

*EVALUATION OF THE HOUSING CENSUS*

Arthur F. Young and staff of the Housing Division

*THE HOUSING EVALUATION PROGRAM - 1950 TO 1990*

The Census Bureau conducted decennial censuses on a regular basis from 1790 to 1940. Although the completeness of the census counts and the accuracy of census data concerned census officials and data users from the very first census, no formal research or evaluation program was a part of a decennial census until after the 1940 Census. During the 1940s, the Census Bureau was confronted with data users needing statistical facts for decision-making. This need for data, accompanied by a growing concern for the accuracy and completeness of census results, led to the development of a formal research and evaluation program that was incorporated into the 1950 decennial census.

The evaluation objectives for the 1950 Census, and all later censuses, have been to provide information for planning future censuses and for making users aware of the sources and magnitude of error in the census.

The housing evaluation programs for the 1950, 1960, 1970, and 1980 censuses were concerned with measurement of coverage error resulting from the erroneous omission or inclusion of housing units in the census. There was also concern for content or response error resulting from incorrect responses to census questions and from errors in recording and processing the responses. Response error is usually defined in terms of gross error (response variance) or net error (response bias). It is measured through re-interview surveys conducted following the census. Definitional problems in determining which vacants should be included/excluded in the housing inventory have limited the bulk of evaluation studies to occupied housing units.

*1950 Census*

The primary method of evaluating both housing coverage and content was through the 1950 Post-Enumeration Survey (PES).

1. Housing Coverage - The 1950 PES measured the completeness of the housing count both in terms of coverage error and errors in applying the definition of a dwelling unit. The results indicated a net undercount of occupied units of 2.5 percent (1.1 million units), with 2.2 percent resulting from missed dwelling units and 0.3 percent from errors in applying the definition of dwelling unit.

---

Arthur F. Young is Chief, Housing Division, U.S. Bureau of Census.

2. Housing Content - The PES evaluated the following housing items: tenure, rooms, water supply, bathing facilities, condition, contract rent, gross rent, and value. Although the results showed that the gross error was high for some items, particularly for value and for gross rent, the net errors for most items were relatively small and the impact on the calculation of medians was minimal.

#### *1960 Census*

Housing coverage and content in 1960 were again evaluated through a re-enumeration of housing units. Randomization studies were also conducted to evaluate the between-interviewer effects for selected data items.

1. Housing Coverage - Coverage evaluation indicated that 3.3 percent of all housing units (1.9 million units) were missed in 1960, with 2.1 percent of occupied units and 11.7 percent of vacant units being missed.
2. Housing Content - The Bureau conducted a separate Content Evaluation Survey (CES) for the reporting of the following characteristics of occupied housing units: rooms, condition, condition and plumbing facilities, access and cooking equipment, bathrooms, bedrooms, bathing facilities, toilet facilities, and water supply. The results showed relatively high indexes of inconsistency (a measure of gross error) for several housing items, particularly for condition of plumbing facilities. However, net difference rates for all items were relatively low.

#### *1970 Census*

The 1970 Census evaluation and research program included a number of studies designed to measure coverage and content errors. Again, randomization studies were conducted to determine between-interviewer effects for selected data items.

1. Housing Coverage - The results of three separate studies conducted to evaluate housing coverage in 1970 showed that 2.5 percent of all housing units were missed (1.7 million units), with 1.4 percent of occupied units and 11.8 of vacant units being missed. The net undercount of occupied units in mail census areas was 1.1 percent.
2. Housing Content - The Bureau conducted a Post-Enumeration Survey (PES) for content from which an analysis of response variance was done for the following selected housing characteristics: heating

fuel, cost of gas and electricity for renters (when paid separately), bathtub or shower facilities, flush toilet facilities, telephone, year structure built, seasonal vacancy status, renters paying extra for water, renters paying extra for fuels and value.

The distributions for the first five items had fairly low levels of response variance; the second five had moderate levels with value having a high response variance. Generally, owners reported more consistently than renters. Responses for occupied units were more consistent than vacant units with respondents in single-unit structures reporting more consistently than those in multi-unit structures.

A response bias analysis conducted for tenure, contract rent, units in structure, bedrooms, piped water, kitchen facilities, heating equipment and rooms, indicated moderate-size biases in the distributions of bedrooms and heating equipment.

Two record check studies were conducted following the 1970 Census: (1) Value of Home Study and (2) Gross Rent Record Check Study. The value study showed a significant understatement of value in the census of between \$160 and \$1,512 on a median of \$27,200.

The gross rent record check showed a substantial overstatement of costs for electricity (15-30 percent) and utility gas (25-50 percent) in the central cities of 5 selected SMSA's. The net effect on gross rent was small, however, from 1.5 to 4.0 percent on median gross rents ranging from \$85 to \$143 in the five cities.

#### *1980 Census*

The 1980 Census research and evaluation program consisted of approximately 40 formal studies conducted during and immediately following the 1980 Census.

1. Housing Coverage - Coverage of housing in the 1980 Census will be estimated from two studies; the Housing Unit Coverage Study and the Housing Unit Overcount Study. Final results are not yet available, but preliminary data show a total missed rate for occupied units comparable to the rate in 1970.

2. Housing Content - The primary vehicle for measuring the accuracy of responses to the 1980 housing items was the content re-interview study. The Bureau conducted a response variance (gross error) analysis for the following housing characteristics:
  - a. tenure
  - b. heating fuel
  - c. mortgage
  - d. year structure built
  - e. property size
  - f. condominium unit
  - g. structure size
  - h. seasonal vacancy status
  - i. property usage (acreage and presence of commercial establishment)
  - k. unit condition (boarded up) status

The study showed that the distributions for the first 3 characteristics had low levels of response variance, and the distributions for the next 5 had moderate levels. The property usage characteristic showed a level in the low end of the high range. The question concerning whether the vacant unit was boarded up had too few cases to make a reliable estimate.

A response bias (net error) analysis, conducted for heating method, plumbing facilities, number of automobiles and number of vans and trucks, showed a moderate level of inconsistency for the census distribution on heating method, with some bias evident in all categories except "no heating equipment." Plumbing facilities had a moderate level of inconsistency with relatively low levels of response bias. The number of automobiles showed a moderate level of inconsistency with significant bias existing in all categories in the distribution. The distribution for number of vans or trucks displayed a level of inconsistency in the high range with substantial bias in most categories.

3. Utility Cost Study - A major study of utility cost reporting was undertaken in conjunction with the 1980 Census. Previous studies indicated that estimates of the average monthly cost of gas and electricity were subject to relatively large response biases (net overreporting) and that the size of the bias varied considerably from area to area.

For this study utility companies in 8 cities were asked to provide information on "average monthly costs" in their March, 1980 bills to a sample of

Young and staff

residential customers. In most of the cities, a significant improvement in electricity and gas reporting occurred when respondents were supplied with average cost information.

#### *Housing Coverage And Content Evaluation In 1990*

The Bureau is now planning its coverage and content evaluation programs for the 1990 Census. The goal is to develop an efficient and effective program which will benefit both the Bureau, in its planning for the year 2000, and data users who are concerned with the accuracy of coverage and content. In order to do this, answers to the following questions are needed from housing data users.

1. What measures of housing coverage are most useful, gross measures of the over- and undercount? Net measures?
2. Past estimates of housing coverage error have been made at relatively large geographic levels. Is this sufficient, or should coverage be measured at lower levels of geography?
3. Is it important to know the components of the overcount or undercount, such as whether the whole structure was missed or all units in the structure were properly counted? Were owners or renters missed? Was the missed rate different for occupied or vacant?
4. For which characteristics are measures of accuracy most needed?
5. Which measures of accuracy and reliability are most useful to data users? Measures of gross error? Measures of net error?
6. In the past, measuring data quality involved a post-enumeration re-interview survey, a record check using administrative records or a match with records from a sample survey. Are there other means by which the quality of the coverage and content data collected in the census can be determined?

#### *ADJUSTMENT OF THE CENSUS*

Those charged with taking the census, from the very first census in 1790 to the present, have been aware of the difficulty in both finding and accurately counting every person in the United States. The undercount associated with the decennial census is not a new problem. What is new is the pressure the Bureau is receiving to adjust the census to correct for the undercount. The pressure to adjust comes from many different directions: the use of census data to distribute federal funds to state and local governments; the use of

census data for redistricting and reapportionment and uses of census data in the one person-one vote court decisions and voting rights laws. The Census Bureau, although it has estimated the undercount for the past several censuses, has never adjusted the census counts. Based on (1) the quality of the 1980 Census (estimates indicated that the net undercount was very small, and (2) the absence of any accurate measure of the number and distribution of illegal aliens, the Bureau decided in December, 1980 not to adjust 1980 Census counts. Since 1980, the Bureau has continued to do research on coverage measurement and the adjustment issue realizing that pressure to adjust the 1990 Census will be present and that a decision to adjust in 1990 may be made by persons outside the Bureau.

Usually when people talk of measuring the undercount or adjusting the census, they are referring to the population counts and characteristics. However, coverage errors also affect census housing data. Attention must be given to the effects of these errors or to the impact that adjusting the population counts and characteristics would have on housing counts and characteristics.

Therefore, as part of the planning process for the 1990 housing census, some basic questions need to be addressed:

1. What are the main uses of housing data, particularly those that involve funds distribution?
2. What effect does the housing undercount have on the uses of these data?
3. If the population count is adjusted, how should it be done? By adding occupied housing units? By adding people to existing housing units? By adding people to group quarters? By some other means?
4. If the population counts are adjusted, but the housing counts are not, what impact would there be on housing content and data relationships? Would the effect be harmful on the following characteristics: persons in housing units, persons per room, household composition, gross rent and shelter costs as a percent of income?
5. If the housing counts are adjusted, should only the total count be adjusted or sub-universes also, for example, by occupied/ vacant? By year-round seasonal? By tenure? By race and ethnicity? Should each characteristic be adjusted?
6. If housing counts were found to be accurate, but some of the characteristics data were in error, should these characteristics data be adjusted based on evaluation studies? At what geographic levels?
7. At what geographic levels should the housing counts be adjusted? U.S. only? States? Cities? Census tracts? Other levels?

Young and staff

8. If the housing counts are adjusted, how should the adjusted figures be published and used? Should the adjustment be the official count with no other number published? Should both the adjusted and the census counts be published?
9. Are there other implications of the adjustment issue that need to be addressed?