

## Appendix B. Source and Accuracy of the Data

SAMPLE DESIGN .....	App-53	Combined sample ratio estimation procedures.....	App-61
Introduction.....	App-53	Mobile home ratio estimation.....	App-61
Designation of AHS-MS sample housing units for the 1989 Survey.....	App-54	Independent total housing unit ratio estimation.....	App-61
1989 AHS-MS original sample selection for the 1970-based area sample of the metropolitan areas .....	App-55	ACCURACY OF THE ESTIMATES.....	App-62
Sample from the 1970-based permit-issuing universe .....	App-55	Nonsampling errors.....	App-62
Sample from the 1970-based new construction universe.....	App-56	AHS-MS content errors.....	App-62
Sample from the 1970-based nonpermit universe .....	App-56	AHS-National content errors .....	App-62
Sample selection for the AHS-MS Coverage Improvement Program .....	App-56	Coverage errors .....	App-63
1985 AHS-MS sample reduction and sample reinstatement.....	App-56	Rounding errors .....	App-63
AHS-MS sample selection for the 1980-based area sample of the metropolitan areas .....	App-57	Sampling errors for the AHS combined sample estimates.....	App-63
Sample from the 1980-based permit-issuing universe .....	App-57	Illustration of the use of the standard error tables .....	App-64
Sample from the 1980-based new construction universe.....	App-58	Differences .....	App-65
Sample from the 1980-based nonpermit universe .....	App-58	Illustration of the computation of the standard error of a difference .....	App-65
1989 AHS-MS sample reduction and sample reinstatement.....	App-58	Medians .....	App-65
AHS-National sample selection.....	App-58	Illustration of the computation of the 90-percent confidence interval of a median .....	App-66
Selection from the 1980 census .....	App-58	Standard error tables.....	App-67
Selection of new construction housing units in permit-issuing areas .....	App-58		
Selection of units from the nonpermit universe .....	App-58		
ESTIMATION .....	App-59		
AHS-MS .....	App-59		
Type M noninterview adjustment.....	App-59		
Type A noninterview adjustment .....	App-59		
Ratio estimation procedure for the 1970-based permit-issuing universe .....	App-59		
Ratio estimation procedure for the 1980-based permit-issuing universe .....	App-60		
AHS-National.....	App-60		
COMBINED SAMPLE WEIGHTING .....	App-60		
Introduction.....	App-60		
Weighting factor adjustment .....	App-60		

### SAMPLE DESIGN

#### Introduction

The estimates for each of the 11 metropolitan areas in this report series (H170/89) are based on data collected from the 1989 American Housing Survey Metropolitan Sample (AHS-MS) and the 1989 American Housing Survey National Sample, which were conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development.

The sample areas covered for metropolitan areas that remained in the AHS sample after survey year 1983 are consistent with the 1983 Office of Management and Budget (OMB) definitions of a metropolitan statistical area (MSA), consolidated metropolitan statistical area (CMSA), or primary metropolitan statistical area (PMSA). In some instances, a given metropolitan area is a combination of primary metropolitan statistical areas and will be referred to as PMSA's. In addition to adding new areas to some

metropolitan samples to comply with the 1983 definitional changes, some new metropolitan areas have been added. Thus, each of the 1989 metropolitan areas will fall into one of three categories:

- a. Areas of the same geographic area as defined for surveys prior to 1984 (i.e., areas in which the 1970 OMB definition of a standard metropolitan statistical area is the same as the 1983 MSA, PMSA, or CMSA definition, 1970-based area)— Dallas, TX PMSA; Los Angeles-Long Beach, CA PMSA; Philadelphia, PA-NJ PMSA; Phoenix, AZ MSA; and San Francisco-Oakland, CA PMSA.
- b. Areas consisting of new area in addition to the 1970-based area—Boston, MA-NH CMSA; Detroit, MI PMSA; Ft. Worth-Arlington, TX PMSA; Minneapolis-St. Paul, MN-WI MSA; and Washington, DC-MD-VA MSA.
- c. Areas that are strictly 1980-based—Tampa-St. Petersburg, FL MSA.

The metropolitan areas selected for the 1989 AHS-MS are interviewed on a rotating basis once every 4 years. Initially, each metropolitan area had an expected sample size of 4,250 or 8,500 housing units, uniformly distributed throughout nine panels (panels 4 through 12). Because of budget constraints, metropolitan areas that had an expected sample size of 8,500 were reduced to 4,250, and panels 11 and 12 were dropped from sample in all metropolitan areas. Hence, the expected sample sizes were lower than the original goal of 4,250 and 8,500 sample units. For all of the 1989 MSA's except Detroit, interviewing was scheduled for April 1989 through October 1989. In Detroit, the interviewing was conducted from May 1989 through August 1989.

In this metropolitan area, 3,405 AHS-MS housing units were eligible for interview. Of these sample housing units, 166 interviews were not obtained, because for occupied sample units, the occupants refused to be interviewed, were not at home after repeated visits, or were unavailable for some other reason; or, for vacant units, no informed respondent could be found after repeated visits. In addition to the AHS-MS housing units eligible for interview, 258 units were visited but were not eligible for interview because they were condemned, unfit, demolished, converted to group quarters use, etc.

The AHS-National sample is interviewed biennially in odd-numbered years. The interviews were a combination of field interviewing and CATI (computer assisted telephone interviews). CATI interviewing was conducted from July 1989 through October 1989 and field interviewing from September 1989 through December 1989. The sample covers 878 counties and independent cities with coverage in each of the 50 States and the District of Columbia. To increase the reliability of the AHS-MS sample estimates, information from AHS-National sample units was used in the estimation process. For each metropolitan

area, interviewed AHS-National units that were located within the 1989 AHS-MS definition of the metropolitan area were used in the estimation procedure. In this metropolitan area, 761 AHS-National units were used.

### **Designation of AHS-MS Sample Housing Units for the 1989 Survey**

The sample housing units designated to be interviewed in the 1989 survey consisted of the following categories, which are described in the following sections.

Housing units that were in the 1970-based area include the following:

- a. All sample housing units that were interviewed in the previous survey and remained in sample after the 1989 reduction. This sample includes housing units that were selected as part of the 1976-1981 Coverage Improvement Program. These coverage improvement cases represented most of the housing units that, until these procedures were implemented, did not have a chance of selection.
- b. All sample housing units that were type A noninterviews (i.e., units eligible to be interviewed) or type B noninterviews (i.e., units not eligible for interview at the time of the survey but which could become eligible in the future) in the previous survey and remained in sample after the 1989 reduction. (For a list of reasons for type A noninterviews, see the facsimile of the 1989 AHS questionnaire, page App-25.)
- c. All sample housing units selected from a listing of new residential construction building permits issued since the previous survey that remained in sample after the 1989 reduction. This sample represented the housing units built in permit-issuing areas since the previous survey.
- d. All sample housing units that were added since the previous survey in sample segments from the nonpermit universe that remained in sample after the 1989 reduction. This sample represented additions to the housing inventory since the previous survey in nonpermit-issuing areas.
- e. In the 1970-based areas of the Boston, MA-NH CMSA; Detroit, MI PMSA; Ft. Worth-Arlington, TX PMSA; Minneapolis-St. Paul, MN-WI MSA; and Washington, DC-MD-VA MSA, all sample housing units selected from the 1980 Census of Population and Housing.
- f. All sample housing units reinstated in sample in 1989. This sample represents units that had been dropped from sample because of sample reductions prior to 1989.

Housing units within new areas added to the metropolitan area in 1980 and metropolitan areas that are in sample for the first time (1980-based area) include the following:

- a. All housing units selected from the 1980 Census of Population and Housing that remained in sample after the 1989 reduction.
- b. All housing units that were selected from a list of new residential construction building permits and remained in sample after the 1989 reduction. This sample represented the housing units built in permit-issuing areas since the 1980 census.
- c. All sample housing units that were selected in sample segments added from the nonpermit universe and remained in sample after the 1989 reduction. This sample represents units enumerated in the 1980 census as well as additions to the housing inventory in nonpermit-issuing areas since the 1980 census.

The following table shows the percent of the AHS-MS old construction sample that is 1970-based and 1980-based for each metropolitan area:

Metropolitan area	Percent 1970-based area	Percent 1980-based area
Boston, MA-NH CMSA.....	70.1	29.9
Dallas, TX PMSA.....	100.0	0.0
Detroit, MI PMSA.....	91.7	8.3
Ft. Worth-Arlington, TX PMSA.....	96.2	3.8
Los Angeles-Long Beach, CA PMSA.....	100.0	0.0
Minneapolis-St. Paul, MN-WI MSA.....	91.6	8.4
Philadelphia, PA-NJ PMSA.....	100.0	0.0
Phoenix, AZ MSA.....	100.0	0.0
San Francisco-Oakland, CA PMSA.....	100.0	0.0
Tampa-St. Petersburg, FL MSA.....	0.0	100.0
Washington, DC-MD-VA MSA.....	93.3	6.7

### 1989 AHS-MS Original Sample Selection for the 1970-Based Area Sample of the Metropolitan Areas

The 1989 AHS-MS original sample for the 1970-based area of the metropolitan areas was selected from two frames: (a) housing units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the 1970-based permit-issuing universe) and (b) housing units constructed in permit-issuing areas since the 1970 census (the 1970-based new construction universe).

In addition, the sample for those metropolitan areas that were not 100-percent permit-issuing in 1970 included a sample selected from a third frame: housing units located in areas not under the jurisdiction of permit-issuing offices (the 1970-based non-permit universe).

In 1970, the Boston, MA-NH CMSA; Los Angeles-Long Beach, CA PMSA; Phoenix, AZ MSA; San Francisco-Oakland, CA PMSA; and Washington, DC-MD-VA MSA were the only metropolitan areas that were 100-percent permit-issuing.

Sampling operations, described in the following paragraphs, were performed separately within the central city and balance using the 1970 OMB definitions of the central city of each metropolitan area for each of the sample frames. The overall sampling rate used to select the sample for each metropolitan area was determined by the size of the sample. Each metropolitan area had a sampling rate about the same for the central city and the balance, since the sample was distributed proportionately between the two, according to the corresponding distribution of total housing units.

### Sample from the 1970-based permit-issuing universe.

The major portion of the sample in each of the metropolitan areas was selected from a file that represented the 20-percent sample of housing units enumerated in permit-issuing areas of the metropolitan areas during the 1970 Census of Population and Housing. This file contained records for occupied housing units, vacant housing units, and housing units in certain special places or group quarters. Sampling operations were done separately for the special place and group quarters records and for the occupied and vacant housing unit records. Before the sample was selected from the occupied and vacant housing unit records, the occupied records were stratified by race of the head of household (non-Black/Black), and the vacant records were stratified into four categories pertaining to the value or rent associated with the vacant housing units. The occupied housing unit records were further stratified so that each unit was assigned to one of 50 strata according to its tenure (owner/renter), family size, and family income category as illustrated by the following table:

Family income	Tenure									
	Owner family size					Renter family size				
	1	2	3	4	5+	1	2	3	4	5+
Under \$3,000.....										
\$3,000 to \$5,999.....										
\$6,000 to \$9,999.....										
\$10,000 to \$14,999.....										
\$15,000 and over.....										

Thus, the occupied housing unit records from the permit-issuing universe were assigned to 1 of 100 strata for either the central city or for the balance, and the vacant housing unit records were assigned to one of the four vacant strata for either the central city or for the balance of the metropolitan areas. A sample selection procedure was then instituted that would produce one-half of the desired sample. However, whenever a record was selected to be in sample, the housing unit record adjacent to it on the file was also selected to be in sample, thereby ensuring the necessary designated sample size.

Before the sample was selected from the group quarters and special place records, the records were stratified by census tract and census enumeration district (ED)

within the central city and within the balance of the metropolitan areas. A sample of special place records was then selected by a procedure that produced one-quarter of the desired sample size. However, at the time of the survey, the housing units at each of the special places were listed and subsampled at a rate that produced an expected four sample units, thereby ensuring the necessary designated sample size.

**Sample from the 1970-based new construction universe.** The second frame from which the metropolitan area sample was selected was a list of new construction building permits issued since 1970 (i.e., the new construction universe). The sample selection from the list of new construction building permits was an independent operation within the metropolitan area. Under clerically selected procedures, the list of permits was stratified by the date the permits were issued, and clusters of an expected four (usually adjacent) housing units were formed. These clusters were then sampled for inclusion at the overall sampling rate. In February 1984, the new construction sampling operation for the 1970-based and 1980-based areas were combined into one computerized system. The universe sampled in the computerized system will be referred to in the estimation section as the 1980-based permit universe. Under the procedures prior to sample selection, the list of permits was stratified by the date of issue, State, 1980 central city and balance, county or minor civil division, and permit office. Clusters of an expected four (usually adjacent) housing units were formed. These clusters were then sampled for inclusion at twice the overall sampling rate. The housing units within each of the clusters were then subsampled so that two of the four housing units originally selected were kept in sample.

**Sample from the 1970-based nonpermit universe.** For those metropolitan areas that were not 100-percent permit-issuing, the remainder of the AHS-MS sample was selected from a frame consisting of areas not under the jurisdiction of permit-issuing offices (i.e., the nonpermit universe). The first step in the sampling operation for the nonpermit universe was the selection of a sample of census enumeration districts. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the metropolitan area.

The probability of selection of an ED was proportionate to the following:

$$\frac{\text{Number of housing units in 1970 census ED} + \frac{\text{Group quarters population in 1970 census ED}}{3}}{4}$$

The sample ED's were then divided into segments (i.e., small land areas with well-defined boundaries having an expected size of four, or a multiple of four, housing units).

At the time of the survey, those segments that did not have an expected size of four were further subdivided to produce an expected four sample housing units. The next step was the selection of one of these segments within each sample ED. All housing units in existence at the time of interview in these selected segments were eligible for sample. Thus, housing units enumerated in the 1970 census as well as housing units built since the 1970 census were included.

### **Sample Selection for the AHS-MS Coverage Improvement Program.**

The AHS-MS Coverage Improvement Program was undertaken to correct certain deficiencies in the metropolitan area sample from the 1970-based permit-issuing universe and the 1970-based new construction universe within the 1970-based area. The coverage deficiencies included the following units:

- a. New construction from building permits issued prior to January 1970, but completed after April 1, 1970.
- b. Mobile homes placed in parks either missed during the 1970 census or established since the 1970 census.
- c. Housing units missed in the 1970 census.
- d. Housing units converted to residential use that were nonresidential at the time of the 1970 census.
- e. Houses that have been moved onto their present site since the 1970 census.
- f. Mobile homes placed outside parks since the 1970 census or vacant at the time of the 1970 census.

For a detailed description of the coverage improvement sample selection process, see earlier reports in the H170 series for the years 1976 through 1981.

### **1985 AHS-MS Sample Reduction and Sample Reinstatement.**

The 1985 AHS-MS sample reduction dropped units from sample, whereas the 1985 AHS-MS sample reinstatement added enumerated units that were previously dropped from sample. The universes involved were (a) the 1970-based permit-issuing universe, (b) the 1970-based new construction universe, and (c) the 1970-based nonpermit universe.

Sample reduction and reinstatement involved dropping or adding (a) individual housing units from the permit-issuing universe, (b) whole clusters from the new construction universe, and (c) whole segments from the nonpermit universe.

The reduction/reinstatement was implemented to achieve two criteria:

- a. A sample size of 8,500 or 4,250 in each metropolitan area.
- b. A sample having an equal number of owners and renters.

To achieve these results, each unit was classified according to the original panel number (the original sample was divided into 12 panels, with one-twelfth of the sample being in each panel) and 1985 tenure (each housing unit was given a 1985 tenure based on the previous year's tenure status). In order to simplify field procedures, panels 1 through 3 (i.e., a random one-fourth of the original sample) were dropped from sample whenever possible. More sample reductions were implemented separately for each 1985 tenure group (using different selection rates) across the remaining panels.

**AHS-MS Sample Selection for the 1980-Based Area Sample of the Metropolitan Areas.**

The sample for new areas added to the 1970-based metropolitan areas, and metropolitan areas in sample for the first time that in 1980 were 100-percent permit-issuing, was selected from two frames: (a) housing units enumerated in the 1980 Census of Population and Housing in areas under the jurisdiction of permit-issuing areas (the 1980-based permit-issuing universe), and (b) housing units constructed in permit-issuing areas since the 1980 census (1980-based new construction universe).

In addition, the sample for those metropolitan areas that were not 100-percent permit-issuing in 1980 included a sample from a third frame: housing units not under the jurisdiction of permit-issuing offices (1980-based nonpermit universe).

In 1980, the Boston, MA-NH CMSA; Ft. Worth-Arlington, TX PMSA; Minneapolis-St. Paul, MN-WI MSA; and Washington, DC-MD-VA MSA were the only metropolitan areas that added new areas that were not 100-percent permit-issuing. To satisfy confidentiality requirements in the Boston, MA-NH CMSA; Ft. Worth-Arlington, TX PMSA; and Washington, DC-MD-VA MSA, it was necessary to supplement the existing sample within the 1970-based area. The additional housing units were selected separately for each metropolitan area from the 1980-based permit-issuing universe.

**Sample from the 1980-based permit-issuing universe.**

The major portion of the sample in each metropolitan area was selected from a file that represented all the housing units enumerated in permit-issuing areas during the 1980 Census of Population and Housing. This file contained records for occupied housing units, vacant housing units, and housing units in group quarters. Sampling operations were done separately for noninstitutionalized group quarters and for all other housing units in permit-issuing areas. In addition, in order that an equal number of owner and

renter housing units were selected in each metropolitan area, a selection rate that differed by tenure group was used. Before the sample was selected, the housing units that were not classified as group quarters were stratified into 60 categories by tenure, contract rent, value, and number of rooms as illustrated by the following table:

Contract rent and value	Number of rooms		
	1-3	4-5	6+
<b>RENTER</b>			
Contract rent:			
Less than \$100 .....			
\$100 to \$149 .....			
\$150 to \$199 .....			
\$200 to \$249 .....			
\$250 to \$299 .....			
\$300 to \$349 .....			
\$350 to \$399 .....			
\$400 or more .....			
Not available .....			
<b>OWNER</b>			
Value:			
Less than \$20,000 .....			
\$20,000 to \$29,999 .....			
\$30,000 to \$34,999 .....			
\$35,000 to \$39,999 .....			
\$40,000 to \$49,999 .....			
\$50,000 to \$64,999 .....			
\$65,000 to \$79,999 .....			
\$80,000 to \$99,999 .....			
\$100,000 to \$149,999 .....			
\$150,000 or more .....			
Not available .....			

The group quarters housing units were grouped into two strata: institutionalized group quarters and noninstitutionalized group quarters.

The following sample selection procedures were then implemented separately within the central city and balance of the metropolitan area. For the Boston, MA-NH CMSA; Ft. Worth-Arlington, TX PMSA; and Washington, DC-MD-VA MSA, the sample selections were implemented separately by the 1970-based and 1980-based areas. All units were sorted by the 1980 central city and balance, stratum, State, district office, ED, and census serial number. The sample selection procedure was then implemented separately for (a) institutionalized group quarters and nongroup quarters housing units and (b) noninstitutionalized group quarters

Individual housing units were selected for the nongroup quarters, but each institutionalized group quarters had one chance of selection. Before the sample selection for the noninstitutionalized group quarters was implemented, the following measure of size was calculated for each record:

$$(1/4) \times \frac{\text{Total group quarters population}}{2.75}$$

The noninstitutionalized group quarters were then selected proportionate to the measure of size.

**Sample selection from the 1980-based new construction universe.** The second frame from which the metropolitan area sample was selected was a list of new construction building permits issued since 1980 (i.e., the new construction universe). The sample selection from the list of new construction building permits was an independent operation within each metropolitan area. This operation was described in the discussion of the 1970-based new construction universe.

**Sample from the 1980-based nonpermit universe.** For those metropolitan areas that were not 100-percent permit-issuing, the remainder of the AHS-MS sample was selected from a frame consisting of areas not under the jurisdiction of permit-issuing offices (i.e., the 1980-based nonpermit universe). The first step in the sampling operation for the nonpermit universe was the selection of a sample of census ED's within these areas (using the overall sampling rate). Prior to this sample selection, the ED's were sorted by State, district office, and enumeration district number. The probability of selection of an ED was proportionate to the following:

$$\frac{\text{Number of housing units in 1980 census ED} + \frac{\text{Noninstitutionalized group quarters population in 1980 census ED}}{2.75}}{4}$$

The sample ED's were then divided into segments (i.e., small land areas with well-defined boundaries having an expected size of four, or a multiple of four, housing units). At the time of the survey, those segments that did not have an expected size of four housing units were further subdivided to produce an expected four sample housing units. Following the division, a segment from each sample ED was selected. All housing units in existence at the time of interview in these selected segments were eligible for sample. Thus, housing units enumerated in the 1980 census as well as housing units built since the 1980 census are included.

### 1989 AHS-MS Sample Reduction and Sample Reinstatement

When these metropolitan areas were interviewed in 1985, 5 of the 11 had an expected sample size of 8,500 distributed throughout panels 4 through 12; in addition, panels 11 and 12 were dropped before interviewing was completed in these large metropolitan areas, further reducing the sample size. The remaining six metropolitan areas had an expected sample size of 4,250 in 1985; in these areas, one or both of panels 11 and 12 were also dropped. In addition, for the large metropolitan areas, the sample size was reduced from 8,500 to 4,250 by randomly selecting half of the original panels 4 through 12 to be dropped. Furthermore, there was some reassignment of units between

panels 9 and 10 and panels 11 and 12 so that all the units interviewed in 1989 also had a prior interview. In addition, panels 11 and 12 were later dropped because of budgetary concerns.

### AHS-National Sample Selection

This sample was set up as a multistage design in which the United States was divided into areas made up of counties and independent cities called primary sampling units (PSU's). These PSU's were grouped into strata consisting of one or more PSU's, and then one PSU was selected from each stratum to represent all PSU's in that stratum.

**Selection from the 1980 census.** Sample units were selected from 1980 census units in these PSU's at an overall sampling rate of 1 in 2,148. The procedure for sampling housing units in a given area depended on (a) the completeness of addresses and (b) the degree of monitoring of new construction by permits.

In areas where addresses were mostly complete and where new construction is monitored by permits, a sample was selected from a list of housing units that received the long-form questionnaire in the 1980 census. This list was based on housing and geographic information on the housing unit.

In areas where at least 4 percent of the addresses were incomplete or inadequate, or where new construction was not monitored by building permits (most rural areas), a sample of 1980 "long-form questionnaire" census units was selected in several steps:

- The areas were grouped, and a sample of areas was chosen.
- A segment was selected within each sample area.
- A sample of housing units that received 1980 census long forms was selected within the segment.

**Selection of new construction housing units in permit-issuing areas.** The sample of new construction was selected from issued building permits so that the units were expected to be completed after April 1, 1980. The sampling procedure was similar to that of AHS-MS; however, the subsampling rate used was 1 in 4.

**Selection of units from the nonpermit universe.** Housing units added to the inventory since the 1980 census were represented using two methods:

- Within-structure additions, which are units in structures that contained at least one unit enumerated in the 1980 census
- Whole-structure additions, which include units in structures that contained no units enumerated in the 1980 census.

Additional information concerning the 1989 AHS-National survey is available in the current housing report series H150/89.

## ESTIMATION

The 1989 AHS-Metropolitan Area sample produced estimates pertaining to characteristics of the housing inventory at the time of the interview (i.e., the 1989 housing inventory). The combined estimates used information from both the AHS-MS and AHS-National samples (i.e., the combined sample estimates).

### AHS-MS

Prior to performing estimation procedures using the combined sample, the AHS-MS sample housing units were weighted according to a one-stage ratio estimation procedure. Before the implementation of the ratio estimation procedure, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for Type M and Type A noninterviews.

**Type M noninterview adjustment.** The Type M noninterviews are sample units that were dropped because of selection by another survey. These noninterviews occur in: (a) the 1980-based permit-issuing area universe; (b) the 1980-based nonpermit-issuing area universe; and (c) the 1980-based new construction universe.

The adjustment was done separately for the above universes for the central city and balance for each metropolitan area. The adjustment was equal to the following:

$$\frac{\text{AHS-MS sample estimate of 1980 housing units in the cell} + \text{Weighted count of Type M noninterviewed housing units}}{\text{AHS-MS sample estimate of 1980 housing units in the cell}}$$

**Type A noninterview adjustment.** Type A noninterviews are sample units for which occupants were not home; occupants refused to be interviewed; or occupants were unavailable for some other reason.

The adjustment was done on occupied units and was computed separately for units in the 1980-based permit-issuing area universe; new construction; and all other housing units (this includes the 1970-based permit-issuing universe, the 1970-based and 1980-based nonpermit-issuing universes, and the 1970-based new construction housing units built prior to the last survey).

For units in the 1980-based permit-issuing universe, a Type A noninterview adjustment factor was computed separately for each of the 62 strata used in the sample selection process, by central city and balance. For new construction units, a Type A noninterview adjustment factor was computed separately for each central city and

balance. For all other units, a Type A noninterview adjustment factor was calculated separately by tenure and 1970 central city and balance for each of the following:

- a. Twenty-four noninterview cells for sample housing units from the permit-issuing universe (each cell was derived from one or more of the 50 different strata used in the 1970-based permit-issuing universe, illustrated earlier).
- b. One noninterview cell for new construction housing units.
- c. One noninterview cell for mobile homes or trailers from the nonpermit-issuing universe.
- d. One noninterview cell for units that were not mobile homes or trailers from the nonpermit-issuing universe.
- e. Three noninterview cells for units from the coverage improvement universe.
- f. One noninterview cell for units classified as vacants at the time of the 1970 census.
- g. One noninterview cell for units classified as group quarters at the time of the 1970 census.

Within a given cell, the Type A noninterview adjustment factor was equal to the following ratio using the basic weight times the Type M noninterview adjustment factor for the sample weight:

$$\frac{\text{Weighted count of interviewed housing units} + \text{Weighted count of Type A noninterviewed housing units}}{\text{Weighted count of interviewed housing units}}$$

**Ratio estimation procedure for the 1970-based permit-issuing universe.** The following ratio estimation procedure was employed for all sample housing units from the permit-issuing universe. This factor was computed separately for all sample housing units within each 1970-based permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

$$\frac{\text{1970 census count of housing units from the 1970-based permit-issuing universe in the corresponding cell}}{\text{AHS-MS sample estimate of 1970-based housing units from the permit-issuing universe in the corresponding cell}}$$

For each metropolitan area, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file (long forms) of housing units enumerated in areas under the jurisdiction of permit-issuing offices.

The denominators of the ratio estimation factors were then obtained from weighted estimates of all the AHS-MS sample housing units from the 1970-based permit-issuing universe using the existing weight (i.e., the basic weight times the Type A noninterview adjustment). The computed ratio estimation factor was then applied to the existing weight for each sample housing unit within the corresponding ratio estimation cells. This ratio estimation procedure was introduced to correct the probabilities of selection for samples in each of the strata used in the sample selection of the 1970-based permit-issuing universe. Prior to the AHS-MS sample selection within each metropolitan area, housing units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. The same probability of selection was then applied to the remaining units to select the AHS-MS sample. Since the number of housing units deleted from the AHS-MS universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata was introduced during the sample selection process.

**Ratio estimation procedure for the 1980-based permit-issuing universe.** The following ratio estimation procedure was employed for all sample units from the 1980-based permit-issuing universe. This factor was computed separately for all metropolitan areas within each 1980-based permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor was equal to the following:

$$\frac{\text{1980 census count of housing units from the 1980-based permit-issuing universe in the corresponding cell}}{\text{AHS-MS sample estimate of 1980-based housing units from the permit-issuing universe in the corresponding cell}}$$

For each metropolitan area, the numerator of the ratio was obtained from the 1980 Census of Population and Housing 100-percent file of housing units enumerated in areas under the jurisdiction of permit-issuing offices. The denominator of the ratio was obtained from weighted estimates of all the AHS-MS sample housing units within the corresponding ratio estimation categories using the existing weight (i.e., the basic weight times the Type M noninterview adjustment factor times the Type A noninterview adjustment factor).

The computed ratio estimation factor was then applied to the existing weight for each sample housing unit within the corresponding ratio estimation categories.

This ratio estimation procedure was introduced to adjust the sample estimate in each of the strata used in the sample selection of the 1980-based permit issuing universe to an independent estimate (1980 census count) for the strata. This adjustment was necessary since some sample units were dropped during processing.

## AHS-National

Before implementing estimation procedures using the AHS-National units for the combined sample, the AHS-National sample units were assigned a weight that reflected the probability of selection for the unit. The AHS-National weighting procedure then made adjustments for units that could not be interviewed for a variety of reasons. For each of these adjustments, a factor was computed and applied to the appropriate units.

The first of these adjustments was done for permit segments only to account for permits that could not be sampled and units that could not be located. These were represented by all other units in permit segments including both interviews and noninterviews (excluding "unable to locate" noninterviews).

The second of the adjustments was done for units in structures built before April 1, 1980. It was done to account for units that could not be located. The unlocatable units were represented by both interviews and noninterviews (excluding "unable to locate" noninterviews).

The last of these adjustments was done to account for units that could not be interviewed because either no one was home after repeated visits or the respondent refused to be interviewed. When prior year AHS or 1980 census data were available, this information was used to determine the noninterview adjustment cell. The cells included characteristics such as tenure, geography, units in structure, and number of rooms. When these data were not available, adjustment factors were computed separately using more general characteristics such as type of area and type of housing unit (i.e., mobile home, nonmobile home). Additional information on the AHS-National weighting procedure can be found in the current housing reports H150/89 series.

## COMBINED SAMPLE WEIGHTING

### Introduction

The estimates for the combined sample were obtained by summing the sample weights of interviewed AHS-MS and AHS-National units. For AHS-MS sample units, the starting weight was obtained after the AHS-MS ratio estimation procedure. For AHS-National units, the starting weight was obtained after the Type A noninterview adjustment. In order to account for the use of two different samples representing one metropolitan area, weighting factors were assigned to each unit prior to the combined sample ratio estimation procedures.

### Weighting Factor Adjustment

The weighting factor adjustment was computed separately for each metropolitan area by sample design (AHS-MS or AHS-National) according to "new construction" or "old

construction" classification. New construction was defined as units built in permit-issuing areas since the 1980 census; old construction units were then categorized by tenure classification (renter/owner).

For a given characteristic, the AHS-MS weighting factor adjustment was a function of the sample size in each survey and the variance associated with each survey's estimates.

The corresponding weighting factor was then applied to the existing weight of each AHS-MS and AHS-National sample unit, and the weights were then combined according to characteristic (i.e., AHS-MS new construction + AHS-National new construction, etc.).

### Combined Sample Ratio Estimation Procedures

For the three ratio estimate procedures described below, each metropolitan area was subdivided into geographic areas consisting of a combination of counties.

**Mobile home ratio estimation.** The following ratio estimation procedure was applied in the Dallas, TX PMSA; Fort Worth-Arlington, TX PMSA; Los Angeles-Long Beach, CA PMSA; Phoenix, AZ MSA; and Tampa-St. Petersburg, FL MSA:

Independent estimate of mobile homes  
for the corresponding geographic subdivision  
of the metropolitan area

---

Sample estimate of mobile homes  
for the corresponding geographic subdivision  
of the metropolitan area

The numerator of this ratio was determined using data from the 1980 census and the 1990 census. The denominator was obtained using the existing weight of AHS sample mobile home units (i.e., the starting weight times the combined sample weighting factor).

**Independent total housing unit ratio estimation.** The following ratio estimation procedure was applied in the Dallas, TX PMSA; Fort Worth-Arlington, TX PMSA; Los Angeles-Long Beach, CA PMSA; Phoenix, AZ MSA; and Tampa-St. Petersburg, FL MSA.

Independent estimate of the occupied housing inventory  
(excluding mobile homes) for the corresponding  
geographic subdivision of the metropolitan area

---

Sample estimate of the occupied housing inventory  
(excluding mobile homes) for the corresponding  
geographic subdivision of the metropolitan area

The independent estimates of occupied housing units that were used as the numerator of this ratio are described below. The denominator was obtained by using the existing weight of AHS sample units (excluding mobile homes).

The methodology used to derive the independent estimates of occupied housing units used a three-step procedure. In step one, the Census Bureau's State household estimates for July 1, 1988, and July 1, 1989, were used to extrapolate State household estimates for July 15, 1989 (the midpoint of the survey interview period—June 23, 1989, for Detroit) and April 1, 1990 (the date of the 1990 Decennial Census of Population and Housing).

In step two, the proportion of the July 1, 1985, to April 1, 1990, State household growth that occurred during the July 1, 1985, to July 15, 1989, time period was estimated using these estimates.

In step three, the July 15, 1989, independent county estimates of households (E) for each county in a metropolitan area were produced using the following formula. The county totals were then summed to the geographic subdivision level.

$$E = G + P(F - G)$$

where P = the proportion derived in step two above.

F = County counts of households for April 1, 1990, from the 1990 Decennial Census of Population and Housing.

G = County estimates of households as of July 1, 1985.

The above three-step procedure was used because after the 1980 census the only available independent estimates of households on a smaller than State level (county level) were the estimates as of July 1, 1985, and the 1990 census counts as of April 1, 1990. The weighting was designed to be consistent with the 1990 census without relying on a linear interpolation between July 1, 1985, and April 1, 1990, since it's probably not reasonable to assume linear growth for such a long time period in many metropolitan areas.

The survey estimate of occupied mobile homes after application of the mobile home ratio estimation factor described above was then subtracted from this independent estimate of occupied housing units. The resulting estimate of occupied housing units, excluding mobile homes, was used as the numerator for this ratio estimation.

**Independent total housing unit ratio estimation.** The following ratio estimation procedure was applied for all other areas except those listed above.

Independent estimate of occupied housing inventory  
for the corresponding geographic subdivision  
of the metropolitan area

---

Sample estimate of occupied housing inventory  
for the corresponding geographic subdivision  
of the metropolitan area

The independent estimates of occupied housing units that were used as the numerator of this ratio were derived using the three-step procedure described above. The denominator was obtained by using the existing weight of AHS sample units (i.e., the starting weight times the combined sample weighting factor).

The computed ratio estimation factors were then applied to all appropriate housing units (including vacant units) in the corresponding geographic area of each metropolitan area, and the resulting product was used as the final weight for tabulation purposes.

The effect of these ratio estimation procedures was to reduce the sampling error for most statistics below what would have been obtained by simply weighting the results of the sample by the inverse of the probability of selection. Since the housing population of the sample differed somewhat, by chance, from the metropolitan area as a whole, it can be expected that the sample estimates will be improved when the sample housing population, or different portions of it, is brought into agreement with known good estimates of the metropolitan area housing population.

## ACCURACY OF THE ESTIMATES

There are two types of possible errors associated with estimates based on data from sample surveys—nonsampling and sampling errors. A description of the nonsampling and sampling errors associated with the AHS sample estimates follows.

### Nonsampling Errors

In general, nonsampling errors can be attributed to many sources: inability to obtain information about all cases; definitional difficulties; differences in the interpretation of questions; inability or unwillingness of respondents to provide correct information; mistakes in recording or coding the data; other errors of collection, response, processing, and coverage; and estimation for missing data. Nonsampling errors are not unique to sample surveys since they can, and do, occur in complete censuses as well.

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1989 AHS-Metropolitan Area sample. In the following sections, the major sources of nonsampling errors will be discussed.

**AHS-MS content errors.** A Content Reinterview Program was done for the 1989 AHS-Metropolitan Area sample units. A sample of these units was revisited, and answers to some of the questions on the questionnaire were obtained again. The original interview and reinterview were

assumed to be two independent readings and, thus, were the basis for the measurement of the accuracy of the data collected from interviewed units.

The 1989 Content Reinterview Program served as an interviewer quality check and a quality analysis of particular survey questions. Some of the interviewers were selected for the quality check, which reviewed the interviewers' proficiency in properly evaluating the items listed below. The other portion of the reinterview program was performed to ensure that certain survey questions elicited consistent responses from the interviewed households. These reinterview items and their response variability are discussed below.

The six interviewer items reviewed were (1) correct unit visited; (2) area segment coverage; (3) living quarters classification; (4) tenure; (5) interview status; and (6) household composition.

The AHS-MS survey items reviewed generally fell into three categories: (1) major repairs; (2) mortgage; and (3) mobility. The results of this reinterview program, however, are not available at this time.

Although the results of the 1989 Content Reinterview Program are not available, past reinterview programs have shown that certain items are likely to produce moderate or high response variability. Response variability is defined as a measure of consistency between the original survey response to an item and the reinterview response to that item. Moderate levels of variability indicate that the response error is not insignificant in comparison to the sampling error. High variability indicates that the response errors are very significant in relation to the sampling errors with which they are associated; therefore, caution should be used when considering estimates of these characteristics. The 1985 Content Reinterview Program had five items that exhibited high variability: (1) major repairs over \$500 each; (2) payments the same throughout mortgage; (3) area lived at age 16; (4) preferred place to live in 5 years; and (5) size of lot. Prior-year results for the 1989 metropolitan areas can be found in the Census Bureau publication series H170 for the year 1985.

**AHS-National content errors.** A Content Reinterview Program was conducted for the AHS-National households as well. A subsample of the original households was revisited, and certain questions from the original questionnaire were asked again. The original and reinterview were assumed to be two independent readings and, thus, were the basis for the measurement of the response error of the AHS estimates of mortgage items. The reinterview also served as a check for interviewer evaluation and quality control. The AHS-National reinterview program performed an interviewer quality check using questions similar to those described above.

Reinterview studies were also conducted in conjunction with previous AHS-National and AHS-MS enumerations. These studies included items dealing primarily with poor housing quality, attitudes about the neighborhood, and

certain housing costs. The following table shows the items that had higher levels of inconsistency. While not all of these questions were included in the 1989 reinterview studies, questions from previous enumerations were not altered enough to lead one to believe that the level of inconsistent responses would change.

A possible explanation for the results of the reinterview studies, as well as the surveys themselves, is that respondents may lack precise information. Also, since the results

Survey items	Level of inconsistency
Mice and rats .....	Moderate
Real estate taxes .....	Moderate
Cost of real estate taxes .....	Moderate
Prefer to live in same area or somewhere else .....	Moderate
Open cracks or holes on inside of building .....	Moderate to High
Holes in floors .....	Moderate to High
Blown fuses/tripped circuit breakers .....	Moderate to High
Neighborhood conditions: street noise; roads in need of repair; crime; trash, litter, junk in streets or on properties; boarded up/abandoned structures; nonresidential activities; odors, smoke, gas .....	Moderate to High
Satisfactory neighborhood services: police protection; hospitals/health clinics; public transportation; shopping; elementary schools .....	Moderate to High
Electricity cost .....	Moderate to High
Oil, coal, kerosene, wood or other fuel cost .....	Moderate to High
Fire/hazard insurance .....	Moderate to High
Cost of garbage collection .....	Moderate to High
Broken plaster or peeling paint on ceiling and walls .....	High
Working electric outlet in all rooms .....	High
Concealed wiring .....	High
Gas cost .....	High
Cost of water supply and sewage disposal .....	High
Gross income .....	High

of the reinterview studies are derived from sample surveys, there is sampling error associated with these estimates of nonsampling error. The possibility of such errors should be taken into account when considering the results of these studies.

**Coverage errors.** In errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional (nonmobile home or trailer) new construction. Because of time constraints, only those building permits issued more than 7 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for each metropolitan area. However, those permits issued during the last 7 months of the survey do not necessarily represent missed housing units. Because of the relatively short time span involved, it is possible that construction of these housing units was not completed at the time the survey was conducted; consequently, they would not have been eligible for interview. In addition to

these deficiencies, new construction in special places that do not require building permits, such as military bases, is not adequately represented.

AHS misses a significant portion of new mobile homes. It is believed that most of the difference is because of poor coverage of mobile home parks in address ED's. Undercoverage exists for those mobile homes built between the time of the last coverage improvement procedure and the 1980 census. It has been estimated that on a national level as much as 25 percent of those mobile homes built after January 1, 1980, may be missed. Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all housing units located inside these ED's would be represented in the sample. However, it has been established that the AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. It should be noted that since these ED's were recanvassed each time this metropolitan area was surveyed, the number of missed housing units may be considerably less for the 1989 survey.

The mobile home and total housing unit ratio estimation procedures correct for these deficiencies as far as the count of mobile homes and total housing units is concerned (i.e., it adjusts to the best available estimate). However, biases of subtotals would still remain.

**Rounding errors.** For errors associated with processing, rounding of estimates introduces another source of error in the data, the severity of which depends upon the statistics being measured. The effect of rounding is significant relative to the sampling error only for small percentages or small medians when these figures are derived from relatively large bases (e.g., median number of persons per household). This means that confidence intervals formed from the standard errors given may be distorted, and this should be taken into account when considering the results of the survey.

**Sampling Errors for the AHS Combined Sample Estimates**

The particular sample used for this survey is one of a large number of possible samples of the same size that could have been selected using the same sample design. Even if the same questionnaires, instructions, and interviewers were used, estimates from each of the different samples would differ from one another. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and, thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates because of sampling and nonsampling errors, but it does not measure

as such any systematic biases in the data. Therefore, the accuracy of the estimates depends upon the standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the average result of all possible samples.

The average result of all possible samples either is or is not contained in any particular computed interval. However, for a particular sample, one can say with specified confidence that the average result of all possible samples is included in the constructed interval.

The figures presented in the tables that follow (see page App-67) are approximations to the standard errors of various estimates shown in this report for this metropolitan area. In order to derive standard errors that would be applicable to a wide variety of items and also could be prepared at a moderate cost, a number of approximations were required. As a result, the tables of standard errors provide an indication of the order of magnitude of the standard errors rather than precise standard errors for any specific item.

Table I presents the standard errors applicable to estimates of characteristics of the 1989 housing inventory. Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table.

The reliability of an estimated percentage, computed by using the sample data for both numerator and denominator, depends upon both the size of the percentage and the size of the total upon which the percentage is based. Estimated percentages are relatively more reliable than the corresponding estimates of the numerators of the percentages, particularly if the percentages are 50 percent or more.

Table II presents the standard errors of estimated percentages for the 1989 housing inventory. Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in these tables.

Included in tables I and II are estimates of standard errors for estimates of zero and zero percent. These estimates of standard errors are considered overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when estimates of zero are obtained.

For ratios, 100 (x/y), where x is not a subclass of y, table II underestimates the standard error of the ratio when there is little or no correlation between x and y. For this type of ratio, a better approximation of the standard error may be

obtained by letting the standard error of the ratio be approximately equal to the following:

$$(100) \frac{x}{y} \sqrt{\frac{S_x^2}{x^2} + \frac{S_y^2}{y^2}}$$

- where: x = the numerator of the ratio
- y = the denominator of the ratio
- s<sub>x</sub> = the standard error of the numerator
- s<sub>y</sub> = the standard error of the denominator

**Illustration of the Use of the Standard Error Tables**

Table 2-1 of this report shows that in the Boston, MA-NH, metropolitan area, there were 914,500 owner-occupied housing units. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 14,090. The following interpolation procedure was used.

The information presented in the following table was extracted from table I. The entry for "x" is the one sought.

Size of estimate	Standard error
900,000 .....	14,120
914,500 .....	x
1,000,000 .....	13,890

The entry of "x" is determined as follows by vertically interpolating between 14,120 and 13,890.

$$\begin{aligned} 914,500 - 900,000 &= 14,500 \\ 1,000,000 - 900,000 &= 100,000 \\ 14,120 + \frac{14,500}{100,000} (13,890 - 14,120) &= 14,090 \end{aligned}$$

Consequently, the 90-percent confidence interval, as shown by these data, is from 891,960 to 937,040 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1989 owner-occupied housing units lies within a range computed in this way would be correct for roughly 90 percent of all possible samples.

Table 2-3 also shows that of 914,500 owner-occupied housing units, 217,700 or 23.8 percent had two bedrooms. Interpolation using table II of this appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 23.8 percent is approximately 1.0 percentage points. The following interpolation procedure was used.

The information presented in the following table was extracted from table II.

Base of percentage	Estimated percentage		
	10 or 90	23.8	25 or 75
900,000 .....	0.7	a	1.0
914,500 .....		p	
1,000,000 .....	0.7	b	0.9

1. The entry for cell "a" is determined by horizontal interpolation between 0.7 and 1.0.

$$\begin{aligned}
 23.8 - 10.0 &= 13.8 \\
 25.0 - 10.0 &= 15.0 \\
 0.7 + \frac{13.8}{15.0} (1.0 - 0.7) &= 1.0
 \end{aligned}$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.7 and 0.9.

$$\begin{aligned}
 23.8 - 10.0 &= 13.8 \\
 25.0 - 10.0 &= 15.0 \\
 0.7 + \frac{13.8}{15.0} (0.9 - 0.7) &= 0.9
 \end{aligned}$$

3. The entry for "p" is then determined by vertical interpolation between 1.0 and 0.9.

$$\begin{aligned}
 914,500 - 900,000 &= 14,500 \\
 1,000,000 - 900,000 &= 100,000 \\
 1.0 + \frac{14,500}{100,000} (0.9 - 1.0) &= 1.0
 \end{aligned}$$

The result is a standard error of 1.0 percentage points. Consequently, the 90-percent confidence interval, as shown by these data, is from 22.2 to 25.4 percent.

**Differences**

The standard errors shown are not directly applicable to differences between two sample estimates. They are quite accurate for the difference between estimates of the same characteristics in two different metropolitan areas or the difference between separate and uncorrelated characteristics in the same metropolitan area. If there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error, but if there is a high negative correlation, the formula will underestimate the true standard error.

**Illustration of the computation of the standard error of a difference.** Table 2-3 of this report shows that in the Boston, MA-NH, metropolitan area, there were 420,700 owner-occupied housing units with three bedrooms. Thus, the apparent difference, as shown by these data, between owner-occupied units with two bedrooms and owner-occupied units with three bedrooms is 203,000. Table I,

with interpolation, shows that the standard error of 420,700 is approximately 12,300, and the standard error of 217,700 is approximately 9,520. Therefore, the standard error of the estimated difference of 203,000 is about 15,550.

$$15,550 = \sqrt{(12,300)^2 + (9,520)^2}$$

Consequently, the 90-percent confidence interval for the 203,000 difference is from 178,120 to 227,880 housing units. Therefore, a conclusion that the average estimate derived from all possible samples, of this difference, lies within a range computed in this way would be correct for roughly 90 percent of all possible samples. Thus, we can conclude with 90-percent confidence that the number of 1989 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied housing units with two bedrooms since the 90-percent confidence interval does not include zero or negative values.

**Medians**

For medians presented in certain tables, the sampling error depends on the size of the base and on the distribution upon which the median is based. An approximate method for measuring the reliability of the estimated median is to determine an interval about the estimated median so that there is a stated degree of confidence, such that the average median from all possible samples lies within the interval. The following procedure may be used to estimate confidence limits of a median based on sample data:

1. From table II, determine the standard error of a 50-percent characteristic on the base of the median.
2. Add to and subtract from 50 percent the standard error determined in step 1.
3. Using the distribution of the characteristics, determine the confidence interval corresponding to the two points established in step 2. To find the lower endpoint of the confidence interval, it is necessary to know into which interval of the distribution the lower percentage limit falls. Similarly, to find the upper endpoint of the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these distribution intervals could be different, although this will not happen very often.

A 1.6 standard-error confidence interval may be determined by finding the values corresponding to 50 percent plus and minus 1.6 times the standard error determined in step 1. For about 90 out of 100 possible samples, the average median from all possible samples would lie between these two values.

**Illustration of the computation of the 90-percent confidence interval of a median.** Table 3-13 of this report shows the median monthly housing cost for owner-occupied housing units is \$663. The base of this distribution is

obtained by subtracting "mortgage payment not reported" from the number of occupied units. Using this definition, the base is equal to 804,600.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 804,600 is approximately 1.2 percentage points.
2. To obtain a 90-percent confidence interval on the estimated median, initially add to and subtract from 50 percent 1.6 times the standard error determined in step 1. This yields percentage limits of 48.1 and 51.9.
3. From the distribution for monthly housing cost in table 3-13, the \$600 to \$700 interval for owner-occupied housing units contains the 48.1 percent derived in step 2. About 373,400 housing units, or 46.4 percent, fall below this interval, and 45,900 housing units, or 5.7 percent, fall within this interval. By linear interpolation,

the lower limit of the 90-percent confidence interval is found to be about \$630.

$$600 + (700 - 600) \frac{48.1 - 46.4}{5.7} = 630$$

4. Similarly, the \$600 to \$700 interval for owner-occupied housing units contains the 51.9 percent derived in step 2. As stated before 373,400 housing units (46.4 percent) fall below this interval, and 45,900 housing units or 5.7 percent, fall within this interval. The upper limit of the 90-percent confidence interval is found to be about \$696.

$$600 + (700 - 600) \frac{51.9 - 46.4}{5.7} = 696$$

Thus, the 90-percent confidence interval ranges from \$630 to \$696.

5. Finally, note that the medians shown in this report are calculated from unrounded data and then rounded. Thus, they may differ from the medians calculated from the grouped data in the tables of this report.

**Table I. Standard Errors for Estimated Number of Housing Units in the 1989 Boston, MA CMSA**

Size of estimate	Standard error <sup>1</sup>		
	Combined owner and renter housing units <sup>2</sup>	Owner housing units <sup>3</sup>	Renter housing units <sup>4</sup>
0.....	480	480	470
400.....	480	480	470
700.....	580	580	580
1,000.....	690	690	690
2,500.....	1,100	1,100	1,090
5,000.....	1,550	1,550	1,530
10,000.....	2,190	2,190	2,170
25,000.....	3,440	3,440	3,410
50,000.....	4,830	4,830	4,790
75,000.....	5,870	5,870	5,820
100,000.....	6,720	6,730	6,670
150,000.....	8,100	8,110	8,030
200,000.....	9,200	9,200	9,120
250,000.....	10,110	10,110	10,020
300,000.....	10,880	10,880	10,780
350,000.....	11,530	11,540	11,430
400,000.....	12,090	12,100	11,990
450,000.....	12,570	12,580	12,460
500,000.....	12,980	12,980	12,860
600,000.....	13,590	13,600	13,480
700,000.....	13,980	13,980	13,860
800,000.....	14,150	14,160	14,030
900,000.....	14,120	14,120	-
1,000,000.....	13,880	13,890	-
1,100,000.....	13,420	-	-
1,200,000.....	12,730	-	-
1,300,000.....	11,750	-	-
1,400,000.....	10,410	-	-
1,500,000.....	8,530	-	-
1,600,000.....	5,620	-	-
1,668,500.....	-	-	-

<sup>1</sup>To compute standard errors for new construction estimates, the standard errors in the table should be multiplied by a factor of 1.1 for owner housing units, 1.1 for renter housing units, and 1.1 for the combined owner and renter housing units.

<sup>2</sup>Some examples that pertain to both owner and renter housing units are: total housing units; all occupied housing units; all year-round housing units; mobile homes or trailers; and total vacant housing units.

<sup>3</sup>The owner housing units pertain to owner-occupied housing units and vacant housing units excluding vacant-for-rent housing units.

<sup>4</sup>The renter housing units pertain to renter-occupied housing units and vacant-for-rent housing units.

**Table II. Standard Errors for Estimated Percentages of Housing Units in the 1989 Housing Inventory of the Boston, MA CMSA**

Base of percentage	Estimated percentage <sup>1</sup>					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
400	54.2	54.2	54.2	54.2	54.2	54.4
700	40.3	40.3	40.3	40.3	40.3	41.1
1,000	32.1	32.1	32.1	32.1	32.1	34.4
2,500	15.9	15.9	15.9	15.9	18.8	21.7
5,000	8.6	8.6	8.6	9.2	13.3	15.4
10,000	4.5	4.5	4.7	6.5	9.4	10.9
25,000	1.9	1.9	3.0	4.1	6.0	6.9
50,000	0.9	1.0	2.1	2.9	4.2	4.9
75,000	0.6	0.8	1.7	2.4	3.4	4.0
100,000	0.5	0.7	1.5	2.1	3.0	3.4
150,000	0.3	0.6	1.2	1.7	2.4	2.8
200,000	0.2	0.5	1.1	1.5	2.1	2.4
250,000	0.2	0.4	0.9	1.3	1.9	2.2
300,000	0.2	0.4	0.9	1.2	1.7	2.0
350,000	0.13	0.4	0.8	1.1	1.6	1.8
400,000	0.12	0.3	0.7	1.0	1.5	1.7
450,000	0.10	0.3	0.7	1.0	1.4	1.6
500,000	0.09	0.3	0.7	0.9	1.3	1.5
600,000	0.08	0.3	0.6	0.8	1.2	1.4
700,000	0.07	0.3	0.6	0.8	1.1	1.3
800,000	0.06	0.2	0.5	0.7	1.1	1.2
900,000	0.05	0.2	0.5	0.7	1.0	1.1
1,000,000	0.05	0.2	0.5	0.7	0.9	1.1
1,100,000	0.04	0.2	0.5	0.6	0.9	1.0
1,200,000	0.04	0.2	0.4	0.6	0.9	1.0
1,300,000	0.04	0.2	0.4	0.6	0.8	1.0
1,400,000	0.03	0.2	0.4	0.6	0.8	0.9
1,500,000	0.03	0.2	0.4	0.5	0.8	0.9
1,600,000	0.03	0.2	0.4	0.5	0.7	0.9
1,668,500	0.03	0.2	0.4	0.5	0.7	0.8

<sup>1</sup>Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than or equal to fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point.

For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.1.