

## Appendix B.

# Sample Design and Estimation

### SAMPLE DESIGN

#### Introduction

The estimates for each of the eight metropolitan areas in this report series (H171/94) are based on data collected from the 1994 American Housing Survey Metropolitan Sample (AHS-MS), which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development.

In most cases, these eight metropolitan areas are consistent with the 1993 Office of Management and Budget (OMB) definitions of the metropolitan statistical area (MA), consolidated metropolitan statistical area (CMSA), or primary metropolitan statistical area (PMSA). The exceptions to this are: Fort Worth, TX, which does not include Hood and Parker Counties from the 1993 OMB definition; Dallas, TX, which does not include Henderson and Hunt Counties from the 1993 OMB definition, and Phoenix, AZ, which does not include Pinal County from the 1993 OMB definition.

The metropolitan areas selected for the 1994 AHS-MS are usually 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, the expected sample sizes were reduced to 4,250 in the metropolitan areas with sample sizes of 8,500. For all of the 1994 MA's, interviewing took place from April 1994 through December 1994.

Table A summarizes the interview activity for the 1994 AHS in each of the metropolitan areas. The table provides the number of eligible units (comprised of completed interviews and noninterviews), and the number of units visited but ineligible for interview.

#### Designation of AHS-MS Sample Housing Units for the 1994 Survey

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

- All sample housing units that were interviewed in the previous survey. 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 which, until these procedures were implemented, did not have a chance of selection.
- All sample housing units that were Type A noninterviews (that is, units eligible to be interviewed) or Type B noninterviews (that is, units not eligible for interview at the time of the survey but which could become eligible in the future) in the previous survey. (For a list of reasons for Type A noninterviews, see the facsimile of the 1994 AHS questionnaire, page A-27.)
- All sample housing units selected from a listing of new residential construction building permits issued since the previous survey. This sample represented the housing units built in permit-issuing areas since the previous survey.

Table A. Description of the American Housing Survey—1994 Metropolitan Sample

Metropolitan statistical area	Units eligible			Units visited, not interviewed <sup>2</sup>
	Total	Interviewed	Not interviewed <sup>1</sup>	
<b>Total</b> .....	<b>32,321</b>	<b>30,847</b>	<b>1,474</b>	<b>1,515</b>
Anaheim-Santa Ana, CA PMSA .....	4,094	3,846	248	116
Buffalo, NY CMSA .....	3,808	3,659	149	176
Dallas, TX PMSA .....	3,917	3,692	225	301
Ft. Worth-Arlington, TX PMSA .....	3,615	3,445	170	256
Milwaukee, WI PMSA .....	3,888	3,712	176	202
Phoenix, AZ MSA .....	4,401	4,150	251	161
Riverside-San Bernardino-Ontario, CA PMSA .....	4,640	4,489	151	194
San Diego, CA MSA .....	3,958	3,854	104	109

<sup>1</sup>Sample units were visited but occupants were not at home after repeated visits or were unavailable for some other reasons; or, for vacant housing units, no informed respondent could be found.

<sup>2</sup>Sample units were visited but did not provide information relevant to the housing inventory. This category includes sample units that were found not to be in the sampling frame.

- d. All sample housing units that were added since the previous survey in sample segments from the nonpermit universe. This sample represented additions to the housing inventory since the previous survey in nonpermit-issuing areas.

**1994 AHS-MS Original Sample Selection**

The 1994 AHS-MS original sample for the metropolitan areas which, in 1970, were 100-percent permit-issuing 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 areas (the 1970-based permit-issuing universe).
- 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 which 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 nonpermit universe).

In 1970, the Anaheim-Santa Ana, CA PMSA; Phoenix, AZ MSA; Riverside-San Bernardino-Ontario, CA PMSA; and San Diego, CA 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 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 table B.

**Table B. 1970 Housing Unit Strata**

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 one 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 also was selected to be in sample, thereby insuring 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 which produced an expected four sample units, thereby insuring the necessary designated sample size.

**Sample from the 1970-based new construction uni-**

**verse.** The second frame from which the metropolitan area sample was selected was a list of new construction building permits issued since 1970 (that is, the new construction universe). The sample selection from the list of new construction building permits was an independent operation within the metropolitan area. Under clerical selection 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 area became computerized. Under these 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 (that is, 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 (that is, 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 AHS-Metropolitan Area sample from the 1970-based permit-issuing universe and the 1970-based new construction universe within 1970-based area. The coverage deficiencies included the following types of 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 reports in the H170 series for the years 1976 through 1981.

**AHS-MS Initial 80-Redesign Sample Reduction and Sample Reinstatement**

The AHS-MS sample reduction dropped units from sample, whereas the 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.

As part of the 1980 redesign, the metropolitan areas in sample were scheduled for their initial interview in 1984, 1985, 1986 or 1987. 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 tenure which was based on the previous year's tenure status. In order to simplify field procedures, panels 1 through 3 (that is, a random one-fourth of the original sample) were dropped from sample whenever possible. More sample reductions were implemented separately for each tenure group (using different selection rates across the remaining panels).

**1988 AHS-MS Sample Reduction for Buffalo and Milwaukee**

These 1988 metropolitan areas had an expected sample size of 4,250 housing units uniformly distributed throughout nine panels (panels 4-12). Because of budget constraints, panel 4 was dropped from sample in 1988.

**1988 AHS-MS Sample Reduction for Dallas, Ft. Worth, and Phoenix**

These 1989 metropolitan areas had an expected sample size of 4,250 housing units uniformly distributed throughout nine panels (panels 4-12). Because of budget constraints, panel 11 and 12 were dropped from sample in 1989.

### 1991 AHS-MS Sample Reduction for San Diego

This 1991 metropolitan area had an expected sample size of 4,250 housing units uniformly distributed throughout nine panels (panels 4-12). Because of budget constraints, panel 12 was dropped from sample in 1991.

### 1994 AHS-MS Sample Reduction and Sample Reinstatement

For the current survey year, all eight MSA's had an expected sample size of 4,250 housing units. Panels that had been dropped in previous enumerations were originally reinstated. However, because of budget constraints, all panel 12 cases were later dropped from sample.

In addition, in the Ft. Worth-Arlington, TX PMSA, Parker County was dropped from the sample.

### ESTIMATION

The 1994 AHS-Metropolitan Area sample produced estimates pertaining to characteristics of the housing inventory at the time of the interview (that is, the 1994 housing inventory).

In addition, the 1994 AHS-MS produced estimates of the characteristics of units that were lost from the housing inventory since the 1988 survey for the Buffalo, NY CMSA and Milwaukee, WI PMSA; the 1989 survey for the Dallas, TX PMSA, Ft. Worth-Arlington, TX PMSA, and Phoenix, AZ MSA; the 1990 survey for the Anaheim-Santa Ana, CA PMSA and Riverside-San Bernardino-Ontario, CA PMSA; and the 1991 survey of the San Diego, CA MSA. These estimates are referred to as building loss estimates.

### Current Housing Inventory Estimates

**AHS-MS Weighting.** The AHS-MS sample housing units were weighted according to a three-step ratio estimation procedure. Before the implementation of the ratio estimation procedures, the basic weight (that is, 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 which were dropped because of permit unavailability. These noninterviews occur in the new construction universe.

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

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

### Type A Noninterview Adjustment

Type A noninterviews are sample units for which (a) occupants were not home, (b) occupants refused to be interviewed, or (c) occupants were unavailable for some other reason.

The adjustment was done on occupied units and was computed separately for the following:

- a. New construction.
- b. All other housing units (this includes the 1970-based permit-issuing universe, the 1970-based nonpermit-issuing universes, and the 1970-based new construction housing units built prior to the last survey).

For new construction units a Type A noninterview adjustment factor was computed separately for each of the 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 for the sample weight:

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

**AHS-MS 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 weights (that is, 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.

**Additional Ratio Estimation Procedures**

For the two ratio estimation 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 all areas:

$$\frac{\text{Independent estimate of mobile homes for the corresponding geographic subdivision of the metropolitan area}}{\text{Sample estimate of mobile homes for the corresponding geographic subdivision of the metropolitan area}}$$

The numerator of this ratio was determined using data from the 1990 census. The denominator was obtained using the existing weight of AHS sample mobile home units (that is, the basic weight times Type M factor times the Type A factor times the permit-issuing ratio estimate factor).

**Independent total housing unit ratio estimation.** The following ratio estimation procedure was applied in all areas:

$$\frac{\text{Independent estimate of the total housing inventory (excluding mobile homes) from the corresponding geographic subdivision of the metropolitan area}}{\text{Sample estimate of the total housing inventory (excluding mobile homes) from the corresponding geographic subdivision of the metropolitan area}}$$

The numerator of this ratio was determined from 1990 census data. The denominator was obtained using the existing weight of AHS sample units, excluding mobile homes (that is, the basic weight times the Type M factor times the Type A factor times the permit-issuing ratio estimate factor).

The computed ratio estimation factors were then applied to all appropriate housing 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 housing population, or different portions of it, is brought into agreement with known good estimates of the metropolitan area housing population.

**Building Loss Estimates**

Sample building loss units from the AHS-MS data were weighted using a three-stage ratio estimation procedure. Before the implementation of the ratio estimation procedures, the basic weight was adjusted to account for panel drops and Type M and Type A noninterviews.

**1994 Building loss adjustment factor.** Building loss estimates incorporate an adjustment unique to the building loss data. Panel 4 was dropped from the sample in 1988, panels 11 and 12 were dropped from sample in 1989, and panel 12 was dropped from sample in 1991. Panel 12 remained out of sample in all areas in 1994. Since we did not collect data on dropped units in 1988, 1989, or 1991, we cannot use them to make estimates of housing characteristics of building losses. Thus, sample housing units, from prior year dropped panels, that were losses to the housing inventory in 1994 were dropped from the building loss sample.

Since not all of the nine panels in sample for 1994 were used to make building loss estimates, the probability of selection was reduced. Consequently, the tables containing building loss data in these publications reflect this adjustment.

**Type M noninterview adjustment.** A description of this factor can be found in the previous section describing the AHS-MS weighting for the housing inventory. For building loss estimates, the Type M factor that was calculated the year the loss unit was interviewed (1988, 1989, 1990, or 1991) was used to compute the final weight.

**Type A noninterview adjustment.** A description of this factor can be found in the previous section describing

AHS-MS weighting for the housing inventory. For building loss estimates, a separate Type A noninterview factor was computed using only loss units and data from the prior year enumeration.

**AHS-MS ratio estimation procedure for the 1970-based permit-issuing universe.** A description of this factor can be found in the previous section describing the AHS-MS weighting for the housing inventory. For building loss estimates, the 1970-based permit-issuing factor that was calculated the year the loss unit was interviewed (1988, 1989, 1990, or 1991) was used to compute the final weight.

**Mobile home ratio estimation.** A description of this factor can be found in the previous section describing the AHS-MS weighting for the housing inventory. For building loss estimates, the mobile home ratio estimation factor that was calculated the year the loss unit was interviewed (1988, 1989, 1990, or 1991) was used to compute the final weight.

**Independent total housing unit ratio estimation.** A description of this factor can be found in the previous section describing the AHS-MS weighting for the housing inventory. For building loss estimates, the independent total housing unit ratio estimation factor that was calculated the year the loss unit was interviewed (1988, 1989, 1990, or 1991) was used to compute the final weight.