, Hawaii, and Puerto Rico were similar to those for Series A, but they also included some special cross-classifications. For example, the Puerto Rico dwelling units were classified according to the composition of the exterior walls and roof. Tabulations for the Virgin Islands, which had only 8,000 dwelling units, consisted only of verifying the hand count made in the field.

Survey of Residential Financing

A sample of owner-occupied properties and rental properties (including vacant) was selected for the Survey of Residential Financing, and schedule forms (P51, P52, P53, P54) were filled in the field after the regular enumeration was over (see page).

The questionnaires were received from the field in five groups: mortgaged, not mortgaged, not within the scope of the survey, not selected for personal enumeration, and nonresponse. The forms were examined to see if they had been properly classified, and those representing properties with more than half their floor space devoted to business purposes were removed. In a second step, the schedules were examined to be sure that the correct form was used, that the schedule was made out for the property in the sample, and that income data were complete. The questions were then edited and, where necessary, coded. The data card, which carried information for the owner from the Population and Housing (P1) Schedule, was completed and coded. Owner and lender schedules for the same properties were matched and compared to be sure that entries were consistent. Finally, the schedules were sorted into three groups according to the type of property: owner-occupied; rental, small; and rental, large.

A basic card was punched from information on the questionnaire and on the data card. Entries on this punch card were edited mechanically to eliminate inconsistencies. A work card was reproduced mechanically from the basic punch card, but some of the entries on the basic card were combined into recodes, which were punched on the work card. A number of ratios were computed mechanically and punched on the work card.

All tabulations for this survey were made from the work card. These tabulations resulted in basic data for 25 large standard metropolitan areas, the United States, and its four regions. They also produced detailed cross-classifications for analytical purposes for the United States, for the properties inside the standard metropolitan areas, and for properties outside such areas.

Agriculture

When the agriculture enumeration forms were removed from the enumerators' portfolios, they were counted, recorded on a check list, arranged by enumeration districts, and placed in special portfolios. Each special portfolio contained about 400 questionnaires. The preliminary count of the questionnaires was

recorded on the portfolio label, which also carried the names of the State and county. An average of about five portfolios were needed for a county. The portfolios were kept in a central file, from which they were sent to the appropriate processing unit. Control records were maintained on a county basis.

The first step in preparing the questionnaires for punching involved general editing and coding. In the second step, the general editing and coding were reviewed, the questionnaires in the sample were edited and coded, and problems were solved. In the third step, the questionnaires were arranged, numbered, and counted to obtain final control totals.

Approximately 900 persons worked for an average of six months to prepare the questionnaires for punching and tabulating. These forms were all processed in Washington.

General Editing and Coding

Each questionnaire was carefully edited. It was checked for completeness, consistency of response, and for agreement with other forms. Figures which had been entered carelessly were rewritten, and fractional answers were converted to common denominators. In the later stages of this check, some items were coded.

The first step in the general editing and coding operations was to match the Landlord-Tenant Operations Questionnaires (A3) and the Irrigation Questionnaires (I-1) with the Agriculture Questionnaires (A1) for the same farms. During this matching process, the questionnaires were separated into two groups: (1) those for landlord-tenant areas, and (2) those not for landlord-tenant areas. Instruction manuals were provided for each of these two groups, and they were edited and coded separately.

Codes were entered on the Agriculture Questionnaires (A1) for color of operator, tenure, size of farm, irrigation (in 17 Western States, Arkansas, Louisiana, and Florida only), and the month the farmer began to operate the place. Other items were also coded if the responses were so varied that the card punch operator could not efficiently punch them directly from the questionnaire. These included such items as "other poultry" (ducks, geese, guineas, pheasants, pigeons, and quail), and miscellaneous crops. Space for codes had been provided on the Agriculture Questionnaire in columns which were usually labeled "For Office Use Only."

In the landlord-tenant areas, a plantation might be located in two or more enumeration districts or in two or more counties. Duplicate reporting of tenants and land was therefore a possibility despite the precise rules given to the enumerators. Agriculture Questionnaires for adjoining enumeration districts and counties were examined to detect and eliminate any that were for land reported on another Agriculture Questionnaire. Most of the duplication was discovered when the Landlord-Tenant Operations Questionnaires were separated from the Agriculture Questionnaires and alphabetized by name of operator.

On the Landlord-Tenant Operations Questionnaires (A3), each unit was given a tenure code to indicate whether it was operated by the owner, a cropper, a cash tenant, etc. The landlord-tenant operation was coded as a "multiple-unit" operation if it covered a home farm and one or more units rented to sharecroppers or if it covered no home farm and two or more units rented to sharecroppers. If the tenants were not sharecroppers, the operation was not classified as a multiple unit.

The questionnaires were then checked to see if the entire landlord-tenant operation met the criteria for a large farm. If it did, the home farm (if any) was designated as a large farm.

A special editing section checked county and State identification on the Landlord-Tenant Operations Questionnaires and entered a color and tenure code for multiple-unit operations. It also supplied omitted tenure codes for individual subunits. Figures for acreage, production, value, etc., for the tenants were added to get totals for the over-all operation. Finally, each Landlord-Tenant Operations Questionnaire was numbered. These questionnaires were kept in the front of the first portfolio for the county.

General editing and coding clerks brought problems to the attention of technical assistants and members of the professional staff by filling a reference note. This note was attached to an Agriculture Questionnaire when it was for a place which did not appear to be a farm according to the Census definition, or for a place which appeared to be a large farm but was not so classified. The note was used for farms in non-irrigation States which reported irrigation works and for Landlord-Tenant Operations Questionnaires which covered multiple-unit operations outside the landlord-tenant area. It was also attached to questionnaires which appeared to be duplicates and to others which presented

problems not to be handled by editing and coding clerks, according to the instructions.

Review of Editing and Coding and Disposition of Problem Cases

In this operation, general editing and coding were reviewed, additional reference notes were prepared for problem questionnaires, problems were disposed of, questionnaires for farms in the sample were separated from questionnaires not in the sample, coding was completed, and specific enumeration problems were investigated.

In the review of general editing and coding, codes were verified and reference notes were checked to make sure that all problem cases were listed. Initially, every entry made by an editor-coder was checked, but after his work reached a specified quality, only a sample was reviewed.

Problem cases were referred to technical assistants or members of the professional staff. They reviewed the materials and decided how the cases should be handled.

All forms except the Agriculture Questionnaires (A1) were then removed from the portfolios and filed. Agriculture Questionnaires for places not qualifying as farms under the Census definition were stamped "Reject" and moved to the front of the portfolio. The sample and large farm Agriculture Questionnaires were placed in special portfolios.

The sample and large farm questionnaires were edited and coded for economic class and type of farm. If the sample questions had not been answered, the information was requested by mail. Any operator of a large farm who was not accounted for in the reviewing process was asked to fill an Agriculture Questionnaire.

If the enumerator failed to follow the prescribed method of selecting the sample farms, sample biases were introduced. To detect such biases, a check was made on the selection of farms. Corrections were made if the sample did not meet prescribed specifications.

A number of other checks were made. Reports for farms with land in two counties were examined to make sure that the farm (or part of it) was not enumerated in both counties. If large inventories but no sales of livestock or poultry were reported, other questionnaires filled by the same enumerator were reviewed to see if he had failed to understand the instructions. Similarly, enumerator bias was checked in questionnaires showing production of cash crops (such as cotton or tobacco) but no sales for those crops, and in questionnaires showing unusually high yields per acre.

Arranging, Numbering, and Counting

The Agriculture Questionnaires had been arranged by enumeration districts within each county in the previous operations. Separate portfolios were set up for large farms, sample farms, and other farms within each county; divider sheets were inserted to separate the reports for each minor civil division. The questionnaires in each portfolio were numbered with an automatic numbering machine and simultaneously counted. Numbers for the nonsample questionnaires started with "0001," those for the sample questionnaires started with "8,001," and those for large farms were numbered with a prefix "X" (X001, X002, etc.). The results of the counts were recorded on a label on the front of the portfolio and were used as an over-all control totals in subsequent operations.

Punching the Agriculture Cards

Thirteen card types were used to record the entries on the Agriculture (A1) Questionnaire (see Appendix B). Items were arranged on the cards according to their subject matter and their relationships to each other. The columns common to all cards--card identification, geographic identification, questionnaire number, etc.--were initially established and other necessary columns were added. The number of columns used for each item was based on the experience of previous censuses.

Two land use cards were designed. The one for the Eastern States included the irrigation questions; but the one for the Western States was so designed that it could be related to a separate card which carried the more extensive information on irrigation for the Western States.

For other subjects, however, the differences between areas were not great enough to require an additional card. One card was therefore designed to carry all the information on the subject. If some topic did not apply to a particular area, the columns for that topic were not punched on the card. An "X" was printed in the code column of the questionnaire in the

¹See page 7 for definition of a large farm and page 20 for Census definition of a farm.

position of the missing topic to tell the punch card operator that column was to be skipped.⁴

The different card types used for the Agriculture Questionnaire (A1) and the number punched for each are shown in Table 3, Appendix D.

One card form was punched from the Landlord-Tenant Operations Questionnaire (A3). This card had information on the entire landlord-tenant operation--land owned and rented by the landlord, land rented to tenants, land in crops, production and acreage of different crops (cotton, tobacco or rice, peanuts, corn for grain), horses and mules owned by the landlord and tenants, and the number of tenants. About 147,000 cards were punched.

Punch card operators worked in units of about 40 persons. Each operator in the unit usually worked on one to three specific card types--land in farms, corn and sorghums, etc. Portfolios were assigned to the operator in lots of two, those for counties with the highest priority being assigned first. When an operator punched his card types, he gave the portfolios to the punch operator for another type or to a verifier. This process continued until all card types for the portfolios were completed.

When all cards for a county had been punched, the different card types were boxed separately. Slips identifying the portfolios and county were placed in the card boxes.

At the peak of the agriculture punching operations, more than a thousand persons were employed--700 punch card operators, 200 verifying operators, 68 clerks, 12 typists, and 45 supervisors. Operators punched 647 cards per day, on an average. Verifiers averaged about 316 cards per day.

Mechanical Editing of the Punched Cards

Mechanical editing greatly simplified examination for errors. For most of the card types, unusual or inconsistent reports could be detected in one run through the Electronic Statistical Machine. Inconsistencies were disclosed by cross-classifying two or more entries on the card; unusual situations were found by comparing an entry on the card with a fixed number which was set in the machine. Some of the cards which were selected were for reports showing: bushels of wheat sold greater than bushels harvested; acres irrigated greater than acres in the farm; number of cows milked greater than number of milk cows; bushels of corn harvested for grain but no acreage; 10 or more bales of cotton per acre; pasture receipts of \$5,000 or more; and 10 or more hired workers.

Cards with impossible or misspelled codes were also selected during the mechanical edit. For example, only two entries in the column for telephone had meaning--a "1" if the farm had a telephone and a "0" if it did not. Any other numbers punched in that column were impossible codes, and the cards having them were withdrawn.

The selected cards were machine listed, and the listed items were examined for possible errors. Other items on the card sometimes explained the entry or indicated the correct one. If they did not, the questionnaire was examined. If changes were necessary, new cards were punched. If no change was necessary, an explanation was entered on the listing sheet. Then, if the figures were questioned in the critical examination of the tabulations, the listing sheets usually provided the explanation. For most counties, the number of cards selected did not exceed 7 percent.

All cards for large farms were listed and examined. Because these cards were to be used in the economic area tabulations, punching of codes on color-tenure, size of farm, economic class, and type of farm was checked. The totals for these groups were used as controls in the tabulations.

After the cards were corrected, they were again arranged by card type. They were then counted, and the counts were checked to the totals obtained after punching. Each card type group was then returned to the proper county file.

Tabulating the Agriculture Data

The cards punched from the Agriculture Questionnaires (A1) were sorted mechanically into the groups for which totals were needed--minor civil divisions, size of farm, etc. The cards for each county were tabulated as a unit, and each tabulation sheet carried separate totals for each group into which the cards had been sorted. Totals for the county were obtained by adding the printed totals for each minor civil division, each size of farm group, etc. The figures on the tabulation sheets were critically examined by the subject-matter specialists before they were transferred to tables.

⁴For discussion of different versions of the Agriculture Questionnaire (A1), see U. S. Bureau of the Census, *U. S. Census of Agriculture: 1950*, Vol. II, General Report, Statistics by Subjects, p. xiii.

After the county tabulations, the cards for large farms and sample farms were tabulated to produce data for each of the 363 State economic areas. These data were in four parts: color-tenure, class of farm, size of farm, and type of farm. Separate totals were printed for sample and large farms, and data for the sample farms were expanded by mechanical weighting techniques.

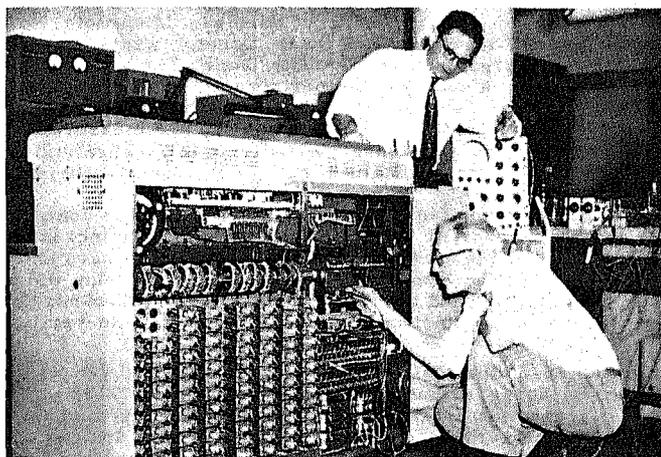
Fourteen tabulations were made from the landlord-tenant operations card. Two were used to get the characteristics of these operations by counties and twelve to get them by economic subregions. In the tabulations for subregions, the general characteristics were cross-classified with each of the following: color-tenure of operator, cropland harvested, kind of subunit, number of sharecropper subunits, size of unit, and type of farm.

Irrigation

Data on irrigation from the Agriculture Questionnaire (A1) and the Irrigation (I-1) Questionnaire were edited simultaneously. Where the reply to question 227 on the Agriculture Questionnaire indicated that the farm had its own supply works for irrigation, an Irrigation (I-1) Questionnaire was required. Related questions on the two forms were compared in the editing process and significant differences were reconciled. The Irrigation (I-1) Questionnaires were marked "Reject" if (1) the farm had less than one acre of irrigated land; (2) the irrigation works consisted only of a supply ditch or pipeline less than one mile in length, which was also used to obtain water from a partnership ditch, ditch company, or other irrigation enterprise; or (3) the irrigation supply works consisted only of a small reservoir or "overnight pond" that regulated the daily delivery of water from an irrigation enterprise.

The Agriculture Questionnaires for farms reporting irrigated land were then coded to indicate the supplier of water as follows: (1) farm operated own supply works and obtained water from no other source, (2) farm operated own supply works and also obtained water from another source, and (3) farm obtained all water from one or more irrigation enterprises and did not have own supply works.

The Irrigation (I-1) Questionnaires were edited for internal consistency and coded. Codes were entered to indicate the source of water (pumped well, spring, natural stream, etc.); whether all, part, or none was pumped; and whether all, part, or none was obtained from irrigation enterprises. A frequency distribution code for the number of irrigated acres was also entered. Data were then summarized and posted in the shaded boxes on the questionnaire.



Working on a machine in the Census Machine Development Laboratory.

Most of the Irrigation (I-2) Questionnaires were filled before the regular enumeration. When irrigation information from the Agriculture (A1) Questionnaires became available, the District Office made a follow-up card for each irrigation enterprise listed on those forms. Special irrigation enumerators then completed Irrigation (I-2) Questionnaires for those enterprises which had not been enumerated. The number of farms and the amount of irrigated land reported on the I-2 Questionnaire were compared with those data on the follow-up card. When the I-2 Questionnaire differed by one or more farms or by 10 percent of land, the District Office rechecked the information to reconcile the difference.

Complete field checking was not possible, however, so the Washington Office made further examinations. The I-2 Questionnaires were checked for internal consistency and compared to

the follow-up cards. Acreage by enumeration districts and by primary supplemental sources on the follow-up cards was adjusted to the information on the I-2 Questionnaire. Data were then posted to a work sheet, and maps showing irrigated land by counties and drainage basins were prepared.

Several codes were entered on the I-2 Questionnaire. These were for: type of enterprise; identification of drainage basin; type of water; whether or not the enterprise obtained 10 percent or more of the water from other enterprises, or used 10 percent or more of the water furnished by other enterprises; size of enterprise (number of acres irrigated); source of water; delivery of water direct or through other enterprises; and season of irrigation.

The checking procedure for large enterprises was more detailed than that for small ones, because large ones had to be divided into several parts to get statistics by counties and by drainage basins. To maintain an accurate count of enterprises, a special inclusion or exclusion code was used. This code prevented counting each part of an enterprise in a county or drainage basin as a whole enterprise.

Technicians separated parts of the enterprises and allocated the parts to the proper area on the basis of the location of the irrigated land in a county or drainage basin. Additional I-2 Questionnaires were prepared when the land irrigated by one enterprise lay in more than one county.

Comparisons with information from other sources were made to insure completeness and accuracy. Data on the I-2 Questionnaires for the larger storage reservoirs were compared with those in published lists of water-storage reservoirs. Representatives of the Bureau of Reclamation, which supplied part or all of the water for nearly 20 percent of the total irrigated acreage in the 17 Western States, checked the information obtained on its projects.

After the questionnaires were edited, they were arranged, numbered, and counted. The counts were posted on a label affixed to the portfolio.

Four cards (A, B, C, and D) were needed to record the data from the two irrigation questionnaires. Data from both questionnaires were punched on cards A, B, and C, but the D card was punched only from the I-2 Questionnaire. The cards and the number punched are shown in Table 3, Appendix D.

The data were tabulated by county, by drainage basin, by type of enterprise (single farm, mutual, commercial, etc.), by size of enterprise, and by type of water (surface, ground, etc.).

Drainage

The Drainage Questionnaire (D1) was used in the 10 States (Delaware, Indiana, Iowa, Kentucky, Michigan, Minnesota, North Dakota, Ohio, Oklahoma, and South Dakota) where drainage projects were organized on a county basis. The questionnaire was designed for a county as a reporting unit, and was especially edited to obtain the area of drained land in each county. The drained area was outlined on a county map and measured with a planimeter. This procedure yielded a more accurate figure than would have been obtained by asking county officials to estimate the area.

The planimeter reading of the drained area was compared with the area reported in the Census of 1940. The increase in acreage was then related to the amount spent by drainage enterprises between the censuses. If a direct relationship did not exist, both the map and the financial data were rechecked. Correspondence was sometimes needed to get data to verify or correct the maps.

In the 30 States that had organized drainage districts, the number of acres drained by each enterprise of 500 or more drained acres was entered on the D2 Drainage Questionnaire. The enumerator then plotted the location of the drained area on a map of the county. These areas were also measured with a planimeter during the editing, and the computed acreage was compared to the amount entered on the D2 Questionnaire to see if the two figures were in reasonable agreement.

The D2 Questionnaires were examined for consistency and completeness. Letters were written to drainage officials and others in an attempt to supply missing information and to resolve discrepancies for the larger enterprises.

Statistics for each county and State were compared with those from the 1940 Census of Drainage. Location of the drained areas in the two censuses was also compared. Most of the large discrepancies were discovered by testing a simple relationship. In this relationship, it was assumed that the indebtedness at the beginning of the period plus construction and maintenance costs during the period should be approximately equal to the tax collections and special assessments during the period plus the outstanding indebtedness at the end of the period.

Data were posted to county and State tables from the questionnaires. Detailed drainage data were hand-tabulated, using intermediate transcription sheets.

Post-Enumeration Survey

The Post-Enumeration Survey (PES) was designed primarily to check the completeness of coverage and the accuracy and consistency of response in the Censuses of Population, Housing, and Agriculture. The PES enumeration took place after the censuses had been taken, but before the field organization had been dispersed.

PES questionnaires were edited in the field as well as in Washington. A specially trained PES editor in each field office examined all questionnaires, chiefly for completeness. In Washington, the questionnaires were reviewed, again primarily for completeness. If faulty questionnaires were discovered, they were returned to the field for completion or correction.

Following the initial editing operations, selected data were transcribed from the PES questionnaires so that "record check" studies could be made. In these studies an attempt was made to match information on the PES questionnaire with information for the same person on the records of other government agencies. Information from the following records was compared with the PES data: (1) birth certificates (to study age), (2) 14th Decennial Census records (1920) for persons 30 years old and over (also to study age), (3) records of the Immigration and Naturalization Service (to study citizenship of foreign-born persons), (4) records of the Veterans Administration (to compare veteran status entries), (5) records of the Bureau of Old-Age and Survivors Insurance (to study wage and salary income and industry), and (6) records of the Bureau of Internal Revenue (to study income).

The record checks were made by transcribing identification data from the PES questionnaires and sending the transcripts to the cooperating agency. That agency entered on the transcript the information it had for that person. The transcript was then returned to the Census Bureau where corresponding data from the Census and the PES were entered on it. Stringent restrictions were imposed to maintain the confidential character of the data in these operations.

After the record check transcripts were prepared, the PES questionnaires were given a final editing, and selected entries were coded for punching. Census questionnaires were then matched with the accompanying PES questionnaires and the Census codes for selected characteristics were transcribed to the PES questionnaires. When a household, person, or farm was missed in the Census and enumerated in PES, the Census schedules were searched to be sure of the omission.

The last operation before punching was to code such information as results of the record checks and reasons for difference in Census and PES agriculture reports.

Seven punch cards were needed for the population and housing data from the Decennial Census, the PES, and the records of cooperating agencies. The data on these cards were transferred to magnetic tape and were processed on the Univac. The resulting tabulations provided a basis for estimating the completeness of the Census coverage and the accuracy of the Census data.

Agriculture data required four punch cards. In the first series of tabulations, the amount of difference between the PES and the Census data was measured. From the second series of tabulations, analytical tables were prepared.

Mechanical Equipment

Many types of machines were needed to process the data collected in the 1950 Censuses. Some of them were developed and built by the Bureau of the Census; others were rented from International Business Machines Corporation (IBM). The machines available and used are listed in Table E.

The functions performed by the mechanical equipment, from a subject-matter point of view, are determining recodes, calculating, editing, and tabulating. The three main operations required to carry out these functions include the preparation of the punch cards; the arrangement of the cards into groups for mechanical editing, tabulating, or reproducing; and the tabulating to permit editing or to prepare final results. For convenience, the machines are grouped according to these three operations, but it will be evident that some machines are used for more than one purpose.

Preparing Punch Cards

Holes may be punched in the card in several ways. A punch machine operator may read the information from the schedule and punch it in the card. A machine can reproduce automatically

April 24, 1981

History Files

FBohme

WGBH-TV, through John Hopper, CSVD, asked the cost of UNIVAC I, installed.*

Joe Daly, citing Morris Hansen's testimony at House hearings in 1954 (p. 65), reports that the official contract cost was \$436,625, and totalled \$701,000, installed.

Installation included the cost of peripheral equipment, including air conditioning and a standby power supply, as well as engineering and tapes. Morris Ullman reports that a good deal of this expense was absorbed by the participating divisions, and the whole thing probably totalled \$1.25 million. Eli Marks agrees that this is a reasonable figure.

Gave Hopper the \$701,000 figure.

*James Mc Pherson
129 Taylor
Paris, Ky. 40361*

Ullman 320-3470

Daly 577-0319

Marks 657-3126

** Delivery, accepted March 31, 1951.*

in one card holes punched in another. Another machine can determine a recode from combinations of entries punched in the card and punch the recode in another column of the card. Another will divide one entry by another and punch the result in the card. One machine can be set to punch the same entries in a group of cards. Another can punch automatically a summary card giving totals for a number of cards. These machines are described in more detail below.

Individual cards were punched on electrically-powered, manually-operated machines (see punch cards in Appendix B). The keypunch machine used most in the Population and Housing Censuses was the IBM Numerical Punch #016. This machine has 14 keys, of which 12 are for the different punching positions in each column of the card (0 to 9, X, and V), and are electrically actuated. (One key is for spacing and the other is for ejecting an incomplete card.) The card is automatically fed into the machine so that column 1 is punched first, then column 2, etc. A light depression of a key causes a punch to be electrically driven through the card, cutting a hole in the column under the punch. As each column is punched, the card is advanced automatically to the next column. As soon as column 80 is punched, the card is ejected and stacked, and the next card automatically enters the machine. An "X-skip-bar" arrangement enables the punch operator to skip several columns for which data are not to be entered.

The keypunch machine used most for the Census of Agriculture was the IBM Numerical Punch #024. This machine is operated in the same way as the #016, but it is newer and faster. Skipping and duplicating are controlled by a punched program card which is mounted on a program drum.

In punching population data from the Population and Housing (F1) Schedule, a special schedule holder (the "Richards" copyholder) was used to hold the schedule and indicate the line of information being punched. Every fifth person enumerated on that schedule was in the sample, and an additional line of information for him appeared at the bottom of the schedule. Shifting from the regular schedule line to the sample information at the bottom and then back to the next regular schedule line was done automatically with the Richards copyholder. The more conventional "Pres-to-line" copyholder was also used to hold questionnaires during the punching operations. A special method of inserting the population schedule to simplify the punching of the sample information was devised. The agriculture copyholder was a special container developed by the Census Bureau to hold the agriculture questionnaires. This was used with a series of special "masks" which showed the data needed for the particular card type for which the mask was designed.

Punching was verified on Census-owned mechanical verifiers, which are similar to card-punching machines. The operator depresses the keys as in punching, and plungers pass through the holes in the card and permit it to advance through the machine. If the keys depressed by the operator do not agree with the holes already in the card, the card does not advance and the verifying operator checks the punched entry against the questionnaire. If the original punch was wrong, an entirely new card is punched.

The IBM Reproducing Punch #514 reproduces in one card the holes punched in another card. It reads all or part of the information in one set of cards and punches it in the same or a different arrangement into another set of cards. It can also be used for gang-punching, in which information common to a group of cards is punched into each card in the group. In the 1950 Censuses, this machine made duplicate sets of cards so that one set could be used to make regular tabulations and the other to make special tabulations.

The Census Recode Machine was developed by the Bureau of the Census with the assistance of the National Bureau of Standards. It was constructed in the Census Bureau's laboratories. In determining recodes and in punching and verifying them in one operation, it does the work formerly done by three separate machines, and it punches at more than four times the speed of other punching machines. It can also edit and select cards which have a size or value which is higher than, lower than, or equal to a certain figure. It can detect cards with blank columns and cards having columns with double punches.

The IBM Electronic Calculating Punch #604 adds or subtracts amounts punched in the same card or in successive cards. It also multiplies and divides and is used to compute averages and percentages. The results are punched automatically in the same or different cards. For example, this machine was used to compute and punch the gross rent recode, for which rent on an unfurnished basis was added to costs of different utilities.

The Census Gang Punch Machine punches codes in groups of cards which require the same punched information. Although the number of columns which can be punched in one run is limited to 18, cards are punched at a rate of 400 per minute--

four times the rate of the IBM Reproducing Punch. Cards are fed end-wise into the machine and can be punched in the first 18 columns on either end of the card. The Census Gang Punch Machine was used to punch the recodes for employment status in the population P card.

When the IBM Duplicating Summary Punch Machine is attached to an IBM Electronic Statistical Machine, it punches totals in a summary card at the same time that the totals are printed on the tabulation sheet. Thus, summary cards are punched for every printed line of information. Running summary cards instead of the individual cards greatly reduced the number of card passes needed to obtain the final tables for the 1950 Censuses.

Arranging the Cards

The punched cards are usually subdivided into groups to simplify obtaining the cross-classifications needed. Population cards, for example, are often sorted by sex before they are tabulated; then each sex group is tabulated separately. Sometimes cards are sorted for editing purposes to disclose inconsistencies or impossible codes. In another type of arrangement, two sets of cards are matched, card by card, before the information from one set is reproduced on the other.

The IBM Sorter #082 sorts the cards according to the entries in one column on the punch card. It has 13 pockets; 12 of them correspond to the 12 punching positions in a card column (0 to 9, X and V). All cards punched "1" in the selected column go into the "1" pocket, all those punched "2" go into the "2" pocket, and so on. The 13th pocket is for rejected cards--those which have no entry in the column.

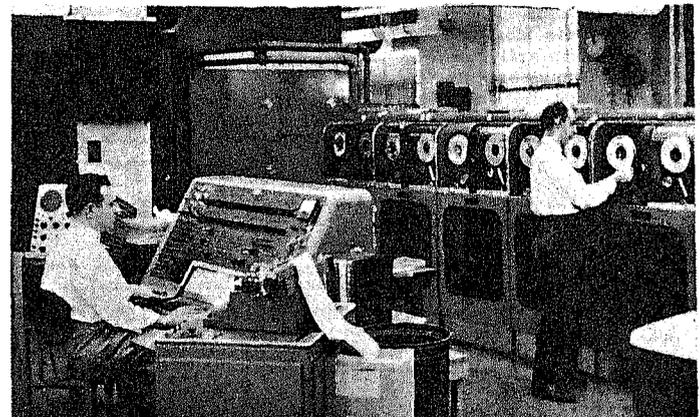
The Census Multi-Column Sorter #488 was developed by the Bureau of the Census for the 1950 Censuses. It sorts cards into groups on the basis of: (1) entries in one or more columns on the punch card, (2) comparison of information punched in one part of the card with information punched in another part, or (3) comparison of a number on the card with a number set in the machine. Acceptable cards are sorted into 12 pockets and rejected cards into a 13th. Thus, the machine sorted the cards into groups for the employment status recode from information punched in 10 columns on the card. It was used in mechanical editing to detect inconsistencies between two entries on a card and to locate cards with impossible codes. The machine can sort cards into groups which are higher than, lower than, or equal to a determined figure. It can also select every Nth card to obtain a sample.

The IBM Collator #077 compares two sets of data punched on a card or it compares one set with a constant; then it separates cards which do not meet the specifications wired into the machine. The machine matches one set of cards with another, merges two sets of cards into one set, selects cards with specified characteristics, and checks sequences. It was used to match housing H and PH cards so that information could be transferred from one to the other.

Tabulating the Cards

All tabulating machines print figures showing either the number of persons, farms, or dwelling units in each classification or the indicated totals. Four types of tabulating machines used for the 1950 Censuses are described below.

The Census Unit Count Machine selects cards, counts characteristics, and prints the results. Its reject mechanism permits



Univac, which was first used for tabulating part of the 1950 censuses.

THE CENSUS OPERATION

Table E.--Machines Used for Punching and Tabulating Operations, 1950 Censuses

Machine type	Maximum number	Census of Population	Census of Housing	Census of Agriculture
IBM Numerical Punch #016.....	1,445	1,015	459	46
IBM Numerical Punch #024.....	691	167	20	530
IBM Alphabetical Duplicating Punch #031	11	8	3	-
IBM Punch Card Verifiers #055.....	159	6	2	151
Census Verifiers #280.....	709	364	210	135
Census Unit Count Machine #581.....	32	27	24	-
Census Multi-Column Sorter #488.....	28	27	15	2
Census Gang Punch.....	12	8	4	-
Census Recode Machine.....	1	1	-	-
IBM Duplicating Summary Punch #524.....	84	30	30	34
IBM Sorter #082.....	103	23	45	36
IBM Electronic Statistical Machine #101	45	28	17	19
IBM Alphabetic Accounting Machine #402.	60	12	24	35
IBM Accounting Machine #407.....	5	1	-	5
IBM Electronic Calculating Punch #604..	6	1	5	1
IBM Reproducing Punch #514.....	82	20	40	34
IBM Collator #077.....	33	4	16	13
IBM Alphabetical Interpreter #552.....	1	1	-	-
Richards Copyholder.....	900	900	-	-
Pres-to-line Copyholder.....	900	491	535	-
Agriculture Copyholder.....	785	-	-	785

it to be used for mechanical editing. The machine has three pockets: the first is for rejected cards, the second is for cards that have been counted, and the third is for cards especially selected. The machine is wired to count cards according to certain characteristics--such as relationship to household head and age--and the 60 counters will print totals for 60 different groups. The card passes into the machine and stops momentarily while those characteristics are read from the punched holes. If the card fulfills the conditions for one of the groups--such as parent of household head, aged 50 to 54--it is counted in that group and goes into the second pocket. If the card does not fulfill the conditions for any of the groups--such as parent of household head, aged 5, for which no group was established--it is rejected and falls in the first pocket. Rejected cards can then be examined and corrected. The machine can also select every Nth card for sampling purposes, and the selection can vary from the 2nd to the 110th card. Moreover, the machine can select the Nth card in a specified group, such as every 15th card for male white married persons or every 10th card for female nonwhite widowed persons. It can also make selections regardless of spacing, such as every 3rd, 10th, 16th, 25th, and 32nd card in a group of 32. With the installation of special "sample wheels" or counters, the machine can be used to inflate sample results to 100 percent. A 20 percent sample wheel, for example, will count 5 for each card.

The IBM Electronic Statistical Machine #101 tabulates, edits, and sorts in one card pass through the machine. It can select cards with inconsistencies between entries, with impossible codes, or with entries higher than, lower than, or equal to a certain figure. In tabulating, it prints not only the totals for the selected groups, but also an identification of each group. Its self-balancing feature enables it to indicate on the tabulation sheet any discrepancy between the sum of the individual columns and the total count. This machine was used extensively in the mechanical edit of housing and agriculture items.

The accounting tabulating machine is designed primarily to accumulate data shown on the cards while the statistical machine primarily makes count distributions of the characteristics punched in the cards. The Alphabetic Accounting Machine #402 lists and tabulates. In listing, the machine prints either alphabetic or numeric information. Information on the

punch card may be read, added, subtracted, compared, or selected, and then printed in the proper place on the report form.

The development and testing of the Univac (Universal Automatic Computer) was the most far-reaching innovation in automatic tabulating equipment for the 1950 Censuses. This machine did not become available until late in the tabulating program and was used for only a small part of the population and housing tabulations. Instructions and data are coded on an 8-channel magnetic tape which is fed into the computer. Information from punched cards can be transferred to that tape by means of a "Card to Tape" machine. After the data are in the Univac, it will duplicate, sort, arrange in sequence, merge, and calculate automatically according to the prepared instructions. Calculations are recorded on other tapes, which are processed on the "Uniprinter." The Uniprinter translates the codes on these tapes and types the results.

Information which requires several runs through punched card equipment can be obtained from the Univac in one run. For example, in the 1950 Census of Population, the Series B tabulations for some States were prepared on the Univac. In one run, the Univac prepared tables which on the punched-card equipment required five separate tabulations, a manual rearrangement, and two sorting operations. With machines like the Univac, future censuses should be processed with considerably greater speed.

Testing the Machines

Each machine used in the 1950 Censuses was tested daily in a specified manner to insure accurate mechanical operation. For some machines, these tests were simple; but for others, they were more complicated. The tabulating machines were usually tested at the beginning, the middle, and at the end of each day. The test was made with a set of punched cards especially prepared for each tabulation wiring. A master tabulation sheet gave verified results of running the test deck. If a run of the cards produced figures different from those on the master sheet, operations were discontinued until the trouble was found, and the machine was again operating correctly.