1979 Metropolitan Appendix B: Source and Accuracy

Table of Contents

Baltimore 3

Buffalo 14

Chicago 25

Cleveland 36

Denver 47

Hartford 58

Honolulu 69

Houston 80

Las Vegas 91

Miami 102

Milwaukee 113

Omaha 124

Portland 135

Raleigh 146

Seattle-Everett 157

Baltimore 1979

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-41
Annual Housing Survey App-41
Designation of sample housing units for the 1979 survey. App-41
Selection of the 1976 AHS-SMSA sample App-41
1976-1979 additions to the housing inventory App-43
Sample selection for the 1976 Coverage Improvement Program. App-43
Coverage improvement for deficiency 1 . . . . . App-43
Coverage improvement for deficiency 2 . . . . . App-43
Coverage improvement for deficiencies 3-6 . . . . . App-43
1970 Census of Population and Housing. App-44
ESTIMATION. App-44
1979 housing inventory App-44
1976-1979 lost housing units . . . App-45
1976 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES. App-46
Nonsampling errors. App-46
1970 census App-46
AHS-SMSA App-46
Coverage errors App-46
Rounding errors. App-46
Sampling errors for the AHS-SMSA sample App-47
Illustration of the use of the standard error tables App-47
Differences. App-48
Illustration of the computation of the standard error of a difference App-48
Medians App-48
Illustration of the computation of the 95-percent confidence interval of a median App-49
Standard error tables App-50

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 15 SMSA's in this report series (H-170-79) are based on data collected from the 1979 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 15 SMSA's, the data were collected during the 12-month period from April 1979 through March 1980 with one-twelfth of the sample housing units being visited each month.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1979 were interviewed previously in either 1975 or 1976 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1979 group of SMSA's, the largest SMSA from three of the four census regions of the United States is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All the remaining 12 SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the 1979 group are: Chicago, Ill.; Houston, Tex.; and Seattle-Everett, Wash.

The remaining SMSA's in the 1979 group are: Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.

In this SMSA, 5,078 housing units were eligible for interview. Of these sample housing units, 196 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 5,078 housing units eligible for interview, 358 units were visited but were not eligible for interview because they were condemned, unfit, demolished, converted to group quarters use, etc.

Designation of sample housing units for the 1979 survey—The sample housing units designated to be interviewed in the 1979 survey consisted of the following categories which are described in detail in the succeeding sections:

1. All sample housing units that were interviewed in the 1976 survey.
2. 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 1976 survey. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1979 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1976 survey. (This sample represented the housing units built in permit-issuing areas, since the 1976 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1976 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1976 survey.)
5. All sample housing units that were selected as part of the 1976 Coverage Improvement Program. (This sample represented most of the housing units which, until 1976, did not have a chance of selection.)

Selection of the 1976 AHS-SMSA sample—In 1970, the sample for the SMSA's which were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in

the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (nonpermit universe). In 1979, the following four SMSA's were 100-percent permit-issuing: Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; and Miami, Fla. The remaining 11 SMSA's contain a sample from the nonpermit universe.

Sampling operations, described in the following paragraphs, were performed separately within the central city and the balance of the SMSA for each of the sample frames. The overall sampling rate used to select the sample for each SMSA was determined by the size of the sample. Thus, for the three largest SMSA's, the overall sampling rate differed by central city and the balance of the SMSA, since the sample was divided equally between the central city and the balance of the SMSA. The remaining SMSA's had an overall sampling rate about the same for the sample selected from both the central city and the balance of the SMSA, since the sample was distributed proportionately between the central city and the balance of the SMSA according to the distribution of the total housing units in each sector.

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA 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 housing unit records were stratified by race of head (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 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household 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, for this SMSA, 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 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. 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 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 SMSA. 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.

The second frame from which this SMSA 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 this SMSA. Prior to sample selection, the list of permits was chronologically 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.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS 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 (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

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 are included.

1976-1979 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1976 survey, was selected to represent housing units built in permit-issuing areas since the 1976 survey. Sampling procedures were identical to those used in selecting the 1970-1976 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1976, to identify any housing units missed in the 1976 survey or any housing units added since the 1976 survey.

Sample selection for the 1976 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

1. New construction housing units from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes and trailers placed in parks either missed during the 1970 census or established since the 1970 census.
3. Housing units missed in the 1970 census.
4. Housing units converted to residential use that were nonresidential at the time of the 1970 census.
5. Housing units that have been moved onto their present site since the 1970 census.
6. Mobile homes and trailers placed outside parks since the 1970 census or vacant at the time of the 1970 census.

For each of the five SMSA's interviewed for the first time in 1975, a full implementation of the Coverage Improvement Program was conducted as a part of the 1979 AHS. For each of the 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with some updating and refining as a part of the 1979 AHS. The following discussion of coverage improvement procedures applies to both the 1976 and 1979 programs. For the 10 (1976) SMSA's, estimates of housing units added by a specific procedure reflect units added in 1976 as well as any additions that resulted from the updating and refining in 1979.

Coverage improvement for deficiency 1—A sample of new construction housing units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 1979 SMSA's. Two different procedures were used. For the SMSA's first interviewed in 1976, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. These new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units).

The first stage sample selection was a sample of permit offices, and the second stage a sample of the 1969 permits within each of the selected permit offices. (In the Honolulu, Hawaii, SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second stage.) For the third stage, structures of size three or more were divided into clusters of an expected size of four housing units and a sample of clusters was selected.

For the SMSA's first interviewed in 1975, housing units whose permits were issued before January 1970, but which were completed after April 1970, were identified from the Survey of Construction (SOC), a survey of building permits conducted monthly by the Bureau of the Census. These housing units were then sampled at one-third the rate of regular AHS units. These procedures added an estimated 2,386 new construction housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiency 2—In permit-issuing areas, a sample of mobile homes and trailers placed in parks that were missed by the census or established after the census was selected in two stages. First, a sample of tracts was selected and canvassed. All parks were listed and then matched back to the 1970 census to identify parks missed by the census and parks established after the census. Second, the parks were divided into clusters of an expected size of four sites and a sample of clusters was selected and interviewed. Each of the sample housing units represented the same number of units that the regular AHS sample housing units represented. Since research showed that there were no mobile home or trailer parks in the central cities of the Chicago, Ill.; Hartford, Conn.; and Portland, Oreg.-Wash.; SMSA's, this procedure was only implemented outside the central cities for these SMSA's. Based on a cost-benefit analysis, this procedure was updated as part of the 1979 AHS only for the Denver, Colo. and Las Vegas, Nev., SMSA's. This procedure added 2,384 housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed housing units were sampled by one of two procedures. The first procedure was designed to represent units from the following types of missed structures (structures that had no chance of selection for the AHS):

1. Structures missed in the 1970 census.
2. Structures that were completely nonresidential in the 1970 census but now contain residential housing units.
3. Mobile homes and trailers that had been placed outside parks since the 1970 census and have a utility hookup, or were on the site during the present survey but not occupied on April 1, 1970, or had no utility hookup but were occupied by persons with no usual residence elsewhere.
4. Housing units that had been moved onto their present site since the 1970 census.

Initially, a subsample of regular AHS units was selected at a rate of 1 in 24 for the Chicago, Ill.; Hartford, Conn.; Miami, Fla.; Milwaukee, Wis.; and Portland, Oreg.-Wash.; SMSA's, and at a rate of 1 in 22.2152 for the 10 remaining SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing units were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and housing units within

these structures were interviewed. In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. Although the Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Houston, Tex.; Omaha, Nebr.-Iowa; and Raleigh, N.C.; SMSA's, were interviewed for the first time in 1976 and had these sampling procedures implemented at that time, some 1976 coverage improvement assignments were not sent out to be interviewed due to time limitations. The sampling was completed as part of the 1979 AHS and this procedure added an estimated 12,652 housing units to the coverage of the housing inventory of this SMSA.

The second procedure was designed to represent missed housing units from structures represented in the AHS. These missed housing units were:

1. Housing units missed in the 1970 census.
2. Nonresidential space converted to residential use since the 1970 census in structures that contained some residential housing units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structures selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 1,694 housing units to the coverage of the housing inventory for this SMSA.

1970 Census of Population and Housing—The estimates pertaining to the 1970 housing inventory (i.e., the housing inventory that existed at the time of the 1970 census) are based on either 20-, 15-, or 5-percent sample data collected in April 1970 for the Decennial Census of Population and Housing. A detailed description of the sample design employed for the 1970 census can be obtained in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

ESTIMATION

The 1979 AHS sample produced two types of estimates for each SMSA: Estimates pertaining to characteristics of the housing inventory at the time of the interview (i.e., the 1979 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1976 (i.e., 1976-1979 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1979 housing inventory—The AHS estimates of characteristics of the 1979 housing inventory were produced using a 2-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and a 3-stage ratio estimation procedure for the remaining 11 SMSA's. Prior to 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 the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of one or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes and trailers from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample housing units from both the nonpermit universe and the coverage improvement universe (if units were not included in any of the previous cells).

The following first-stage 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 permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of housing units enumerated in areas under the jurisdiction of permit-issuing offices. The denominators of the ratios were obtained from weighted estimates of all the AHS sample housing units within the corresponding ratio estimation categories using the existing weight (i.e., the basic weight times the noninterview factor). The computed ratio estimation factor was then applied to the existing weight for each sample housing unit within the corresponding ratio estimation category.

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 permit-issuing universe. Prior to the AHS sample selection within each SMSA, 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 sample. Since the number of housing units deleted from the AHS universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the Houston, Tex., SMSA, a second ratio estimation procedure was employed for new construction housing units outside the Houston central city. Due to rapid growth in the

nonpermit-issuing universe the sampling methods did not adequately represent the new construction outside the Houston central city. Because of this deficiency, separate factors were computed for the following cells:

1. Conventional housing units (non-mobile home or trailer) outside the central city, which were built between the 1970 census and the 1976 survey.
2. Mobile homes and trailers outside the central city with a model year between the 1970 census and 1976 survey.
3. Conventional housing units outside the central city built between the 1976 and 1979 surveys.
4. Mobile homes and trailers outside the central city with a model year between the 1976 and 1979 surveys.

In a specific cell, the ratio estimation factor was equal to the following:

$$\frac{\text{Weighted count of housing units in the cell} + \text{Estimate of the undercoverage in the cell}}{\text{Weighted count of housing units in the cell}}$$

The computed second-stage ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell. For the Baltimore, Md.; Buffalo, N.Y.; Chicago, Ill.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; and Seattle-Everett, Wash.; SMSA's, a second-stage ratio estimation procedure was employed to adjust the central city/balance distribution of the weighted sample estimate of new construction housing units built since the last survey in permit-issuing areas to an independently derived estimate of this distribution.

This ratio estimation factor was calculated separately for the central city and balance of each SMSA and was applied to all new construction housing units from permit-issuing areas within the corresponding sector (central city or balance of the SMSA). This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}{\text{(Sample estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}$$

The independent estimates of new construction were based upon the number of authorized building permits which were determined from the Survey of Construction (SOC). The sample estimates were obtained from the weighted estimate of the AHS-SMSA sample housing units after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

This second-stage ratio estimation procedure was not implemented in the Honolulu, Hawaii; Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's. In the Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C., SMSA's, the central city permit offices, sources of SOC information on permits, had expanded their coverage of permits to include areas greater than those covered by the 1970 census central city definitions. Since the AHS-SMSA survey uses the 1970 definitions, use of this ratio estimation procedure would not correctly adjust the central city/balance distribution of new construction housing units. In the Honolulu, Hawaii, SMSA, this procedure could not be used because the permit office covers the entire SMSA and does not differentiate between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Neb.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and as a third-stage ratio estimation procedure for the remaining 11 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the October 1979 housing inventory of each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sectors. This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the October 1979 housing inventory for the corresponding sector of the SMSA)}}{\text{(AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA)}}$$

The numerator of this ratio was derived using 1970 and 1980 census counts. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight after the first-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and the existing weight after the second-stage ratio estimation procedure for the 11 remaining SMSA's.

The computed ratio estimation factors for the central city and balance of the SMSA's were then applied to the existing weight for all corresponding sample units and the resulting product was used as the final weight for tabulation purposes.

The effect of the total housing unit ratio estimation procedure, as well as the overall 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 SMSA as a whole, it can be expected that the sample estimates will be improved when the sample housing population, or different portions of it, are brought into agreement with known good estimates of the SMSA housing population.

1976-1979 lost housing units—The 1976-1979 lost housing unit (housing unit removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1976 housing inventory, as was described in the 1976 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1976-1979 lost housing units existed, by definition, in the 1976 housing inventory, there was a 1976 housing inventory weight associated with each 1976-1979 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1976-1979 lost housing units.

1976 estimation procedure—This report presents data on the housing characteristics of the 1976 housing inventory from the 1976 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1976.

Ratio estimation procedure of the 1970 Census of Population and Housing—This report presents data on the housing characteristics of the 1970 housing inventory from the 1970 Census of Population and Housing. The statistics based on 1970 census sample data employed a ratio estimation procedure which was applied separately for each of the three census samples. A detailed description of this ratio estimation procedure can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

RELIABILITY OF THE ESTIMATES

There are two types of possible errors associated with estimates based on data from sample surveys—sampling and nonsampling errors. The following is a description of the sampling and nonsampling errors associated with the AHS-SMSA sample and of the nonsampling errors associated with the 1970 census estimates. A description of the sampling errors associated with the sample estimates from the 1970 census can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

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; and other errors of collection, response, processing, 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 1970 Census of Population and Housing and the 1976 AHS-SMSA sample.

1970 census—A number of studies were conducted to measure two types of general errors associated with 1970 census estimates—"coverage" and "content" errors. The "coverage" errors determined how completely housing units were counted in the census and the extent to which occupancy status was erroneously reported. The "content" errors measured the accuracy of the data collected for enumerated housing units. These errors were measured by reinterviews, record checks, and other surveys.

The detailed results of these studies, as well as the methodology employed, are presented in the 1970 Census of Population and Housing Evaluation and Research Program Reports, Series PHC(E)-5, *The Coverage of Housing in the 1970 Census,* and PHC(E)-10, *Accuracy of Data for Selected Housing Characteristics as Measured by Reinterviews.*

AHS-SMSA—Results from the 1979 AHS-SMSA sample reinterview program were not available at the time this report was being prepared. However, a study was conducted for the 1976 AHS-SMSA sample. The results are presented in the Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1976."

Coverage errors—In errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional (non-mobile home or trailer) new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units. Due to 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, in which case, 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, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit.

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 estimated that the 1976 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 for the 1979 survey, the number of missed housing units may be considerably less for 1979.

The final ratio estimation procedure corrects for these deficiencies as far as the count of total housing is concerned; i.e., it adjusts the estimate of the total housing inventory to the best available estimate. However, biases of subtotals would still remain.

Rounding errors—For errors associated with processing, the rounding of estimates introduces another source of error in the data, the severity of which depends on 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-SMSA sample—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 each other. 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 due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by 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, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. 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.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two 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 (page App-50) are approximations to the standard errors of various estimates shown in this report for this SMSA. 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. Standard errors applicable to estimates of characteristics of the 1976 housing inventory can be found in the AHS Series H-170 reports for 1976.

Table I (page App-50) presents the standard errors applicable to estimates of characteristics of the 1979 housing inventory as well as estimates of characteristics of the 1976-1979 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 22,670 for the total SMSA, 13,560 for the central city of the SMSA, and 18,100 for the balance of the SMSA.

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.

Tables II (page App-50) presents the standard errors of estimated percentages for the 1979 housing inventory as well as estimated percentages of the 1976-1979 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

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 as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is 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:

$$(100)\left(\frac{x}{y}\right)\sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}$$

where: x = the numerator of the ratio
y = the denominator of the ratio
σ_x = the standard error of the numerator
σ_y = the standard error of the denominator

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1979 there were 454,300 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 6,070. 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
400,000	6,170
454,300	x
489,800	6,000

The entry for "x" is determined as follows by vertically interpolating between 6,170 and 6,000.

$$454{,}300-400{,}000 = 54{,}300$$
$$489{,}800-400{,}000 = 89{,}800$$

$$6{,}170 + \frac{54{,}300}{89{,}800}(6{,}000-6{,}170) = 6{,}070$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 448,230 to 460,370 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1979 owner-occupied housing units lies within a range computed in this way, would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 444,590 to 464,010 housing units with 90 percent confidence; and that the average estimate lies within the interval from 442,160 to 466,440 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 454,300 owner-occupied housing units, 89,800, or 19.8 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 19.8 percent is approximately 0.8 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	19.8	25 or 75
400,000.	0.7	a	1.0
454,300.		p	
500,000.	0.6	b	0.9

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

$$19.8-10.0 = 9.8$$
$$25.0-10.0 = 15.0$$

$$0.7 + \frac{9.8}{15.0}(1.0-0.7) = 0.9$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.6 and 0.9.

$$19.8-10.0 = 9.8$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{9.8}{15.0}(0.9-0.6) = 0.8$$

3. The entry for "p" is then determined by vertical interpolation between 0.9 and 0.8.

$$454{,}300-400{,}000 = 54{,}300$$
$$500{,}000-400{,}000 = 100{,}000$$

$$0.9 + \frac{54{,}300}{100{,}000}(0.8-0.9) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 19.0 to 20.6 percent; the 90-percent confidence interval is from 18.5 to 21.1 percent; and the 95-percent confidence interval is from 18.2 to 21.4 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. 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. Due to the overlap of the 1976 and 1979 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1979 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1979 there were 263,600 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 173,800. Table I shows the standard error of 263,600 is approximately 5,810, and the standard error of 89,800 is approximately 3,900. Therefore, the standard error of the estimated difference of 173,800 is about 7,000.

$$7{,}000 = \sqrt{(5{,}810)^2 + (3{,}900)^2}$$

Consequently, the 68-percent confidence interval for the 173,800 difference is from 166,800 to 180,800 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 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 162,600 to 185,000 housing units, and the 95-percent confidence interval is from 159,800 to 187,800 housing untis. Thus, we can conclude with 95 percent confidence that the number of 1979 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied housing units with two bedrooms, since the 95-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 deter-

mine an interval about the estimated median so that there is a stated degree of confidence 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. These two distribution intervals could be different.

For about 68 out of 100 possible samples, the average median from all possible samples would lie between these two values.

A two-standard-error confidence interval may be determined by finding the values corresponding to 50 percent plus and minus twice the standard error determined in step 1. For about 95 out of 100 possible samples, the average median from all possible samples would lie between these two values.

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.9. The base of the distribution from which this median was determined is 454,300 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 454,300 is approximately 1.0 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.0 and 52.0.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with 3 persons (for purposes of calculating the median, the category of 3 persons is considered to be from 2.5 to 3.5 persons) contains the 48.0 percent derived in step 2. About 191,100 housing units or 42.1 percent fall below this interval, and 99,200 housing units or 21.8 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(48.0-42.1)}{21.8} = 2.8$$

Similarly, the interval for owner-occupied housing units with 3 persons contains the 52.0 percent derived in step 2. About 191,100 housing units or 42.1 percent fall below this interval, and 99,200 housing units or 21.8 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(52.0-42.1)}{21.8} = 3.0$$

Thus, the 95-percent confidence interval ranges from 2.8 to 3.0 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1979 Housing Inventory and for Estimated Number of 1976-1979 Lost Units for the Baltimore, Md., SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city
0	190	180	200
100	190	180	200
200	200	190	200
500	310	300	310
700	370	350	370
1,000	440	420	440
2,500	690	660	700
5,000	980	930	990
10,000	1,380	1,310	1,390
25,000	2,160	2,010	2,160
50,000	3,000	2,720	2,980
75,000	3,610	3,160	3,540
100,000	4,100	3,440	3,960
150,000	4,830	3,660	4,530
200,000	5,360	3,460	4,830
250,000	5,740	2,780	4,920
300,000	5,990	700	4,790
302,800	6,000	—	4,780
400,000	6,170	—	3,810
489,800	6,000	—	—
500,000	5,960	—	—
600,000	5,290	—	—
700,000	3,960	—	—
792,600	—	—	—

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.3 for the total SMSA, 1.3 for the central city, and 1.2 for the balance (not in central city) estimates. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero since these estimates were derived from census data which are not subject to sampling error. However, these estimates are subject to the nonsampling errors associated with the 1970 and 1980 censuses and with the interpolation procedures used to derive these estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1979 Housing Inventory and for Estimated Percentages of 1976-1979 Lost Housing Units for the Baltimore, Md., SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
200	49.7	49.7	49.7	49.7	49.7	49.7
500	28.3	28.3	28.3	28.3	28.3	31.4
700	22.0	22.0	22.0	22.0	23.0	26.6
1,000	16.5	16.5	16.5	16.5	19.2	22.2
2,500	7.3	7.3	7.3	8.4	12.2	14.1
5,000	3.8	3.8	4.3	6.0	8.6	9.9
10,000	1.9	1.9	3.1	4.2	6.1	7.0
25,000	0.8	0.9	1.9	2.7	3.8	4.4
50,000	0.4	0.6	1.4	1.9	2.7	3.1
75,000	0.3	0.5	1.1	1.5	2.2	2.6
100,000	0.2	0.4	1.0	1.3	1.9	2.2
150,000	0.13	0.4	0.8	1.1	1.6	1.8
200,000	0.10	0.3	0.7	0.9	1.4	1.6
250,000	0.08	0.3	0.6	0.8	1.2	1.4
300,000	0.07	0.3	0.6	0.8	1.1	1.3
400,000	0.05	0.2	0.5	0.7	1.0	1.1
500,000	0.04	0.2	0.4	0.6	0.9	1.0
600,000	0.03	0.2	0.4	0.5	0.8	0.9
700,000	0.03	0.2	0.4	0.5	0.7	0.8
800,000	0.02	0.2	0.3	0.5	0.7	0.8

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than 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 1.3 for the total SMSA, 1.3 for the central city, and 1.2 for the balance (not in central city).

Buffalo 1979

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-41
Annual Housing Survey App-41
Designation of sample housing units for the 1979 survey. App-41
Selection of the 1976 AHS-SMSA sample App-41
1976-1979 additions to the housing inventory App-43
Sample selection for the 1976 Coverage Improvement Program. App-43
Coverage improvement for deficiency 1 . . . . . App-43
Coverage improvement for deficiency 2 . . . . . App-43
Coverage improvement for deficiencies 3-6 . . . . . App-43
1970 Census of Population and Housing. App-44
ESTIMATION. App-44
1979 housing inventory App-44
1976-1979 lost housing units . . . App-45
1976 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES. App-46
Nonsampling errors. App-46
1970 census App-46
AHS-SMSA App-46
Coverage errors App-46
Rounding errors. App-46
Sampling errors for the AHS-SMSA sample App-47
Illustration of the use of the standard error tables App-47
Differences. App-48
Illustration of the computation of the standard error of a difference App-48
Medians App-48
Illustration of the computation of the 95-percent confidence interval of a median App-49
Standard error tables App-50

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 15 SMSA's in this report series (H-170-79) are based on data collected from the 1979 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 15 SMSA's, the data were collected during the 12-month period from April 1979 through March 1980 with one-twelfth of the sample housing units being visited each month.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1979 were interviewed previously in either 1975 or 1976 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1979 group of SMSA's, the largest SMSA from three of the four census regions of the United States is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All the remaining 12 SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the 1979 group are: Chicago, Ill.; Houston, Tex.; and Seattle-Everett, Wash.

The remaining SMSA's in the 1979 group are: Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.

In this SMSA, 5,008 housing units were eligible for interview. Of these sample housing units, 213 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 5,008 housing units eligible for interview, 232 units were visited but were not eligible for interview because they were condemned, unfit, demolished, converted to group quarters use, etc.

Designation of sample housing units for the 1979 survey—The sample housing units designated to be interviewed in the 1979 survey consisted of the following categories which are described in detail in the succeeding sections.

1. All sample housing units that were interviewed in the 1976 survey.
2. 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 1976 survey. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1979 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1976 survey. (This sample represented the housing units built in permit-issuing areas, since the 1976 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1976 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1976 survey.)
5. All sample housing units that were selected as part of the 1976 Coverage Improvement Program. (This sample represented most of the housing units which, until 1976, did not have a chance of selection.)

Selection of the 1976 AHS-SMSA sample—In 1970, the sample for the SMSA's which were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in

the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (nonpermit universe). In 1979, the following four SMSA's were 100-percent permit-issuing: Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; and Miami, Fla. The remaining 11 SMSA's contain a sample from the nonpermit universe.

Sampling operations, described in the following paragraphs, were performed separately within the central city and the balance of the SMSA for each of the sample frames. The overall sampling rate used to select the sample for each SMSA was determined by the size of the sample. Thus, for the three largest SMSA's, the overall sampling rate differed by central city and the balance of the SMSA, since the sample was divided equally between the central city and the balance of the SMSA. The remaining SMSA's had an overall sampling rate about the same for the sample selected from both the central city and the balance of the SMSA, since the sample was distributed proportionately between the central city and the balance of the SMSA according to the distribution of the total housing units in each sector.

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA 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 housing unit records were stratified by race of head (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 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household 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, for this SMSA, 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 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. 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 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 SMSA. 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.

The second frame from which this SMSA 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 this SMSA. Prior to sample selection, the list of permits was chronologically 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.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS 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 (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

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 are included.

1976-1979 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1976 survey, was selected to represent housing units built in permit-issuing areas since the 1976 survey. Sampling procedures were identical to those used in selecting the 1970-1976 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1976, to identify any housing units missed in the 1976 survey or any housing units added since the 1976 survey.

Sample selection for the 1976 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

1. New construction housing units from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes and trailers placed in parks either missed during the 1970 census or established since the 1970 census.
3. Housing units missed in the 1970 census.
4. Housing units converted to residential use that were nonresidential at the time of the 1970 census.
5. Housing units that have been moved onto their present site since the 1970 census.
6. Mobile homes and trailers placed outside parks since the 1970 census or vacant at the time of the 1970 census.

For each of the five SMSA's interviewed for the first time in 1975, a full implementation of the Coverage Improvement Program was conducted as a part of the 1979 AHS. For each of the 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with some updating and refining as a part of the 1979 AHS. The following discussion of coverage improvement procedures applies to both the 1976 and 1979 programs. For the 10 (1976) SMSA's, estimates of housing units added by a specific procedure reflect units added in 1976 as well as any additions that resulted from the updating and refining in 1979.

Coverage improvement for deficiency 1—A sample of new construction housing units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 1979 SMSA's. Two different procedures were used. For the SMSA's first interviewed in 1976, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. These new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units).

The first stage sample selection was a sample of permit offices, and the second stage a sample of the 1969 permits within each of the selected permit offices. (In the Honolulu, Hawaii, SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second stage.) For the third stage, structures of size three or more were divided into clusters of an expected size of four housing units and a sample of clusters was selected.

For the SMSA's first interviewed in 1975, housing units whose permits were issued before January 1970, but which were completed after April 1970, were identified from the Survey of Construction (SOC), a survey of building permits conducted monthly by the Bureau of the Census. These housing units were then sampled at one-third the rate of regular AHS units. These procedures added an estimated 772 new construction housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiency 2—In permit-issuing areas, a sample of mobile homes and trailers placed in parks that were missed by the census or established after the census was selected in two stages. First, a sample of tracts was selected and canvassed. All parks were listed and then matched back to the 1970 census to identify parks missed by the census and parks established after the census. Second, the parks were divided into clusters of an expected size of four sites and a sample of clusters was selected and interviewed. Each of the sample housing units represented the same number of units that the regular AHS sample housing units represented. Since research showed that there were no mobile home or trailer parks in the central cities of the Chicago, Ill.; Hartford, Conn.; and Portland, Oreg.-Wash.; SMSA's, this procedure was only implemented outside the central cities for these SMSA's. Based on a cost-benefit analysis, this procedure was updated as part of the 1979 AHS only for the Denver, Colo. and Las Vegas, Nev., SMSA's. This procedure added 667 housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed housing units were sampled by one of two procedures. The first procedure was designed to represent units from the following types of missed structures (structures that had no chance of selection for the AHS):

1. Structures missed in the 1970 census.
2. Structures that were completely nonresidential in the 1970 census but now contain residential housing units.
3. Mobile homes and trailers that had been placed outside parks since the 1970 census and have a utility hookup, or were on the site during the present survey but not occupied on April 1, 1970, or had no utility hookup but were occupied by persons with no usual residence elsewhere.
4. Housing units that had been moved onto their present site since the 1970 census.

Initially, a subsample of regular AHS units was selected at a rate of 1 in 24 for the Chicago, Ill.; Hartford, Conn.; Miami, Fla.; Milwaukee, Wis.; and Portland, Oreg.-Wash.; SMSA's, and at a rate of 1 in 22.2152 for the 10 remaining SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing units were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and housing units within

these structures were interviewed. In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. Although the Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Houston, Tex.; Omaha, Nebr.-Iowa; and Raleigh, N.C.; SMSA's, were interviewed for the first time in 1976 and had these sampling procedures implemented at that time, some 1976 coverage improvement assignments were not sent out to be interviewed due to time limitations. The sampling was completed as part of the 1979 AHS and this procedure added an estimated 5,282 housing units to the coverage of the housing inventory of this SMSA.

The second procedure was designed to represent missed housing units from structures represented in the AHS. These missed housing units were:

1. Housing units missed in the 1970 census.
2. Nonresidential space converted to residential use since the 1970 census in structures that contained some residential housing units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structures selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 1,773 housing units to the coverage of the housing inventory for this SMSA.

1970 Census of Population and Housing—The estimates pertaining to the 1970 housing inventory (i.e., the housing inventory that existed at the time of the 1970 census) are based on either 20-, 15-, or 5-percent sample data collected in April 1970 for the Decennial Census of Population and Housing. A detailed description of the sample design employed for the 1970 census can be obtained in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

ESTIMATION

The 1979 AHS sample produced two types of estimates for each SMSA: Estimates pertaining to characteristics of the housing inventory at the time of the interview (i.e., the 1979 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1976 (i.e., 1976-1979 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1979 housing inventory—The AHS estimates of characteristics of the 1979 housing inventory were produced using a 2-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and a 3-stage ratio estimation procedure for the remaining 11 SMSA's. Prior to 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 the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of one or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes and trailers from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample housing units from both the nonpermit universe and the coverage improvement universe (if units were not included in any of the previous cells).

The following first-stage 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 permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of housing units enumerated in areas under the jurisdiction of permit-issuing offices. The denominators of the ratios were obtained from weighted estimates of all the AHS sample housing units within the corresponding ratio estimation categories using the existing weight (i.e., the basic weight times the noninterview factor). The computed ratio estimation factor was then applied to the existing weight for each sample housing unit within the corresponding ratio estimation category.

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 permit-issuing universe. Prior to the AHS sample selection within each SMSA, 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 sample. Since the number of housing units deleted from the AHS universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the Houston, Tex., SMSA, a second ratio estimation procedure was employed for new construction housing units outside the Houston central city. Due to rapid growth in the

nonpermit-issuing universe the sampling methods did not adequately represent the new construction outside the Houston central city. Because of this deficiency, separate factors were computed for the following cells:

1. Conventional housing units (non-mobile home or trailer) outside the central city, which were built between the 1970 census and the 1976 survey.
2. Mobile homes and trailers outside the central city with a model year between the 1970 census and 1976 survey.
3. Conventional housing units outside the central city built between the 1976 and 1979 surveys.
4. Mobile homes and trailers outside the central city with a model year between the 1976 and 1979 surveys.

In a specific cell, the ratio estimation factor was equal to the following:

$$\frac{\text{Weighted count of housing units in the cell} + \text{Estimate of the undercoverage in the cell}}{\text{Weighted count of housing units in the cell}}$$

The computed second-stage ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell. For the Baltimore, Md.; Buffalo, N.Y.; Chicago, Ill.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; and Seattle-Everett, Wash.; SMSA's, a second-stage ratio estimation procedure was employed to adjust the central city/balance distribution of the weighted sample estimate of new construction housing units built since the last survey in permit-issuing areas to an independently derived estimate of this distribution.

This ratio estimation factor was calculated separately for the central city and balance of each SMSA and was applied to all new construction housing units from permit-issuing areas within the corresponding sector (central city or balance of the SMSA). This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}{\text{(Sample estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}$$

The independent estimates of new construction were based upon the number of authorized building permits which were determined from the Survey of Construction (SOC). The sample estimates were obtained from the weighted estimate of the AHS-SMSA sample housing units after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

This second-stage ratio estimation procedure was not implemented in the Honolulu, Hawaii; Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's. In the Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C., SMSA's, the central city permit offices, sources of SOC information on permits, had expanded their coverage of permits to include areas greater than those covered by the 1970 census central city definitions. Since the AHS-SMSA survey uses the 1970 definitions, use of this ratio estimation procedure would not correctly adjust the central city/balance distribution of new construction housing units. In the Honolulu, Hawaii, SMSA, this procedure could not be used because the permit office covers the entire SMSA and does not differentiate between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Neb.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and as a third-stage ratio estimation procedure for the remaining 11 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the October 1979 housing inventory of each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sectors. This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the October 1979 housing inventory for the corresponding sector of the SMSA)}}{\text{(AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA)}}$$

The numerator of this ratio was derived using 1970 and 1980 census counts. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight after the first-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and the existing weight after the second-stage ratio estimation procedure for the 11 remaining SMSA's.

The computed ratio estimation factors for the central city and balance of the SMSA's were then applied to the existing weight for all corresponding sample units and the resulting product was used as the final weight for tabulation purposes.

The effect of the total housing unit ratio estimation procedure, as well as the overall 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 SMSA as a whole, it can be expected that the sample estimates will be improved when the sample housing population, or different portions of it, are brought into agreement with known good estimates of the SMSA housing population.

1976-1979 lost housing units—The 1976-1979 lost housing unit (housing unit removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1976 housing inventory, as was described in the 1976 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1976-1979 lost housing units existed, by definition, in the 1976 housing inventory, there was a 1976 housing inventory weight associated with each 1976-1979 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1976-1979 lost housing units.

1976 estimation procedure—This report presents data on the housing characteristics of the 1976 housing inventory from the 1976 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1976.

Ratio estimation procedure of the 1970 Census of Population and Housing—This report presents data on the housing characteristics of the 1970 housing inventory from the 1970 Census of Population and Housing. The statistics based on 1970 census sample data employed a ratio estimation procedure which was applied separately for each of the three census samples. A detailed description of this ratio estimation procedure can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

RELIABILITY OF THE ESTIMATES

There are two types of possible errors associated with estimates based on data from sample surveys—sampling and nonsampling errors. The following is a description of the sampling and nonsampling errors associated with the AHS-SMSA sample and of the nonsampling errors associated with the 1970 census estimates. A description of the sampling errors associated with the sample estimates from the 1970 census can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

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; and other errors of collection, response, processing, 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 1970 Census of Population and Housing and the 1976 AHS-SMSA sample.

1970 census—A number of studies were conducted to measure two types of general errors associated with 1970 census estimates—"coverage" and "content" errors. The "coverage" errors determined how completely housing units were counted in the census and the extent to which occupancy status was erroneously reported. The "content" errors measured the accuracy of the data collected for enumerated housing units. These errors were measured by reinterviews, record checks, and other surveys.

The detailed results of these studies, as well as the methodology employed, are presented in the 1970 Census of Population and Housing Evaluation and Research Program Reports, Series PHC(E)-5, *The Coverage of Housing in the 1970 Census,* and PHC(E)-10, *Accuracy of Data for Selected Housing Characteristics as Measured by Reinterviews.*

AHS-SMSA—Results from the 1979 AHS-SMSA sample reinterview program were not available at the time this report was being prepared. However, a study was conducted for the 1976 AHS-SMSA sample. The results are presented in the Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1976."

Coverage errors—In errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional (non-mobile home or trailer) new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units. Due to 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, in which case, 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, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit.

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 estimated that the 1976 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 for the 1979 survey, the number of missed housing units may be considerably less for 1979.

The final ratio estimation procedure corrects for these deficiencies as far as the count of total housing is concerned; i.e., it adjusts the estimate of the total housing inventory to the best available estimate. However, biases of subtotals would still remain.

Rounding errors—For errors associated with processing, the rounding of estimates introduces another source of error in the data, the severity of which depends on 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-SMSA sample—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 each other. 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 due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by 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, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. 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.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two 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 (page App-50) are approximations to the standard errors of various estimates shown in this report for this SMSA. 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. Standard errors applicable to estimates of characteristics of the 1976 housing inventory can be found in the AHS Series H-170 reports for 1976.

Table I (page App-50) presents the standard errors applicable to estimates of characteristics of the 1979 housing inventory as well as estimates of characteristics of the 1976-1979 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 11,840 for the total SMSA, 6,320 for the central city of the SMSA, and 9,960 for the balance of the SMSA.

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 (page App-50) presents the standard errors of estimated percentages for the 1979 housing inventory as well as estimated percentages of the 1976-1979 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

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 as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is 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:

$$(100)\left(\frac{x}{y}\right)\sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}$$

where: x = the numerator of the ratio
y = the denominator of the ratio
σ_x = the standard error of the numerator
σ_y = the standard error of the denominator

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1979 there were 288,100 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 3,380. 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
250,000	3,470
288,100	x
300,000	3,350

The entry for "x" is determined as follows by vertically interpolating between 3,470 and 3,350.

$$288{,}100-250{,}000 = 38{,}100$$
$$300{,}000-250{,}000 = 50{,}000$$

$$3{,}470 + \frac{38{,}100}{50{,}000}(3{,}350-3{,}470) = 3{,}380$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 284,720 to 291,480 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1979 owner-occupied housing units lies within a range computed in this way, would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 282,690 to 293,510 housing units with 90 percent confidence; and that the average estimate lies within the interval from 281,340 to 294,860 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 288,100 owner-occupied housing units, 58,100, or 20.2 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 20.2 percent is approximately 0.7 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	20.2	25 or 75
250,000	0.6	a	0.9
288,100		p	
300,000	0.6	b	0.8

1. The entry for cell "a" is determined by horizontal interpolation between 0.6 and 0.9

$$20.2-10.0 = 10.2$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{10.2}{15.0}(0.9-0.6) = 0.8$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.6 and 0.8.

$$20.2-10.0 = 10.2$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{10.2}{15.0}(0.8-0.6) = 0.7$$

3. The entry for "p" is then determined by vertical interpolation between 0.8 and 0.7.

$$288{,}100-250{,}000 = 38{,}100$$
$$300{,}000-250{,}000 = 50{,}000$$

$$0.8 + \frac{38{,}100}{50{,}000}(0.7-0.8) = 0.7$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 19.5 to 20.9 percent; the 90-percent confidence interval is from 19.1 to 21.3 percent; and the 95-percent confidence interval is from 18.8 to 21.6 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. 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. Due to the overlap of the 1976 and 1979 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1979 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1979 there were 158,500 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 100,400. Table I shows the standard error of 158,500 is approximately 3,290, and the standard error of 58,100 is approximately 2,270. Therefore, the standard error of the estimated difference of 100,400 is about 4,000.

$$4{,}000 = \sqrt{(3{,}290)^2 + (2{,}270)^2}$$

Consequently, the 68-percent confidence interval for the 100,400 difference is from 96,400 to 104,400 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 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 94,000 to 106,800 housing units, and the 95-percent confidence interval is from 92,400 to 108,400 housing units. Thus, we can conclude with 95 percent confidence that the number of 1979 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied housing units with two bedrooms, since the 95-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 deter-

mine an interval about the estimated median so that there is a stated degree of confidence 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 known into which interval of the distribution the upper percentage limit falls. These two distribution intervals could be different.

For about 68 out of 100 possible samples, the average median from all possible samples would lie between these two values.

A two-standard-error confidence interval may be determined by finding the values corresponding to 50 percent plus and minus twice the standard error determined in step 1. For about 95 out of 100 possible samples, the average median from all possible samples would lie between these two values.

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.8. The base of the distribution from which this median was determined is 288,100 housing units.

1.0 — 1.1

1. Interpolation using table II shows that the standard error of 50 percent on a base of 288,100 is approximately 0.9 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.2 and 51.8. 50 ± 2(1.6)
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with 3 persons (for purposes of calculating the median, the category of 3 persons is considered to be from 2.5 to 3.5 persons) contains the 48.2 percent derived in step 2. About 129,900 housing units or 45.1 percent fall below this interval, and 50,900 housing units or 17.7 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(48.2-45.1)}{17.7} = 2.7 \text{ person}$$

Similarly, the interval for owner-occupied housing units with 3 persons also contains the 51.8 percent derived in step 2. About 129,900 housing units or 45.1 percent fall below this interval, and 50,900 housing units or 17.7 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(51.8-45.1)}{17.7} = 2.9 \text{ person}$$

Thus, the 95-percent confidence interval ranges from 2.7 to 2.9 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1979 Housing Inventory and for Estimated Number of 1976-1979 Lost Units for the Buffalo, N.Y., SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city
0	100	100	100
100	100	100	100
200	140	140	140
500	230	220	220
700	270	260	270
1,000	320	310	320
2,500	500	490	500
5,000	710	690	710
10,000	1,000	960	990
25,000	1,560	1,440	1,530
50,000	2,140	1,830	2,060
75,000	2,540	1,960	2,410
100,000	2,840	1,890	2,630
150,000	3,240	810	2,820
157,000	3,280	—	2,830
200,000	3,440	—	2,720
250,000	3,470	—	2,290
300,000	3,350	—	1,220
315,400	3,280	—	—
400,000	2,500	—	—
472,300	—	—	—

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2 for the total SMSA, 1.3 for the central city, and 1.2 for the balance (not in central city) estimates. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero since these estimates were derived from census data which are not subject to sampling error. However, these estimates are subject to the nonsampling errors associated with the 1970 and 1980 censuses and with the interpolation procedures used to derive these estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1979 Housing Inventory and for Estimated Percentages of 1976-1979 Lost Housing Units for the Buffalo, N.Y., SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
200	33.9	33.9	33.9	33.9	33.9	35.8
500	17.0	17.0	17.0	17.0	19.6	22.6
700	12.8	12.8	12.8	12.8	16.6	19.1
1,000	9.3	9.3	9.3	9.6	13.9	16.0
2,500	3.9	3.9	4.4	6.1	8.8	10.1
5,000	2.0	2.0	3.1	4.3	6.2	7.2
10,000	1.0	1.0	2.2	3.0	4.4	5.1
25,000	0.4	0.6	1.4	1.9	2.8	3.2
50,000	0.2	0.5	1.0	1.4	2.0	2.3
75,000	0.14	0.4	0.8	1.1	1.6	1.8
100,000	0.10	0.3	0.7	1.0	1.4	1.6
150,000	0.17	0.3	0.6	0.8	1.1	1.3
200,000	0.15	0.2	0.5	0.7	1.0	1.1
250,000	0.04	0.2	0.4	0.6	0.9	1.0
300,000	0.03	0.2	0.4	0.6	0.8	0.9
400,000	0.03	0.2	0.3	0.5	0.7	0.8
500,000	0.02	0.14	0.3	0.4	0.6	0.7

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than 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 1.2 for the total SMSA, 1.3 for the central city, and 1.2 for the balance (not in central city).

700
520
180

Chicago 1979

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1979 survey App-42
Selection of the 1975 AHS-SMSA sample App-42
1975-1979 additions to the housing inventory App-44
Sample selection for the 1979 Coverage Improvement Program. App-44
Coverage improvement for deficiency 1 . . . . . App-44
Coverage improvement for deficiency 2 . . . . . App-44
Coverage improvement for deficiencies 3-6 . . . . . App-44
1970 Census of Population and Housing. App-45
ESTIMATION. App-45
1979 housing inventory App-45
1975-1979 lost housing units . . . App-46
1975 estimation procedure App-47
Ratio estimation procedure of the 1970 Census of Population and Housing App-47
RELIABILITY OF THE ESTIMATES. App-47
Nonsampling errors. App-47
1970 census App-47
AHS-SMSA App-47
Coverage errors App-47
Rounding errors. App-47
Sampling errors for the AHS-SMSA sample App-48
Illustration of the use of the standard error tables App-48
Differences. App-49
Illustration of the computation of the standard error of a difference App-49
Medians App-49
Illustration of the computation of the 95-percent confidence interval of a median App-50
Standard error tables App-51

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 15 SMSA's in this report series (H-170-79) are based on data collected from the 1979 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 15 SMSA's, the data were collected during the 12-month period from April 1979 through March 1980 with one-twelfth of the sample housing units being visited each month.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1979 were interviewed previously in either 1975 or 1976 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1979 group of SMSA's, the largest SMSA from three of the four census regions of the United States is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All the remaining 12 SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the 1979 group are: Chicago, Ill.; Houston, Tex.; and Seattle-Everett, Wash.

The remaining SMSA's in the 1979 group are: Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.

In this SMSA, 15,280 housing units were eligible for interview. Of these sample housing units, 930 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 15,280 housing units eligible for interview, 976 units were visited but were not eligible for interview because they were condemned, unfit, demolished, converted to group quarters use, etc.

Designation of sample housing units for the 1979 survey—The sample housing units designated to be interviewed in the 1979 survey consisted of the following categories which are described in detail in the succeeding sections.

1. All sample housing units that were interviewed in the 1975 survey.
2. 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 1975 survey. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1979 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1975 survey. (This sample represented the housing units built in permit-issuing areas, since the 1975 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1975 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1975 survey.)
5. All sample housing units that were selected as part of the 1979 Coverage Improvement Program. (This sample represented most of the housing units which, until 1979, did not have a chance of selection.)

Selection of the 1975 AHS-SMSA sample—In 1970, the sample for the SMSA's which were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in

the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). In 1979, the following four SMSA's were 100-percent permit-issuing: Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; and Miami, Fla. The remaining 11 SMSA's contain a sample from the nonpermit universe.

Sampling operations, described in the following paragraphs, were performed separately within the central city and the balance of the SMSA for each of the sample frames. The overall sampling rate used to select the sample for each SMSA was determined by the size of the sample. Thus, for the three largest SMSA's, the overall sampling rate differed by central city and the balance of the SMSA, since the sample was divided equally between the central city and the balance of the SMSA. The remaining SMSA's had an overall sampling rate about the same for the sample selected from both the central city and the balance of the SMSA, since the sample was distributed proportionately between the central city and the balance of the SMSA according to the distribution of the total housing units in each sector.

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA 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 housing unit records were stratified by race of head (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 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household 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, for this SMSA, 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 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. 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 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 SMSA. 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.

The second frame from which this SMSA 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 this SMSA. Prior to sample selection, the list of permits was chronologically 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.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS 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 (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

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 are included.

1975-1979 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1975 survey, was selected to represent housing units built in permit-issuing areas since the 1975 survey. Sampling procedures were identical to those used in selecting the 1970-1975 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1975, to identify any housing units missed in the 1975 survey or any housing units added since the 1975 survey.

Sample selection for the 1979 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

1. New construction housing units from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes and trailers placed in parks either missed during the 1970 census or established since the 1970 census.
3. Housing units missed in the 1970 census.
4. Housing units converted to residential use that were nonresidential at the time of the 1970 census.
5. Housing units that have been moved onto their present site since the 1970 census.
6. Mobile homes and trailers placed outside parks since the 1970 census or vacant at the time of the 1970 census.

For each of the five SMSA's interviewed for the first time in 1975, a full implementation of the Coverage Improvement Program was conducted as a part of the 1979 AHS. For each of the 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with some updating and refining as a part of the 1979 AHS. The following discussion of coverage improvement procedures applies to both the 1976 and 1979 programs. For the 10 (1976) SMSA's, estimates of housing units added by a specific procedure reflect units added in 1976 as well as any additions that resulted from the updating and refining in 1979.

Coverage improvement for deficiency 1—A sample of new construction housing units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 1979 SMSA's. Two different procedures were used. For the SMSA's first interviewed in 1976, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. These new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units).

The first stage sample selection was a sample of permit offices, and the second stage a sample of the 1969 permits within each of the selected permit offices. (In the Honolulu, Hawaii, SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second stage.) For the third stage, structures of size three or more were divided into clusters of an expected size of four housing units and a sample of clusters was selected.

For the SMSA's first interviewed in 1975, housing units whose permits were issued before January 1970, but which were completed after April 1970, were identified from the Survey of Construction (SOC), a survey of building permits conducted monthly by the Bureau of the Census. These housing units were then sampled at one-third the rate of regular AHS units. These procedures added an estimated 24,078 new construction housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiency 2—In permit-issuing areas, a sample of mobile homes and trailers placed in parks that were missed by the census or established after the census was selected in two stages. First, a sample of tracts was selected and canvassed. All parks were listed and then matched back to the 1970 census to identify parks missed by the census and parks established after the census. Second, the parks were divided into clusters of an expected size of four sites and a sample of clusters was selected and interviewed. Each of the sample housing units represented the same number of units that the regular AHS sample housing units represented. Since research showed that there were no mobile home or trailer parks in the central cities of the Chicago, Ill.; Hartford, Conn.; and Portland, Oreg.-Wash.; SMSA's, this procedure was only implemented outside the central cities for these SMSA's. Based on a cost-benefit analysis, this procedure was updated as part of the 1979 AHS only for the Denver, Colo. and Las Vegas, Nev., SMSA's.

Coverage improvement for deficiencies 3-6—The remaining missed housing units were sampled by one of two procedures. The first procedure was designed to represent units from the following types of missed structures (structures that had no chance of selection for the AHS):

1. Structures missed in the 1970 census.
2. Structures that were completely nonresidential in the 1970 census but now contain residential housing units.
3. Mobile homes and trailers that had been placed outside parks since the 1970 census and have a utility hookup, or were on the site during the present survey but not occupied on April 1, 1970, or had no utility hookup but were occupied by persons with no usual residence elsewhere.
4. Housing units that had been moved onto their present site since the 1970 census.

Initially, a subsample of regular AHS units was selected at a rate of 1 in 24 for the Chicago, Ill.; Hartford, Conn.; Miami, Fla.; Milwaukee, Wis.; and Portland, Oreg.-Wash.; SMSA's, and at a rate of 1 in 22.2152 for the 10 remaining SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing units were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and housing units within these structures were interviewed. In cases where the interviewer workload would have been too great, a representative subsample

of units within these structures was selected. Although the Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Houston, Tex.; Omaha, Nebr.-Iowa; and Raleigh, N.C.; SMSA's, were interviewed for the first time in 1976 and had these sampling procedures implemented at that time, some 1976 coverage improvement assignments were not sent out to be interviewed due to time limitations. The sampling was completed as part of the 1979 AHS and this procedure added an estimated 35,809 housing units to the coverage of the housing inventory of this SMSA.

The second procedure was designed to represent missed housing units from structures represented in the AHS. These missed housing units were:

1. Housing units missed in the 1970 census.
2. Nonresidential space converted to residential use since the 1970 census in structures that contained some residential housing units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structures selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 27,807 housing units to the coverage of the housing inventory for this SMSA.

1970 Census of Population and Housing—The estimates pertaining to the 1970 housing inventory (i.e., the housing inventory that existed at the time of the 1970 census) are based on either 20-, 15-, or 5-percent sample data collected in April 1970 for the Decennial Census of Population and Housing. A detailed description of the sample design employed for the 1970 census can be obtained in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

ESTIMATION

The 1979 AHS sample produced two types of estimates for each SMSA: Estimates pertaining to characteristics of the housing inventory at the time of the interview (i.e., the 1979 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1975 (i.e., 1975-1979 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1979 housing inventory—The AHS estimates of characteristics of the 1979 housing inventory were produced using a 2-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and a 3-stage ratio estimation procedure for the remaining 11 SMSA's. Prior to 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 the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of one or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes and trailers from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample housing units from both the nonpermit universe and the coverage improvement universe (if units were not included in any of the previous cells).

The following first-stage 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 permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of housing units enumerated in areas under the jurisdiction of permit-issuing offices. The denominators of the ratios were obtained from weighted estimates of all the AHS sample housing units within the corresponding ratio estimation categories using the existing weight (i.e., the basic weight times the noninterview factor). The computed ratio estimation factor was then applied to the existing weight for each sample housing unit within the corresponding ratio estimation category.

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 permit-issuing universe. Prior to the AHS sample selection within each SMSA, 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 sample. Since the number of housing units deleted from the AHS universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the Houston, Tex., SMSA, a second ratio estimation procedure was employed for new construction housing units outside the Houston central city. Due to rapid growth in the

nonpermit-issuing universe the sampling methods did not adequately represent the new construction outside the Houston central city. Because of this deficiency, separate factors were computed for the following cells:

1. Conventional housing units (non-mobile home or trailer) outside the central city, which were built between the 1970 census and the 1976 survey.
2. Mobile homes and trailers outside the central city with a model year between the 1970 census and 1976 survey.
3. Conventional housing units outside the central city built between the 1976 and 1979 surveys.
4. Mobile homes and trailers outside the central city with a model year between the 1976 and 1979 surveys.

In a specific cell, the ratio estimation factor was equal to the following:

$$\frac{\text{Weighted count of housing units in the cell} + \text{Estimate of the undercoverage in the cell}}{\text{Weighted count of housing units in the cell}}$$

The computed second-stage ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell. For the Baltimore, Md.; Buffalo, N.Y.; Chicago, Ill.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; and Seattle-Everett, Wash.; SMSA's, a second-stage ratio estimation procedure was employed to adjust the central city/balance distribution of the weighted sample estimate of new construction housing units built since the last survey in permit-issuing areas to an independently derived estimate of this distribution.

This ratio estimation factor was calculated separately for the central city and balance of each SMSA and was applied to all new construction housing units from permit-issuing areas within the corresponding sector (central city or balance of the SMSA). This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}{\text{(Sample estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}$$

The independent estimates of new construction were based upon the number of authorized building permits which were determined from the Survey of Construction (SOC). The sample estimates were obtained from the weighted estimate of the AHS-SMSA sample housing units after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

This second-stage ratio estimation procedure was not implemented in the Honolulu, Hawaii; Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's. In the Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C., SMSA's, the central city permit offices, sources of SOC information on permits, had expanded their coverage of permits to include areas greater than those covered by the 1970 census central city definitions. Since the AHS-SMSA survey uses the 1970 definitions, use of this ratio estimation procedure would not correctly adjust the central city/balance distribution of new construction housing units. In the Honolulu, Hawaii, SMSA, this procedure could not be used because the permit office covers the entire SMSA and does not differentiate between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Neb.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and as a third-stage ratio estimation procedure for the remaining 11 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the October 1979 housing inventory of each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sectors. This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the October 1979 housing inventory for the corresponding sector of the SMSA)}}{\text{(AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA)}}$$

The numerator of this ratio was derived using 1970 and 1980 census counts. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight after the first-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and the existing weight after the second-stage ratio estimation procedure for the 11 remaining SMSA's.

The computed ratio estimation factors for the central city and balance of the SMSA's were then applied to the existing weight for all corresponding sample units and the resulting product was used as the final weight for tabulation purposes.

The effect of the total housing unit ratio estimation procedure, as well as the overall 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 SMSA as a whole, it can be expected that the sample estimates will be improved when the sample housing population, or different portions of it, are brought into agreement with known good estimates of the SMSA housing population.

1975-1979 lost housing units—The 1975-1979 lost housing unit (housing unit removed from the inventory) estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1975 housing inventory, as was described in the 1975 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas*. Since the 1975-1979 lost housing units existed, by definition, in the 1975 housing inventory, there was a 1975 housing inventory weight associated with each 1975-1979 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1975-1979 lost housing units.

1975 estimation procedure—This report presents data on the housing characteristics of the 1975 housing inventory from the 1975 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1975.

Ratio estimation procedure of the 1970 Census of Population and Housing—This report presents data on the housing characteristics of the 1970 housing inventory from the 1970 Census of Population and Housing. The statistics based on 1970 census sample data employed a ratio estimation procedure which was applied separately for each of the three census samples. A detailed description of this ratio estimation procedure can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

RELIABILITY OF THE ESTIMATES

There are two types of possible errors associated with estimates based on data from sample surveys—sampling and nonsampling errors. The following is a description of the sampling and nonsampling errors associated with the AHS-SMSA sample and of the nonsampling errors associated with the 1970 census estimates. A description of the sampling errors associated with the sample estimates from the 1970 census can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

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; and other errors of collection, response, processing, 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 1970 Census of Population and Housing and the 1975 AHS-SMSA sample.

1970 census—A number of studies were conducted to measure two types of general errors associated with 1970 census estimates—"coverage" and "content" errors. The "coverage" errors determined how completely housing units were counted in the census and the extent to which occupancy status was erroneously reported. The "content" errors measured the accuracy of the data collected for enumerated housing units. These errors were measured by reinterviews, record checks, and other surveys.

The detailed results of these studies, as well as the methodology employed, are presented in the 1970 Census of Population and Housing Evaluation and Research Program Reports, Series PHC(E)-5, *The Coverage of Housing in the 1970 Census,* and PHC(E)-10, *Accuracy of Data for Selected Housing Characteristics as Measured by Reinterviews.*

AHS-SMSA—Results from the 1979 AHS-SMSA sample reinterview program were not available at the time this report was being prepared. However, a study was conducted for the 1975 AHS-SMSA sample. The results are presented in the Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1975."

Coverage errors—In errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional (non-mobile home or trailer) new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units. Due to 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, in which case, 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, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit.

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 estimated that the 1975 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 for the 1979 survey, the number of missed housing units may be considerably less for 1979.

The final ratio estimation procedure corrects for these deficiencies as far as the count of total housing is concerned; i.e., it adjusts the estimate of the total housing inventory to the best available estimate. However, biases of subtotals would still remain.

Rounding errors—For errors associated with processing, the rounding of estimates introduces another source of error in the data, the severity of which depends on 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-SMSA sample—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 each other. 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 due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by 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, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. 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.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two 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 (page App-51) are approximations to the standard errors of various estimates shown in this report for this SMSA. 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. Standard errors applicable to estimates of characteristics of the 1975 housing inventory can be found in the AHS Series H-170 reports for 1975.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1979 housing inventory as well as estimates of characteristics of the 1975-1979 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 57,150 for the total SMSA, 30,200 for the central city of the SMSA, and 48,370 for the balance of the SMSA.

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 (page App-51) presents the standard errors of estimated percentages for the 1979 housing inventory as well as estimated percentages of the 1975-1979 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

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 as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is 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:

$$(100)\left(\frac{x}{y}\right)\sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}$$

where: x = the numerator of the ratio
y = the denominator of the ratio
σ_x = the standard error of the numerator
σ_y = the standard error of the denominator

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1979 there were 1,448,400 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 11,830. 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
1,447,700	11,830
1,448,400	x
1,500,000	11,770

The entry for "x" is determined as follows by vertically interpolating between 11,830 and 11,770.

$$1{,}448{,}400 - 1{,}447{,}700 = 700$$
$$1{,}500{,}000 - 1{,}447{,}700 = 52{,}300$$

$$11{,}830 + \frac{700}{52{,}300}(11{,}770 - 11{,}830) = 11{,}830$$

Consequently the 68-percent confidence interval, as shown by these data, is from 1,436,570 to 1,460,230 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1979 owner-occupied housing units lies within a range computed in this way, would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 1,429,470 to 1,467,330 housing units with 90 percent confidence; and that the average estimate lies within the interval from 1,424,740 to 1,472,060 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 1,448,400 owner-occupied housing units, 371,300, or 25.6 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 25.6 percent is approximately 0.5 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	25 or 75	25.6	50
1,400,000	0.5	a	0.6
1,448,400		p	
1,500,000	0.5	b	0.6

1. The entry for cell "a" is determined by horizontal interpolation between 0.5 and 0.6.

$$25.6 - 25.0 = 0.6$$
$$50.0 - 25.0 = 25.0$$

$$0.5 + \frac{0.6}{25.0}(0.6 - 0.5) = 0.5$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.5 and 0.6.

$$25.6 - 25.0 = 0.6$$
$$50.0 - 25.0 = 25.0$$

$$0.5 + \frac{0.6}{25.0}(0.6 - 0.5) = 0.5$$

3. The entry for "p" is determined to be 0.5 with vertical interpolation not required between identical values.

Consequently, the 68-percent confidence interval, as shown by these data, is from 25.1 to 26.1 percent; the 90-percent confidence interval is from 24.8 to 26.4 percent; and the 95-percent confidence interval is from 24.6 to 26.6 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. 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. Due to the overlap of the 1975 and 1979 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1975 and 1979 characteristics

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1979 there were 732,600 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 361,300. Table I shows the standard error of 732,600 is approximately 10,660, and the standard error of 371,300 is approximately 8,270. Therefore, the standard error of the estimated difference of 361,300 is about 13,490.

$$13{,}490 = \sqrt{(10{,}660)^2 + (8{,}270)^2}$$

Consequently, the 68-percent confidence interval for the 361,300 difference is from 347,810 to 374,790 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 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 339,720 to 382,880 housing units, and the 95-percent confidence interval is from 334,320 to 388,280 housing units. Thus, we can conclude with 95 percent confidence that the number of 1979 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied housing units with two bedrooms, since the 95-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 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 which interval of the distribution the lower percentage limit falls into. Similarly, to find the upper endpoint of the confidence interval, it is necessary to know which interval of the distribution the upper percentage limit falls into. These two distribution intervals could be different, although this will not happen very often.

For about 68 out of 100 possible samples, the average median from all possible samples would lie between these two values.

A two-standard-error confidence interval may be determined by finding the values corresponding to 50 percent plus and minus twice the standard error determined in step 1. For about 95 out of 100 possible samples, the average median from all possible samples would lie between these two values.

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 3.0. The base of the distribution from which this median was determined is 1,448,400 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 1,448,400 is approximately 0.6 percentage points.

÷2
724,200

2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.8 and 51.2.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with 3 persons (for purposes of calculating the median, the category of 3 persons is considered to be from 2.5 to 3.5 persons) contains the 48.8 percent derived in step 2. About 604,200 housing units or 41.7 percent fall below this interval, and 260,300 housing units or 18.0 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(48.8-41.7)}{18.0} = 2.9$$

Similarly, the interval for owner-occupied housing units with 3 persons contains the 51.2 percent derived in step 2. About 604,200 housing units or 41.7 percent fall below this interval, and 260,300 housing units or 18.0 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(51.2-41.7)}{18.0} = 3.0$$

Thus, the 95-percent confidence interval ranges from 2.9 to 3.0 percent.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1979 Housing Inventory and for Estimated Number of 1975-1979 Lost Units for the Chicago, Ill., SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size estimate	Standard error[1]		
	SMSA	In central city	Not in central city
0	220	210	210
100	220	210	210
200	220	210	210
500	330	320	320
700	390	380	380
1,000	460	450	460
2,500	730	720	720
5,000	1,040	1,010	1,020
10,000	1,470	1,430	1,440
25,000	2,310	2,240	2,270
50,000	3,250	3,140	3,180
75,000	3,960	3,800	3,860
100,000	4,550	4,330	4,420
150,000	5,520	5,180	5,310
200,000	6,310	5,840	6,010
250,000	6,980	6,360	6,590
300,000	7,570	6,770	7,060
400,000	8,550	7,360	7,790
500,000	9,340	7,680	8,290
600,000	9,990	7,770	8,580
700,000	10,520	7,630	8,710
800,000	10,950	7,250	8,660
900,000	11,290	6,580	8,450
1,000,000	11,550	5,540	8,050
1,100,000	11,740	3,820	7,440
1,175,800	11,830	—	6,810
1,200,000	11,850	—	6,560
1,300,000	11,890	—	5,270
1,400,000	11,870	—	3,110
1,447,700	11,830	—	—
1,500,000	11,770	—	—
1,750,000	11,210	—	—
2,000,000	10,130	—	—
2,225,000	8,310	—	—
2,500,000	5,040	—	—
2,623,500	—	—	—

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.3 for the total SMSA, 1.3 for the central city, and 1.2 for the balance (not in central city) estimates. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero since these estimates were derived from census data which are not subject to sampling error. However, these estimates are subject to the nonsampling errors associated with 1970 and 1980 censuses and with the interpolation procedures used to derive these estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1979 Housing Inventory and for Estimated Percentages of 1975-1979 Lost Housing Units for the Chicago, Ill., SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
500	30.1	30.1	30.1	30.1	30.1	32.8
700	23.6	23.6	23.6	23.6	24.0	27.8
1,000	17.7	17.7	17.7	17.7	20.1	23.2
2,500	7.9	7.9	7.9	8.8	12.7	14.7
5,000	4.1	4.1	4.5	6.2	9.0	10.4
10,000	2.1	2.1	3.2	4.4	6.4	7.3
25,000	0.9	0.9	2.0	2.8	4.0	4.6
50,000	0.4	0.7	1.4	2.0	2.8	3.3
75,000	0.3	0.5	1.2	1.6	2.3	2.7
100,000	0.2	0.5	1.0	1.4	2.0	2.3
150,000	0.14	0.4	0.8	1.1	1.6	1.9
200,000	0.11	0.3	0.7	1.0	1.4	1.6
250,000	0.19	0.3	0.6	0.9	1.3	1.5
300,000	0.17	0.3	0.6	0.8	1.2	1.3
400,000	0.15	0.2	0.5	0.7	1.0	1.2
500,000	0.04	0.2	0.5	0.6	0.9	1.0
600,000	0.04	0.2	0.4	0.6	0.8	0.9
700,000	0.03	0.2	0.4	0.5	0.8	0.9
800,000	0.03	0.2	0.4	0.5	0.7	0.8
900,000	0.02	0.15	0.3	0.5	0.7	0.8
1,000,000	0.02	0.15	0.3	0.4	0.6	0.7
1,100,000	0.02	0.14	0.3	0.4	0.6	0.7
1,200,000	0.02	0.13	0.3	0.4	0.6	0.7
1,300,000	0.02	0.13	0.3	0.4	0.6	0.6
1,400,000	0.02	0.12	0.3	0.4	0.5	0.6
1,500,000	0.01	0.12	0.3	0.4	0.5	0.6
1,750,000	0.01	0.11	0.2	0.3	0.5	0.6
2,000,000	0.01	0.10	0.2	0.3	0.4	0.5
2,250,000	0.01	0.10	0.2	0.3	0.4	0.5
2,500,000	0.01	0.09	0.2	0.3	0.4	0.5
3,000,000	0.01	0.08	0.2	0.3	0.4	0.4

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than 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 1.3 for the total SMSA, 1.3 for the central city, and 1.2 for the balance (not in central city).

Cleveland 1979

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-41
Annual Housing Survey App-41
Designation of sample housing units for the 1979 survey. App-41
Selection of the 1976 AHS-SMSA sample App-41
1976-1979 additions to the housing inventory App-43
Sample selection for the 1976 Coverage Improvement Program. App-43
Coverage improvement for deficiency 1 . . . . . App-43
Coverage improvement for deficiency 2 . . . . . App-43
Coverage improvement for deficiencies 3-6 . . . . . App-43
1970 Census of Population and Housing App-44
ESTIMATION. App-44
1979 housing inventory App-44
1976-1979 lost housing units . . . App-45
1976 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES. App-46
Nonsampling errors. App-46
1970 census App-46
AHS-SMSA App-46
Coverage errors App-46
Rounding errors. App-46
Sampling errors for the AHS-SMSA sample App-47
Illustration of the use of the standard error tables App-47
Differences. App-48
Illustration of the computation of the standard error of a difference App-48
Medians App-48
Illustration of the computation of the 95-percent confidence interval of a median App-49
Standard error tables App-50

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 15 SMSA's in this report series (H-170-79) are based on data collected from the 1979 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 15 SMSA's, the data were collected during the 12-month period from April 1979 through March 1980 with one-twelfth of the sample housing units being visited each month.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1979 were interviewed previously in either 1975 or 1976 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1979 group of SMSA's, the largest SMSA from three of the four census regions of the United States is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All the remaining 12 SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the 1979 group are: Chicago, Ill.; Houston, Tex.; and Seattle-Everett, Wash.

The remaining SMSA's in the 1979 group are: Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.

In this SMSA, 5,020 housing units were eligible for interview. Of these sample housing units, 259 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 5,020 housing units eligible for interview, 195 units were visited but were not eligible for interview because they were condemned, unfit, demolished, converted to group quarters use, etc.

Designation of sample housing units for the 1979 survey—The sample housing units designated to be interviewed in the 1979 survey consisted of the following categories which are described in detail in the succeeding sections.

1. All sample housing units that were interviewed in the 1976 survey.
2. 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 1976 survey. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1979 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1976 survey. (This sample represented the housing units built in permit-issuing areas, since the 1976 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1976 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1976 survey.)
5. All sample housing units that were selected as part of the 1976 Coverage Improvement Program. (This sample represented most of the housing units which, until 1976, did not have a chance of selection.)

Selection of the 1976 AHS-SMSA sample—In 1970, the sample for the SMSA's which were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in

the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (nonpermit universe). In 1979, the following four SMSA's were 100-percent permit-issuing: Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; and Miami, Fla. The remaining 11 SMSA's contain a sample from the nonpermit universe.

Sampling operations, described in the following paragraphs, were performed separately within the central city and the balance of the SMSA for each of the sample frames. The overall sampling rate used to select the sample for each SMSA was determined by the size of the sample. Thus, for the three largest SMSA's, the overall sampling rate differed by central city and the balance of the SMSA, since the sample was divided equally between the central city and the balance of the SMSA. The remaining SMSA's had an overall sampling rate about the same for the sample selected from both the central city and the balance of the SMSA, since the sample was distributed proportionately between the central city and the balance of the SMSA according to the distribution of the total housing units in each sector.

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA 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 housing unit records were stratified by race of head (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 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household 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, for this SMSA, 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 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. 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 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 SMSA. 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.

The second frame from which this SMSA 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 this SMSA. Prior to sample selection, the list of permits was chronologically 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.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS 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 (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

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 are included.

1976-1979 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1976 survey, was selected to represent housing units built in permit-issuing areas since the 1976 survey. Sampling procedures were identical to those used in selecting the 1970-1976 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1976, to identify any housing units missed in the 1976 survey or any housing units added since the 1976 survey.

Sample selection for the 1976 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

1. New construction housing units from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes and trailers placed in parks either missed during the 1970 census or established since the 1970 census.
3. Housing units missed in the 1970 census.
4. Housing units converted to residential use that were nonresidential at the time of the 1970 census.
5. Housing units that have been moved onto their present site since the 1970 census.
6. Mobile homes and trailers placed outside parks since the 1970 census or vacant at the time of the 1970 census.

For each of the five SMSA's interviewed for the first time in 1975, a full implementation of the Coverage Improvement Program was conducted as a part of the 1979 AHS. For each of the 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with some updating and refining as a part of the 1979 AHS. The following discussion of coverage improvement procedures applies to both the 1976 and 1979 programs. For the 10 (1976) SMSA's, estimates of housing units added by a specific procedure reflect units added in 1976 as well as any additions that resulted from the updating and refining in 1979.

Coverage improvement for deficiency 1—A sample of new construction housing units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 1979 SMSA's. Two different procedures were used. For the SMSA's first interviewed in 1976, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. These new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units).

The first stage sample selection was a sample of permit offices, and the second stage a sample of the 1969 permits within each of the selected permit offices. (In the Honolulu, Hawaii, SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second stage.) For the third stage, structures of size three or more were divided into clusters of an expected size of four housing units and a sample of clusters was selected.

For the SMSA's first interviewed in 1975, housing units whose permits were issued before January 1970, but which were completed after April 1970, were identified from the Survey of Construction (SOC), a survey of building permits conducted monthly by the Bureau of the Census. These housing units were then sampled at one-third the rate of regular AHS units. These procedures added an estimated 5,269 new construction housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiency 2—In permit-issuing areas, a sample of mobile homes and trailers placed in parks that were missed by the census or established after the census was selected in two stages. First, a sample of tracts was selected and canvassed. All parks were listed and then matched back to the 1970 census to identify parks missed by the census and parks established after the census. Second, the parks were divided into clusters of an expected size of four sites and a sample of clusters was selected and interviewed. Each of the sample housing units represented the same number of units that the regular AHS sample housing units represented. Since research showed that there were no mobile home or trailer parks in the central cities of the Chicago, Ill.; Hartford, Conn.; and Portland, Oreg.-Wash.; SMSA's, this procedure was only implemented outside the central cities for these SMSA's. Based on a cost-benefit analysis, this procedure was updated as part of the 1979 AHS only for the Denver, Colo. and Las Vegas, Nev., SMSA's. This procedure added no housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed housing units were sampled by one of two procedures. The first procedure was designed to represent units from the following types of missed structures (structures that had no chance of selection for the AHS):

1. Structures missed in the 1970 census.
2. Structures that were completely nonresidential in the 1970 census but now contain residential housing units.
3. Mobile homes and trailers that had been placed outside parks since the 1970 census and have a utility hookup, or were on the site during the present survey but not occupied on April 1, 1970, or had no utility hookup but were occupied by persons with no usual residence elsewhere.
4. Housing units that had been moved onto their present site since the 1970 census.

Initially, a subsample of regular AHS units was selected at a rate of 1 in 24 for the Chicago, Ill.; Hartford, Conn.; Miami, Fla.; Milwaukee, Wis.; and Portland, Oreg.-Wash.; SMSA's, and at a rate of 1 in 22.2152 for the 10 remaining SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing units were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and housing units within

these structures were interviewed. In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. Although the Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Houston, Tex.; Omaha, Nebr.-Iowa; and Raleigh, N.C.; SMSA's, were interviewed for the first time in 1976 and had these sampling procedures implemented at that time, some 1976 coverage improvement assignments were not sent out to be interviewed due to time limitations. The sampling was completed as part of the 1979 AHS and this procedure added an estimated 15,226 housing units to the coverage of the housing inventory of this SMSA.

The second procedure was designed to represent missed housing units from structures represented in the AHS. These missed housing units were:

1. Housing units missed in the 1970 census.
2. Nonresidential space converted to residential use since the 1970 census in structures that contained some residential housing units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structures selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 2,186 housing units to the coverage of the housing inventory for this SMSA.

1970 Census of Population and Housing—The estimates pertaining to the 1970 housing inventory (i.e., the housing inventory that existed at the time of the 1970 census) are based on either 20-, 15-, or 5-percent sample data collected in April 1970 for the Decennial Census of Population and Housing. A detailed description of the sample design employed for the 1970 census can be obtained in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

ESTIMATION

The 1979 AHS sample produced two types of estimates for each SMSA: Estimates pertaining to characteristics of the housing inventory at the time of the interview (i.e., the 1979 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1976 (i.e., 1976-1979 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1979 housing inventory—The AHS estimates of characteristics of the 1979 housing inventory were produced using a 2-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and a 3-stage ratio estimation procedure for the remaining 11 SMSA's. Prior to 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 the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of one or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes and trailers from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample housing units from both the nonpermit universe and the coverage improvement universe (if units were not included in any of the previous cells).

The following first-stage 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 permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of housing units enumerated in areas under the jurisdiction of permit-issuing offices. The denominators of the ratios were obtained from weighted estimates of all the AHS sample housing units within the corresponding ratio estimation categories using the existing weight (i.e., the basic weight times the noninterview factor). The computed ratio estimation factor was then applied to the existing weight for each sample housing unit within the corresponding ratio estimation category.

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 permit-issuing universe. Prior to the AHS sample selection within each SMSA, 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 sample. Since the number of housing units deleted from the AHS universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the Houston, Tex., SMSA, a second ratio estimation procedure was employed for new construction housing units outside the Houston central city. Due to rapid growth in the

nonpermit-issuing universe the sampling methods did not adequately represent the new construction outside the Houston central city. Because of this deficiency, separate factors were computed for the following cells:

1. Conventional housing units (non-mobile home or trailer) outside the central city, which were built between the 1970 census and the 1976 survey.
2. Mobile homes and trailers outside the central city with a model year between the 1970 census and 1976 survey.
3. Conventional housing units outside the central city built between the 1976 and 1979 surveys.
4. Mobile homes and trailers outside the central city with a model year between the 1976 and 1979 surveys.

In a specific cell, the ratio estimation factor was equal to the following:

$$\frac{\text{Weighted count of housing units in the cell} + \text{Estimate of the undercoverage in the cell}}{\text{Weighted count of housing units in the cell}}$$

The computed second-stage ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell. For the Baltimore, Md.; Buffalo, N.Y.; Chicago, Ill.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; and Seattle-Everett, Wash.; SMSA's, a second-stage ratio estimation procedure was employed to adjust the central city/balance distribution of the weighted sample estimate of new construction housing units built since the last survey in permit-issuing areas to an independently derived estimate of this distribution.

This ratio estimation factor was calculated separately for the central city and balance of each SMSA and was applied to all new construction housing units from permit-issuing areas within the corresponding sector (central city or balance of the SMSA). This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}{\text{(Sample estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}$$

The independent estimates of new construction were based upon the number of authorized building permits which were determined from the Survey of Construction (SOC). The sample estimates were obtained from the weighted estimate of the AHS-SMSA sample housing units after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

This second-stage ratio estimation procedure was not implemented in the Honolulu, Hawaii; Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's. In the Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, the central city permit offices, sources of SOC information on permits, had expanded their coverage of permits to include areas greater than those covered by the 1970 census central city definitions. Since the AHS-SMSA survey uses the 1970 definitions, use of this ratio estimation procedure would not correctly adjust the central city/balance distribution of new construction housing units. In the Honolulu, Hawaii, SMSA, this procedure could not be used because the permit office covers the entire SMSA and does not differentiate between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Neb.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and as a third-stage ratio estimation procedure for the remaining 11 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the October 1979 housing inventory of each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sectors. This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the October 1979 housing inventory for the corresponding sector of the SMSA)}}{\text{(AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA)}}$$

The numerator of this ratio was derived using 1970 and 1980 census counts. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight after the first-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and the existing weight after the second-stage ratio estimation procedure for the 11 remaining SMSA's.

The computed ratio estimation factors for the central city and balance of the SMSA's were then applied to the existing weight for all corresponding sample units and the resulting product was used as the final weight for tabulation purposes.

The effect of the total housing unit ratio estimation procedure, as well as the overall 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 SMSA as a whole, it can be expected that the sample estimates will be improved when the sample housing population, or different portions of it, are brought into agreement with known good estimates of the SMSA housing population.

1976-1979 lost housing units—The 1976-1979 lost housing unit (housing unit removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1976 housing inventory, as was described in the 1976 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1976-1979 lost housing units existed, by definition, in the 1976 housing inventory, there was a 1976 housing inventory weight associated with each 1976-1979 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1976-1979 lost housing units.

1976 estimation procedure—This report presents data on the housing characteristics of the 1976 housing inventory from the 1976 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1976.

Ratio estimation procedure of the 1970 Census of Population and Housing—This report presents data on the housing characteristics of the 1970 housing inventory from the 1970 Census of Population and Housing. The statistics based on 1970 census sample data employed a ratio estimation procedure which was applied separately for each of the three census samples. A detailed description of this ratio estimation procedure can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

RELIABILITY OF THE ESTIMATES

There are two types of possible errors associated with estimates based on data from sample surveys—sampling and nonsampling errors. The following is a description of the sampling and nonsampling errors associated with the AHS-SMSA sample and of the nonsampling errors associated with the 1970 census estimates. A description of the sampling errors associated with the sample estimates from the 1970 census can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

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; and other errors of collection, response, processing, 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 1970 Census of Population and Housing and the 1976 AHS-SMSA sample.

1970 census—A number of studies were conducted to measure two types of general errors associated with 1970 census estimates—"coverage" and "content" errors. The "coverage" errors determined how completely housing units were counted in the census and the extent to which occupancy status was erroneously reported. The "content" errors measured the accuracy of the data collected for enumerated housing units. These errors were measured by reinterviews, record checks, and other surveys.

The detailed results of these studies, as well as the methodology employed, are presented in the 1970 Census of Population and Housing Evaluation and Research Program Reports, Series PHC(E)-5, *The Coverage of Housing in the 1970 Census,* and PHC(E)-10, *Accuracy of Data for Selected Housing Characteristics as Measured by Reinterviews.*

AHS-SMSA—Results from the 1979 AHS-SMSA sample reinterview program were not available at the time this report was being prepared. However, a study was conducted for the 1976 AHS-SMSA sample. The results are presented in the Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1976."

Coverage errors—In errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional (non-mobile home or trailer) new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units. Due to 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, in which case, 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, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit.

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 estimated that the 1976 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 for the 1979 survey, the number of missed housing units may be considerably less for 1979.

The final ratio estimation procedure corrects for these deficiencies as far as the count of total housing is concerned; i.e., it adjusts the estimate of the total housing inventory to the best available estimate. However, biases of subtotals would still remain.

Rounding errors—For errors associated with processing, the rounding of estimates introduces another source of error in the data, the severity of which depends on 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-SMSA sample—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 each other. 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 due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by 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, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. 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.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two 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 (page App-50) are approximations to the standard errors of various estimates shown in this report for this SMSA. 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. Standard errors applicable to estimates of characteristics of the 1976 housing inventory can be found in the AHS Series H-170 reports for 1976.

Table I (page App-50) presents the standard errors applicable to estimates of characteristics of the 1979 housing inventory as well as estimates of characteristics of the 1976-1979 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 20,350 for the total SMSA, 12,120 for the central city of the SMSA, and 16,300 for the balance of the SMSA.

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 (page App-50) presents the standard errors of estimated percentages for the 1979 housing inventory as well as estimated percentages of the 1976-1979 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

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 as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is 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:

$$(100)\left(\frac{x}{y}\right)\sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}$$

where: x = the numerator of the ratio
y = the denominator of the ratio
σ_x = the standard error of the numerator
σ_y = the standard error of the denominator

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1979 there were 442,600 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 5,364. 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
400,000	5,510
442,600	x
490,500	5,200

The entry for "x" is determined as follows by vertically interpolating between 5,510 and 5,150.

$$442{,}600-400{,}000 = 42{,}600$$
$$490{,}500-400{,}000 = 90{,}500$$

$$5{,}510 + \frac{42{,}600}{90{,}500}(5{,}200-5{,}510) = 5{,}360$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 437,240 to 447,960 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1979 owner-occupied housing units lies within a range computed in this way, would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 434,024 to 451,176 housing units with 90 percent confidence; and that the average estimate lies within the interval from 431,880 to 453,320 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 442,600 owner-occupied housing units, 95,400, or 21.6 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 21.6 percent is approximately 0.8 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	21.6	25 or 75
400,000.	0.6	a	0.9
442,600.		p	
500,000.	0.6	b	0.8

1. The entry for cell "a" is determined by horizontal interpolation between 0.6 and 0.9.

$$21.6-10.0 = 11.6$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{11.6}{15.0}(0.9-0.6) = 0.8$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.6 and 0.8.

$$21.6-10.0 = 11.6$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{11.6}{15.0}(0.8-0.6) = 0.8$$

3. The entry for "p" is then determined by vertical interpolation between 0.8 and 0.8. This interpolation is not necessary as "a" and "b" have the same value. The formula, however would appear as follows.

$$442{,}600-400{,}000 = 42{,}600$$
$$500{,}000-400{,}000 = 100{,}000$$

$$0.8 + \frac{42{,}600}{100{,}000}(0.8-0.8) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 20.8 to 22.4 percent; the 90-percent confidence interval is from 20.3 to 22.9 percent; and the 95-percent confidence interval is from 20.0 to 23.2 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. 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. Due to the overlap of the 1976 and 1979 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1979 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1979 there were 236,800 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing with three bedrooms is 141,400. Table I shows the standard error of 236,800 is approximately 5,170, and the standard error of 95,400 is approximately 3,720. Therefore, the standard error of the estimated difference of 141,400 is about 6,370.

$$6{,}370 = \sqrt{(5{,}170)^2 + (3{,}720)^2}$$

Consequently, the 68-percent confidence interval for the 141,400 difference is from 135,030 to 147,770 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 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 131,208 to 151,592 housing units, and the 95-percent confidence interval is from 128,660 to 154,140 housing units. Thus, we can conclude with 95 percent confidence that the number of 1979 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied housing units with two bedrooms, since the 95-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 deter-

mine an interval about the estimated median so that there is a stated degree of confidence 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. These two distribution intervals could be different.

For about 68 out of 100 possible samples, the average median from all possible samples would lie between these two values.

A two-standard-error confidence interval may be determined by finding the values corresponding to 50 percent plus and minus twice the standard error determined in step 1. For about 95 out of 100 possible samples, the average median from all possible samples would lie between these two values.

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.8. The base of the distribution from which this median was determined is 442,600 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 442,600 is approximately 1.0 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.0 and 52.0.
3. From the distribution for "persons" in table A-1 of Part A, the interval for owner-occupied housing units with 3 persons (for purposes of calculating the median, the category of 3 persons is considered to be from 2.5 to 3.5 persons) contains the 48.0 percent derived in step 2. About 192,100 housing units or 43.4 percent fall below this interval, and 86,400 housing units or 19.5 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(48.0-43.4)}{19.5} = 2.7$$

Similarly, the interval for owner-occupied housing units with 3 persons contains the 52.0 percent derived in step 2. About 192,100 housing units or 43.4 percent fall below this interval, and 86,400 housing units or 19.5 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(52.0-43.4)}{19.5} = 2.9$$

Thus, the 95-percent confidence interval ranges from 2.7 to 2.9 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1979 Housing Inventory and for Estimated Number of 1976-1979 Lost Units for the Cleveland, Ohio, SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city
0	170	150	180
100	170	150	180
200	180	170	190
500	290	270	300
700	340	320	350
1,000	410	380	420
2,500	650	610	670
5,000	910	850	940
10,000	1,290	1,190	1,330
25,000	2,010	1,820	2,060
50,000	2,790	2,420	2,840
75,000	3,360	2,770	3,380
100,000	3,800	2,940	3,780
150,000	4,470	2,890	4,320
200,000	4,940	2,240	4,610
240,800	5,200	—	4,690
250,000	5,250	—	4,690
300,000	5,450	—	4,570
400,000	5,510	—	3,640
490,500	5,200	—	—
500,000	5,150	—	—
600,000	4,250	—	—
700,000	2,240	—	—
731,300	—	—	—

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2 for the total SMSA, 1.4 for the central city, and 1.1 for the balance (not in central city) estimates. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero since these estimates were derived from census data which are not subject to sampling error. However, these estimates are subject to the nonsampling errors associated with the 1970 and 1980 censuses and with the interpolation procedures used to derive these estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1979 Housing Inventory and for Estimated Percentages of 1976-1979 Lost Housing Units for the Cleveland, Ohio, SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
200	47.3	47.3	47.3	47.3	47.3	47.4
500	26.4	26.4	26.4	26.4	26.4	30.0
700	20.4	20.4	20.4	20.4	21.9	25.3
1,000	15.2	15.2	15.2	15.2	18.4	21.2
2,500	6.7	6.7	6.7	8.0	11.6	13.4
5,000	3.5	3.5	4.1	5.7	8.2	9.5
10,000	1.8	1.8	2.9	4.0	5.8	6.7
25,000	0.7	0.8	1.8	2.5	3.7	4.2
50,000	0.4	0.6	1.3	1.8	2.6	3.0
75,000	0.2	0.5	1.1	1.5	2.1	2.4
100,000	0.2	0.4	0.9	1.3	1.8	2.1
150,000	0.12	0.3	0.8	1.0	1.5	1.7
200,000	0.09	0.3	0.7	0.9	1.3	1.5
250,000	0.07	0.3	0.6	0.8	1.2	1.3
300,000	0.06	0.2	0.5	0.7	1.1	1.2
400,000	0.04	0.2	0.5	0.6	0.9	1.1
500,000	0.04	0.2	0.4	0.6	0.8	0.9
600,000	0.03	0.2	0.4	0.5	0.7	0.9
700,000	0.03	0.2	0.3	0.5	0.7	0.8
800,000	0.02	0.15	0.3	0.4	0.6	0.7

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than 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 1.2 for the total SMSA, 1.4 for the central city, and 1.1 for the balance (not in central city).

Denver 1979

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-41
Annual Housing Survey App-41
Designation of sample housing units for the 1979 survey. App-41
Selection of the 1976 AHS-SMSA sample App-41
1976-1979 additions to the housing inventory App-43
Sample selection for the 1976 Coverage Improvement Program. App-43
Coverage improvement for deficiency 1 . . . . . App-43
Coverage improvement for deficiency 2 . . . . . App-43
Coverage improvement for deficiencies 3-6 . . . . . App-43
1970 Census of Population and Housing. App-44
ESTIMATION. App-44
1979 housing inventory App-44
1976-1979 lost housing units . . . App-45
1976 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES. App-46
Nonsampling errors. App-46
1970 census App-46
AHS-SMSA App-46
Coverage errors App-46
Rounding errors. App-46
Sampling errors for the AHS-SMSA sample App-47
Illustration of the use of the standard error tables App-47
Differences. App-48
Illustration of the computation of the standard error of a difference App-48
Medians App-48
Illustration of the computation of the 95-percent confidence interval of a median App-49
Standard error tables App-50

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 15 SMSA's in this report series (H-170-79) are based on data collected from the 1979 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 15 SMSA's, the data were collected during the 12-month period from April 1979 through March 1980 with one-twelfth of the sample housing units being visited each month.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1979 were interviewed previously in either 1975 or 1976 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1979 group of SMSA's, the largest SMSA from three of the four census regions of the United States is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All the remaining 12 SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the 1979 group are: Chicago, Ill.; Houston, Tex.; and Seattle-Everett, Wash.

The remaining SMSA's in the 1979 group are: Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.

In this SMSA, 5,483 housing units were eligible for interview. Of these sample housing units, 271 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 5,483 housing units eligible for interview, 408 units were visited but were not eligible for interview because they were condemned, unfit, demolished, converted to group quarters use, etc.

Designation of sample housing units for the 1979 survey—The sample housing units designated to be interviewed in the 1979 survey consisted of the following categories which are described in detail in the succeeding sections.

1. All sample housing units that were interviewed in the 1976 survey.
2. 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 1976 survey. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1979 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1976 survey. (This sample represented the housing units built in permit-issuing areas, since the 1976 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1976 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1976 survey.)
5. All sample housing units that were selected as part of the 1976 Coverage Improvement Program. (This sample represented most of the housing units which, until 1976, did not have a chance of selection.)

Selection of the 1976 AHS-SMSA sample—In 1970, the sample for the SMSA's which were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in

the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). In 1979, the following four SMSA's were 100-percent permit-issuing: Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; and Miami, Fla. The remaining 11 SMSA's contain a sample from the nonpermit universe.

Sampling operations, described in the following paragraphs, were performed separately within the central city and the balance of the SMSA for each of the sample frames. The overall sampling rate used to select the sample for each SMSA was determined by the size of the sample. Thus, for the three largest SMSA's, the overall sampling rate differed by central city and the balance of the SMSA, since the sample was divided equally between the central city and the balance of the SMSA. The remaining SMSA's had an overall sampling rate about the same for the sample selected from both the central city and the balance of the SMSA, since the sample was distributed proportionately between the central city and the balance of the SMSA according to the distribution of the total housing units in each sector.

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA 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 housing unit records were stratified by race of head (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 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household 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, for this SMSA, 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 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. 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 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 SMSA. 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.

The second frame from which this SMSA 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 this SMSA. Prior to sample selection, the list of permits was chronologically 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.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS 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 (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

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 are included.

1976-1979 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1976 survey, was selected to represent housing units built in permit-issuing areas since the 1976 survey. Sampling procedures were identical to those used in selecting the 1970-1976 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1976, to identify any housing units missed in the 1976 survey or any housing units added since the 1976 survey.

Sample selection for the 1976 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

1. New construction housing units from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes and trailers placed in parks either missed during the 1970 census or established since the 1970 census.
3. Housing units missed in the 1970 census.
4. Housing units converted to residential use that were nonresidential at the time of the 1970 census.
5. Housing units that have been moved onto their present site since the 1970 census.
6. Mobile homes and trailers placed outside parks since the 1970 census or vacant at the time of the 1970 census.

For each of the five SMSA's interviewed for the first time in 1975, a full implementation of the Coverage Improvement Program was conducted as a part of the 1979 AHS. For each of the 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with some updating and refining as a part of the 1979 AHS. The following discussion of coverage improvement procedures applies to both the 1976 and 1979 programs. For the 10 (1976) SMSA's, estimates of housing units added by a specific procedure reflect units added in 1976 as well as any additions that resulted from the updating and refining in 1979.

Coverage improvement for deficiency 1—A sample of new construction housing units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 1979 SMSA's. Two different procedures were used. For the SMSA's first interviewed in 1976, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. These new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units).

The first stage sample selection was a sample of permit offices, and the second stage a sample of the 1969 permits within each of the selected permit offices. (In the Honolulu, Hawaii, SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second stage.) For the third stage, structures of size three or more were divided into clusters of an expected size of four housing units and a sample of clusters was selected.

For the SMSA's first interviewed in 1975, housing units whose permits were issued before January 1970, but which were completed after April 1970, were identified from the Survey of Construction (SOC), a survey of building permits conducted monthly by the Bureau of the Census. These housing units were then sampled at one-third the rate of regular AHS units. These procedures added an estimated 7,295 new construction housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiency 2—In permit-issuing areas, a sample of mobile homes and trailers placed in parks that were missed by the census or established after the census was selected in two stages. First, a sample of tracts was selected and canvassed. All parks were listed and then matched back to the 1970 census to identify parks missed by the census and parks established after the census. Second, the parks were divided into clusters of an expected size of four sites and a sample of clusters was selected and interviewed. Each of the sample housing units represented the same number of units that the regular AHS sample housing units represented. Since research showed that there were no mobile home or trailer parks in the central cities of the Chicago, Ill.; Hartford, Conn.; and Portland, Oreg.-Wash.; SMSA's, this procedure was only implemented outside the central cities for these SMSA's. Based on a cost-benefit analysis, this procedure was updated as part of the 1979 AHS only for the Denver, Colo. and Las Vegas, Nev., SMSA's. This procedure added 7,893 housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed housing units were sampled by one of two procedures. The first procedure was designed to represent units from the following types of missed structures (structures that had no chance of selection for the AHS):

1. Structures missed in the 1970 census.
2. Structures that were completely nonresidential in the 1970 census but now contain residential housing units.
3. Mobile homes and trailers that had been placed outside parks since the 1970 census and have a utility hookup, or were on the site during the present survey but not occupied on April 1, 1970, or had no utility hookup but were occupied by persons with no usual residence elsewhere.
4. Housing units that had been moved onto their present site since the 1970 census.

Initially, a subsample of regular AHS units was selected at a rate of 1 in 24 for the Chicago, Ill.; Hartford, Conn.; Miami, Fla.; Milwaukee, Wis.; and Portland, Oreg.-Wash.; SMSA's, and at a rate of 1 in 22.2152 for the 10 remaining SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing units were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and housing units within

these structures were interviewed. In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. Although the Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Houston, Tex.; Omaha, Nebr.-Iowa; and Raleigh, N.C.; SMSA's, were interviewed for the first time in 1976 and had these sampling procedures implemented at that time, some 1976 coverage improvement assignments were not sent out to be interviewed due to time limitations. The sampling was completed as part of the 1979 AHS and this procedure added an estimated 5,280 housing units to the coverage of the housing inventory of this SMSA.

The second procedure was designed to represent missed housing units from structures represented in the AHS. These missed housing units were:

1. Housing units missed in the 1970 census.
2. Nonresidential space converted to residential use since the 1970 census in structures that contained some residential housing units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structures selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added no housing units to the coverage of the housing inventory for this SMSA.

1970 Census of Population and Housing—The estimates pertaining to the 1970 housing inventory (i.e., the housing inventory that existed at the time of the 1970 census) are based on either 20-, 15-, or 5-percent sample data collected in April 1970 for the Decennial Census of Population and Housing. A detailed description of the sample design employed for the 1970 census can be obtained in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

ESTIMATION

The 1979 AHS sample produced two types of estimates for each SMSA: Estimates pertaining to characteristics of the housing inventory at the time of the interview (i.e., the 1979 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1976 (i.e., 1976-1979 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1979 housing inventory—The AHS estimates of characteristics of the 1979 housing inventory were produced using a 2-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and a 3-stage ratio estimation procedure for the remaining 11 SMSA's. Prior to 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 the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of one or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes and trailers from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample housing units from both the nonpermit universe and the coverage improvement universe (if units were not included in any of the previous cells).

The following first-stage 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 permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of housing units enumerated in areas under the jurisdiction of permit-issuing offices. The denominators of the ratios were obtained from weighted estimates of all the AHS sample housing units within the corresponding ratio estimation categories using the existing weight (i.e., the basic weight times the noninterview factor). The computed ratio estimation factor was then applied to the existing weight for each sample housing unit within the corresponding ratio estimation category.

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 permit-issuing universe. Prior to the AHS sample selection within each SMSA, 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 sample. Since the number of housing units deleted from the AHS universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the Houston, Tex., SMSA, a second ratio estimation procedure was employed for new construction housing units outside the Houston central city. Due to rapid growth in the

nonpermit-issuing universe the sampling methods did not adequately represent the new construction outside the Houston central city. Because of this deficiency, separate factors were computed for the following cells:

1. Conventional housing units (non-mobile home or trailer) outside the central city, which were built between the 1970 census and the 1976 survey.
2. Mobile homes and trailers outside the central city with a model year between the 1970 census and 1976 survey.
3. Conventional housing units outside the central city built between the 1976 and 1979 surveys.
4. Mobile homes and trailers outside the central city with a model year between the 1976 and 1979 surveys.

In a specific cell, the ratio estimation factor was equal to the following:

$$\frac{\text{Weighted count of housing units in the cell} + \text{Estimate of the undercoverage in the cell}}{\text{Weighted count of housing units in the cell}}$$

The computed second-stage ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell. For the Baltimore, Md.; Buffalo, N.Y.; Chicago, Ill.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; and Seattle-Everett, Wash.; SMSA's, a second-stage ratio estimation procedure was employed to adjust the central city/balance distribution of the weighted sample estimate of new construction housing units built since the last survey in permit-issuing areas to an independently derived estimate of this distribution.

This ratio estimation factor was calculated separately for the central city and balance of each SMSA and was applied to all new construction housing units from permit-issuing areas within the corresponding sector (central city or balance of the SMSA). This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}{\text{(Sample estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}$$

The independent estimates of new construction were based upon the number of authorized building permits which were determined from the Survey of Construction (SOC). The sample estimates were obtained from the weighted estimate of the AHS-SMSA sample housing units after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

This second-stage ratio estimation procedure was not implemented in the Honolulu, Hawaii; Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's. In the Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C., SMSA's, the central city permit offices, sources of SOC information on permits, had expanded their coverage of permits to include areas greater than those covered by the 1970 census central city definitions. Since the AHS-SMSA survey uses the 1970 definitions, use of this ratio estimation procedure would not correctly adjust the central city/balance distribution of new construction housing units. In the Honolulu, Hawaii, SMSA, this procedure could not be used because the permit office covers the entire SMSA and does not differentiate between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Neb.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and as a third-stage ratio estimation procedure for the remaining 11 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the October 1979 housing inventory of each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sectors. This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the October 1979 housing inventory for the corresponding sector of the SMSA)}}{\text{(AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA)}}$$

The numerator of this ratio was derived using 1970 and 1980 census counts. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight after the first-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and the existing weight after the second-stage ratio estimation procedure for the 11 remaining SMSA's.

The computed ratio estimation factors for the central city and balance of the SMSA's were then applied to the existing weight for all corresponding sample units and the resulting product was used as the final weight for tabulation purposes.

The effect of the total housing unit ratio estimation procedure, as well as the overall 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 SMSA as a whole, it can be expected that the sample estimates will be improved when the sample housing population, or different portions of it, are brought into agreement with known good estimates of the SMSA housing population.

1976-1979 lost housing units—The 1976-1979 lost housing unit (housing unit removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1976 housing inventory, as was described in the 1976 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1976-1979 lost housing units existed, by definition, in the 1976 housing inventory, there was a 1976 housing inventory weight associated with each 1976-1979 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1976-1979 lost housing units.

1976 estimation procedure—This report presents data on the housing characteristics of the 1976 housing inventory from the 1976 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS series H-170 reports for 1976.

Ratio estimation procedure of the 1970 Census of Population and Housing—This report presents data on the housing characteristics of the 1970 housing inventory from the 1970 Census of Population and Housing. The statistics based on 1970 census sample data employed a ratio estimation procedure which was applied separately for each of the three census samples. A detailed description of this ratio estimation procedure can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

RELIABILITY OF THE ESTIMATES

There are two types of possible errors associated with estimates based on data from sample surveys—sampling and nonsampling errors. The following is a description of the sampling and nonsampling errors associated with the AHS-SMSA sample and of the nonsampling errors associated with the 1970 census estimates. A description of the sampling errors associated with the sample estimates from the 1970 census can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

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; and other errors of collection, response, processing, 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 1970 Census of Population and Housing and the 1975 AHS-SMSA sample.

1970 census—A number of studies were conducted to measure two types of general errors associated with 1970 census estimates—"coverage" and "content" errors. The "coverage" errors determined how completely housing units were counted in the census and the extent to which occupancy status was erroneously reported. The "content" errors measured the accuracy of the data collected for enumerated housing units. These errors were measured by reinterviews, record checks, and other surveys.

The detailed results of these studies, as well as the methodology employed, are presented in the 1970 Census of Population and Housing Evaluation and Research Program Reports, Series PHC(E)-5, *The Coverage of Housing in the 1970 Census,* and PHC(E)-10, *Accuracy of Data for Selected Housing Characteristics as Measured by Reinterviews.*

AHS-SMSA—Results from the 1979 AHS-SMSA sample reinterview program were not available at the time this report was being prepared. However, a study was conducted for the 1976 AHS-SMSA sample. The results are presented in the Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1976."

Coverage errors—In errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional (non-mobile home or trailer) new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units. Due to 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, in which case, 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, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit.

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 estimated that the 1976 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 for the 1979 survey, the number of missed housing units may be considerably less for 1979.

The final ratio estimation procedure corrects for these deficiencies as far as the count of total housing is concerned; i.e., it adjusts the estimate of the total housing inventory to the best available estimate. However, biases of subtotals would still remain.

Rounding errors—For errors associated with processing, the rounding of estimates introduces another source of error in the data, the severity of which depends on 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-SMSA sample—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 each other. 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 due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by 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, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. 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.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two 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 (page App-50) are approximations to the standard errors of various estimates shown in this report for this SMSA. 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. Standard errors applicable to estimates of characteristics of the 1976 housing inventory can be found in the AHS Series H-170 reports for 1976.

Table I (page App-50) presents the standard errors applicable to estimate of characteristics of the 1979 housing inventory as well as estimates of characteristics of the 1976-1979 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 19,770 for the total SMSA, 10,420 for the central city of the SMSA, and 16,720 for the balance of the SMSA.

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 (page App-50) presents the standard errors of estimated percentages for the 1979 housing inventory as well as estimated percentages of the 1976-1979 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

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 as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is 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:

$$(100)\left(\frac{x}{y}\right)\sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}$$

where: x = the numerator of the ratio
y = the denominator of the ratio
σ_x = the standard error of the numerator
σ_y = the standard error of the denominator

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1979 there were 370,100 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 4,480. 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
300,000	4,590
370,100	x
400,000	4,430

The entry for "x" is determined as follows by vertically interpolating between 4,590 and 4,430.

$$370{,}100-300{,}000 = 70{,}100$$
$$400{,}000-300{,}000 = 100{,}000$$

$$4{,}590 + \frac{70{,}100}{100{,}000}(4{,}430-4{,}590) = 4{,}480$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 365,620 to 374,580 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1979 owner-occupied housing units lies within a range computed in this way, would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 362,930 to 377,270 housing units with 90 percent confidence; and that the average estimate lies within the interval from 361,140 to 379,060 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 370,100 owner-occupied housing units, 91,200, or 24.6 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 24.6 percent is approximately 0.8 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	24.6	25 or 75
300,000.	0.6	a	0.9
370,100.		p	
400,000.	0.6	b	0.8

1. The entry for cell "a" is determined by horizontal interpolation between 0.6 and 0.9.

$$24.6-10.0 = 14.6$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{14.6}{15.0}(0.9-0.6) = 0.9$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.6 and 0.8.

$$24.6-10.0 = 14.6$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{14.6}{15.0}(0.8-0.6) = 0.8$$

3. The entry for "p" is then determined by vertical interpolation between 0.9 and 0.8.

$$370{,}100-300{,}000 = 70{,}100$$
$$400{,}000-300{,}000 = 100{,}000$$

$$0.9 + \frac{70{,}100}{100{,}000}(0.8-0.9) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 23.8 to 25.4 percent; the 90-percent confidence interval is from 23.3 to 25.9 percent; and the 95-percent confidence interval is from 23.0 to 26.2 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. 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. Due to the overlap of the 1976 and 1979 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1979 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1979 there were 152,100 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 60,900. Table I shows the standard error of 152,100 is approximately 3,930, and the standard error of 91,200 is approximately 3,220. Therefore, the standard error of the estimated difference of 60,900 is about 5,080.

$$5{,}080 = \sqrt{(3{,}930)^2 + (3{,}220)^2}$$

Consequently, the 68-percent confidence interval for the 60,900 difference is from 55,820 to 65,980 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 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 52,770 to 69,030 housing units, and the 95-percent confidence interval is from 50,740 to 71,060 housing units. Thus, we can conclude with 95 percent confidence that the number of 1979 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied housing units with two bedrooms, since the 95-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 deter-

mine an interval about the estimated median so that there is a stated degree of confidence 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. These two distribution intervals could be different.

For about 68 out of 100 possible samples, the average median from all possible samples would lie between these two values.

A two-standard-error confidence interval may be determined by finding the values corresponding to 50 percent plus and minus twice the standard error determined in step 1. For about 95 out of 100 possible samples, the average median from all possible samples would lie between these two values.

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.6. The base of the distribution from which this median was determined is 370,100 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 370,100 is approximately 1.0 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.0 and 52.0.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with 3 persons (for purposes of calculating the median, the category of 3 persons is considered to be from 2.5 to 3.5 persons) contains the 48.0 percent derived in step 2. About 174,900 housing units or 47.3 percent fall below this interval, and 71,300 housing units or 19.3 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(48.0-47.3)}{19.3} = 2.5$$

Similarly, the interval for owner-occupied housing units with 3 persons contains the 52.0 percent derived in step 2. About 174,900 housing units or 47.3 percent fall below this interval, and 71,300 housing units or 19.3 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(52.0-47.3)}{19.3} = 2.7$$

Thus, the 95-percent confidence interval ranges from 2.5 to 2.7 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1979 Housing Inventory and for Estimated Number of 1976-1979 Lost Units for the Denver, Colo., SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city
0	130	130	130
100	130	130	130
200	160	160	160
500	260	260	260
700	310	310	300
1,000	370	370	360
2,500	580	580	570
5,000	810	810	810
10,000	1,150	1,130	1,140
25,000	1,790	1,730	1,770
50,000	2,480	2,290	2,410
75,000	2,970	2,600	2,850
100,000	3,350	2,740	3,170
150,000	3,910	2,610	3,550
200,000	4,280	1,770	3,670
226,200	4,410	–	3,650
250,000	4,490	–	3,570
300,000	4,590	–	3,220
400,000	4,430	–	870
405,700	4,410	–	–
500,000	3,740	–	–
600,000	2,010	–	–
631,900	–	–	–

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2 for the total SMSA, 1.6 for the central city, and 1.1 for the balance (not in central city) estimates. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero since these estimates were derived from census data which are not subject to sampling error. However, these estimates are subject to the nonsampling errors associated with the 1970 and 1980 censuses and with the interpolation procedures used to derive these estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1979 Housing Inventory and for Estimated Percentages of 1976-1979 Lost Housing Units for the Denver, Colo., SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
200	40.3	40.3	40.3	40.3	40.3	41.0
500	21.2	21.2	21.2	21.2	22.5	26.0
700	16.1	16.1	16.1	16.1	19.0	21.9
1,000	11.9	11.9	11.9	11.9	15.9	18.4
2,500	5.1	5.1	5.1	7.0	10.1	11.6
5,000	2.6	2.6	3.6	4.9	7.1	8.2
10,000	1.3	1.3	2.5	3.5	5.0	5.8
25,000	0.5	0.7	1.6	2.2	3.2	3.7
50,000	0.3	0.5	1.1	1.6	2.2	2.6
75,000	0.2	0.4	0.9	1.3	1.8	2.1
100,000	0.13	0.4	0.8	1.1	1.6	1.8
150,000	0.09	0.3	0.7	0.9	1.3	1.5
200,000	0.07	0.3	0.6	0.8	1.1	1.3
250,000	0.05	0.2	0.5	0.7	1.0	1.2
300,000	0.04	0.2	0.5	0.6	0.9	1.1
400,000	0.03	0.2	0.4	0.6	0.8	0.9
500,000	0.03	0.2	0.4	0.5	0.7	0.8
600,000	0.02	0.15	0.3	0.4	0.6	0.7
700,000	0.02	0.14	0.3	0.4	0.6	0.7

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than 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 1.2 for the total SMSA, 1.6 for the central city, and 1.1 for the balance (not in central city).

Hartford 1979

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-41
Annual Housing Survey . . . App-41
Designation of sample housing units for the 1979 survey . . App-41
Selection of the 1975 AHS-SMSA sample . . . App-41
1975-1979 additions to the housing inventory . . . App-43
Sample selection for the 1979 Coverage Improvement Program. App-43
Coverage improvement for deficiency 1 . . . App-43
Coverage improvement for deficiency 2 . . . App-43
Coverage improvement for deficiencies 3-6 . . . App-43
1970 Census of Population and Housing . . . App-44
ESTIMATION . . . App-44
1979 housing inventory . . . App-44
1975-1979 lost housing units . . . App-45
1975 estimation procedure . . . App-46
Ratio estimation procedure of the 1970 Census of Population and Housing . . . App-46
RELIABILITY OF THE ESTIMATES . . . App-46
Nonsampling errors . . . App-46
1970 census . . . App-46
AHS-SMSA . . . App-46
Coverage errors . . . App-46
Rounding errors . . . App-46
Sampling errors for the AHS-SMSA sample . . . App-47
Illustration of the use of the standard error tables . . . App-47
Differences . . . App-48
Illustration of the computation of the standard error of a difference . . . App-48
Medians . . . App-48
Illustration of the computation of the 95-percent confidence interval of a median . . . App-49
Standard error tables . . . App-50

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 15 SMSA's in this report series (H-170-79) are based on data collected from the 1979 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 15 SMSA's, the data were collected during the 12-month period from April 1979 through March 1980 with one-twelfth of the sample housing units being visited each month.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1979 were interviewed previously in either 1975 or 1976 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1979 group of SMSA's, the largest SMSA from three of the four census regions of the United States is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All the remaining 12 SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the 1979 group are: Chicago, Ill.; Houston, Tex.; and Seattle-Everett, Wash.

The remaining SMSA's in the 1979 group are: Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.

In this SMSA, 5,211 housing units were eligible for interview. Of these sample housing units, 199 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 5,211 housing units eligible for interview, 283 units were visited but were not eligible for interview because they were condemned, unfit, demolished, converted to group quarters use, etc.

Designation of sample housing units for the 1979 survey—The sample housing units designated to be interviewed in the 1979 survey consisted of the following categories which are described in detail in the succeeding sections.

1. All sample housing units that were interviewed in the 1975 survey.
2. 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 1975 survey. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1979 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1975 survey. (This sample represented the housing units built in permit-issuing areas, since the 1975 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1975 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1975 survey.)
5. All sample housing units that were selected as part of the 1979 Coverage Improvement Program. (This sample represented most of the housing units which, until 1979, did not have a chance of selection.)

Selection of the 1975 AHS-SMSA sample—In 1970, the sample for the SMSA's which were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in

the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). In 1979, the following four SMSA's were 100-percent permit-issuing: Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; and Miami, Fla. The remaining 11 SMSA's contain a sample from the nonpermit universe.

Sampling operations, described in the following paragraphs, were performed separately within the central city and the balance of the SMSA for each of the sample frames. The overall sampling rate used to select the sample for each SMSA was determined by the size of the sample. Thus, for the three largest SMSA's, the overall sampling rate differed by central city and the balance of the SMSA, since the sample was divided equally between the central city and the balance of the SMSA. The remaining SMSA's had an overall sampling rate about the same for the sample selected from both the central city and the balance of the SMSA, since the sample was distributed proportionately between the central city and the balance of the SMSA according to the distribution of the total housing units in each sector.

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA 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 housing unit records were stratified by race of head (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 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household 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, for this SMSA, 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 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. 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 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 SMSA. 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.

The second frame from which this SMSA 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 this SMSA. Prior to sample selection, the list of permits was chronologically 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.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS 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 (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

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 are included.

1975-1979 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1975 survey, was selected to represent housing units built in permit-issuing areas since the 1975 survey. Sampling procedures were identical to those used in selecting the 1970-1975 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1975, to identify any housing units missed in the 1975 survey or any housing units added since the 1975 survey.

Sample selection for the 1979 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

1. New construction housing units from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes and trailers placed in parks either missed during the 1970 census or established since the 1970 census.
3. Housing units missed in the 1970 census.
4. Housing units converted to residential use that were nonresidential at the time of the 1970 census.
5. Housing units that have been moved onto their present site since the 1970 census.
6. Mobile homes and trailers placed outside parks since the 1970 census or vacant at the time of the 1970 census.

For each of the five SMSA's interviewed for the first time in 1975, a full implementation of the Coverage Improvement Program was conducted as a part of the 1979 AHS. For each of the 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with some updating and refining as a part of the 1979 AHS. The following discussion of coverage improvement procedures applies to both the 1976 and 1979 programs. For the 10 (1976) SMSA's, estimates of housing units added by a specific procedure reflect units added in 1976 as well as any additions that resulted from the updating and refining in 1979.

Coverage improvement for deficiency 1—A sample of new construction housing units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 1979 SMSA's. Two different procedures were used. For the SMSA's first interviewed in 1976, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. These new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units).

The first stage sample selection was a sample of permit offices, and the second stage a sample of the 1969 permits within each of the selected permit offices. (In the Honolulu, Hawaii, SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second stage.) For the third stage, structures of size three or more were divided into clusters of an expected size of four housing units and a sample of clusters was selected.

For the SMSA's first interviewed in 1975, housing units whose permits were issued before January 1970, but which were completed after April 1970, were identified from the Survey of Construction (SOC), a survey of building permits conducted monthly by the Bureau of the Census. These housing units were then sampled at one-third the rate of regular AHS units. These procedures added an estimated 3,282 new construction housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiency 2—In permit-issuing areas, a sample of mobile homes and trailers placed in parks that were missed by the census or established after the census was selected in two stages. First, a sample of tracts was selected and canvassed. All parks were listed and then matched back to the 1970 census to identify parks missed by the census and parks established after the census. Second, the parks were divided into clusters of an expected size of four sites and a sample of clusters was selected and interviewed. Each of the sample housing units represented the same number of units that the regular AHS sample housing units represented. Since research showed that there were no mobile home or trailer parks in the central cities of the Chicago, Ill.; Hartford, Conn.; and Portland, Oreg.-Wash.; SMSA's, this procedure was only implemented outside the central cities for these SMSA's. Based on a cost-benefit analysis, this procedure was updated as part of the 1979 AHS only for the Denver, Colo. and Las Vegas, Nev., SMSA's. This procedure added no units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed housing units were sampled by one of two procedures. The first procedure was designed to represent units from the following types of missed structures (structures that had no chance of selection for the AHS):

1. Structures missed in the 1970 census.
2. Structures that were completely nonresidential in the 1970 census but now contain residential housing units.
3. Mobile homes and trailers that had been placed outside parks since the 1970 census and have a utility hookup, or were on the site during the present survey but not occupied on April 1, 1970, or had no utility hookup but were occupied by persons with no usual residence elsewhere.
4. Housing units that had been moved onto their present site since the 1970 census.

Initially, a subsample of regular AHS units was selected at a rate of 1 in 24 for the Chicago, Ill.; Hartford, Conn.; Miami, Fla.; Milwaukee, Wis.; and Portland, Oreg.-Wash.; SMSA's, and at a rate of 1 in 22.2152 for the 10 remaining SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing units were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and housing units within these structures were interviewed. In cases where the interviewer workload would have been too great, a representative subsample

of units within these structures was selected. Although the Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Houston, Tex.; Omaha, Nebr.-Iowa; and Raleigh, N.C.; SMSA's, were interviewed for the first time in 1976 and had these sampling procedures implemented at that time, some 1976 coverage improvement assignments were not sent out to be interviewed due to time limitations. The sampling was completed as part of the 1979 AHS and this procedure added an estimated 4,276 housing units to the coverage of the housing inventory of this SMSA.

The second procedure was designed to represent missed housing units from structures represented in the AHS. These missed housing units were:

1. Housing units missed in the 1970 census.
2. Nonresidential space converted to residential use since the 1970 census in structures that contained some residential housing units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structures selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 981 housing units to the coverage of the housing inventory for this SMSA.

1970 Census of Population and Housing—The estimates pertaining to the 1970 housing inventory (i.e., the housing inventory that existed at the time of the 1970 census) are based on either 20-, 15-, or 5-percent sample data collected in April 1970 for the Decennial Census of Population and Housing. A detailed description of the sample design employed for the 1970 census can be obtained in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

ESTIMATION

The 1979 AHS sample produced two types of estimates for each SMSA: Estimates pertaining to characteristics of the housing inventory at the time of the interview (i.e., the 1979 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1975 (i.e., 1975-1979 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1979 housing inventory—The AHS estimates of characteristics of the 1979 housing inventory were produced using a 2-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and a 3-stage ratio estimation procedure for the remaining 11 SMSA's. Prior to 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 the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of one or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes and trailers from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample housing units from both the nonpermit universe and the coverage improvement universe (if units were not included in any of the previous cells).

The following first-stage 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 permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of housing units enumerated in areas under the jurisdiction of permit-issuing offices. The denominators of the ratios were obtained from weighted estimates of all the AHS sample housing units within the corresponding ratio estimation categories using the existing weight (i.e., the basic weight times the noninterview factor). The computed ratio estimation factor was then applied to the existing weight for each sample housing unit within the corresponding ratio estimation category.

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 permit-issuing universe. Prior to the AHS sample selection within each SMSA, 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 sample. Since the number of housing units deleted from the AHS universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the Houston, Tex., SMSA, a second ratio estimation procedure was employed for new construction housing units outside the Houston central city. Due to rapid growth in the

nonpermit-issuing universe the sampling methods did not adequately represent the new construction outside the Houston central city. Because of this deficiency, separate factors were computed for the following cells:

1. Conventional housing units (non-mobile home or trailer) outside the central city, which were built between the 1970 census and the 1976 survey.
2. Mobile homes and trailers outside the central city with a model year between the 1970 census and 1976 survey.
3. Conventional housing units outside the central city built between the 1976 and 1979 surveys.
4. Mobile homes and trailers outside the central city with a model year between the 1976 and 1979 surveys.

In a specific cell, the ratio estimation factor was equal to the following:

$$\frac{\text{Weighted count of housing units in the cell} + \text{Estimate of the undercoverage in the cell}}{\text{Weighted count of housing units in the cell}}$$

The computed second-stage ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell. For the Baltimore, Md.; Buffalo, N.Y.; Chicago, Ill.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; and Seattle-Everett, Wash.; SMSA's, a second-stage ratio estimation procedure was employed to adjust the central city/balance distribution of the weighted sample estimate of new construction housing units built since the last survey in permit-issuing areas to an independently derived estimate of this distribution.

This ratio estimation factor was calculated separately for the central city and balance of each SMSA and was applied to all new construction housing units from permit-issuing areas within the corresponding sector (central city or balance of the SMSA). This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}{\text{(Sample estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}$$

The independent estimates of new construction were based upon the number of authorized building permits which were determined from the Survey of Construction (SOC). The sample estimates were obtained from the weighted estimate of the AHS-SMSA sample housing units after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

This second-stage ratio estimation procedure was not implemented in the Honolulu, Hawaii; Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's. In the Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, the central city permit offices, sources of SOC information on permits, had expanded their coverage of permits to include areas greater than those covered by the 1970 census central city definitions. Since the AHS-SMSA survey uses the 1970 definitions, use of this ratio estimation procedure would not correctly adjust the central city/balance distribution of new construction housing units. In the Honolulu, Hawaii, SMSA, this procedure could not be used because the permit office covers the entire SMSA and does not differentiate between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Neb.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and as a third-stage ratio estimation procedure for the remaining 11 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the October 1979 housing inventory of each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sectors. This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the October 1979 housing inventory for the corresponding sector of the SMSA)}}{\text{(AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA)}}$$

The numerator of this ratio was derived using 1970 and 1980 census counts. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight after the first-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and the existing weight after the second-stage ratio estimation procedure for the 11 remaining SMSA's.

The computed ratio estimation factors for the central city and balance of the SMSA's were then applied to the existing weight for all corresponding sample units and the resulting product was used as the final weight for tabulation purposes.

The effect of the total housing unit ratio estimation procedure, as well as the overall 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 SMSA as a whole, it can be expected that the sample estimates will be improved when the sample housing population, or different portions of it, are brought into agreement with known good estimates of the SMSA housing population.

1975-1979 lost housing units—The 1975-1979 lost housing unit (housing unit removed from the inventory) estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1975 housing inventory, as was described in the 1975 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1975-1979 lost housing units existed, by definition, in the 1975 housing inventory, there was a 1975 housing inventory weight associated with each 1975-1979 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1975-1979 lost housing units.

1975 estimation procedure—This report presents data on the housing characteristics of the 1975 housing inventory from the 1975 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1975.

Ratio estimation procedure of the 1970 Census of Population and Housing—This report presents data on the housing characteristics of the 1970 housing inventory from the 1970 Census of Population and Housing. The statistics based on 1970 census sample data employed a ratio estimation procedure which was applied separately for each of the three census samples. A detailed description of this ratio estimation procedure can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

RELIABILITY OF THE ESTIMATES

There are two types of possible errors associated with estimates based on data from sample surveys—sampling and nonsampling errors. The following is a description of the sampling and nonsampling errors associated with the AHS-SMSA sample and of the nonsampling errors associated with the 1970 census estimates. A description of the sampling errors associated with the sample estimates from the 1970 census can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

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; and other errors of collection, response, processing, 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 1970 Census of Population and Housing and the 1975 AHS-SMSA sample.

1970 census—A number of studies were conducted to measure two types of general errors associated with 1970 census estimates—"coverage" and "content" errors. The "coverage" errors determined how completely housing units were counted in the census and the extent to which occupancy status was erroneously reported. The "content" errors measured the accuracy of the data collected for enumerated housing units. These errors were measured by reinterviews, record checks, and other surveys.

The detailed results of these studies, as well as the methodology employed, are presented in the 1970 Census of Population and Housing Evaluation and Research Program Reports, Series PHC(E)-5, *The Coverage of Housing in the 1970 Census,* and PHC(E)-10, *Accuracy of Data for Selected Housing Characteristics as Measured by Reinterviews.*

AHS-SMSA—Results from the 1979 AHS-SMSA sample reinterview program were not available at the time this report was being prepared. However, a study was conducted for the 1975 AHS-SMSA sample. The results are presented in the Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1975."

Coverage errors—In errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional (non-mobile home or trailer) new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units. Due to 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, in which case, 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, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit.

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 estimated that the 1975 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 for the 1979 survey, the number of missed housing units may be considerably less for 1979.

The final ratio estimation procedure corrects for these deficiencies as far as the count of total housing is concerned; i.e., it adjusts the estimate of the total housing inventory to the best available estimate. However, biases of subtotals would still remain.

Rounding errors—For errors associated with processing, the rounding of estimates introduces another source of error in the data, the severity of which depends on 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-SMSA sample—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 each other. 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 due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by 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, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. 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.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two 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 (page App-50) are approximations to the standard errors of various estimates shown in this report for this SMSA. 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. Standard errors applicable to estimates of characteristics of the 1975 housing inventory can be found in the AHS Series H-170 reports for 1975.

Table I (page App-50) presents the standard errors applicable to estimates of characteristics of the 1979 housing inventory as well as estimates of characteristics of the 1975-1979 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 7,230 for the total SMSA, 4,350 for the central city of the SMSA, and 5,710 for the balance of the SMSA.

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 (page App-50) presents the standard errors of estimated percentages for the 1979 housing inventory as well as estimated percentages of the 1975-1979 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

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 as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is 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:

$$(100)\left(\frac{x}{y}\right)\sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}$$

where: x = the numerator of the ratio
y = the denominator of the ratio
σ_x = the standard error of the numerator
σ_y = the standard error of the denominator

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1979 there were 142,500 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 1,780. 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
100,000	1,800
142,500	x
150,000	1,780

The entry for "x" is determined as follows by vertically interpolating between 1,800 and 1,780.

$$142{,}500-100{,}000 = 42{,}500$$
$$150{,}000-100{,}000 = 50{,}000$$

$$1{,}800 + \frac{42{,}500}{50{,}000}(1{,}780-1{,}800) = 1{,}780$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 140,720 to 144,280 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1979 owner-occupied housing units lies within a range computed in this way, would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 139,650 to 145,350 housing units with 90 percent confidence; and that the average estimate lies within the interval from 138,940 to 146,060 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 142,500 owner-occupied housing units, 30,800, or 21.6 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 21.6 percent is approximately 0.8 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	21.6	25 or 75
100,000	0.7	a	1.0
142,500		p	
150,000	0.6	b	0.8

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

$$21.6-10.0 = 11.6$$
$$25.0-10.0 = 15.0$$

$$0.7 + \frac{11.6}{15.0}(1.0-0.7) = 0.9$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.6 and 0.8.

$$21.6-10.0 = 11.6$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{11.6}{15.0}(0.8-0.6) = 0.8$$

3. The entry for "p" is then determined by vertical interpolation between 0.9 and 0.8.

$$142{,}500-100{,}000 = 42{,}500$$
$$150{,}000-100{,}000 = 50{,}000$$

$$0.9 + \frac{42{,}500}{50{,}000}(0.8-0.9) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 20.8 to 22.4 percent; the 90-percent confidence interval is from 20.3 to 22.9 percent; and the 95-percent confidence interval is from 20.0 to 23.2 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. 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. Due to the overlap of the 1975 and 1979 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1975 and 1979 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1979 there were 72,100 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 41,300. Table I shows the standard error of 72,100 is approximately 1,670, and the standard error of 30,800 is approximately 1,200. Therefore, the standard error of the estimated difference of 41,300 is about 2,060.

$$2{,}060 = \sqrt{(1{,}670)^2 + (1{,}200)^2}$$

Consequently, the 68-percent confidence interval for the 41,300 difference is from 39,240 to 43,360 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 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 38,000 to 44,600 housing units, and the 95-percent confidence interval is from 37,180 to 45,420 housing units. Thus, we can conclude with 95 percent confidence that the number of 1979 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied housing units with two bedrooms, since the 95-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 deter-

mine an interval about the estimated median so that there is a stated degree of confidence 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. These two distribution intervals could be different.

For about 68 out of 100 possible samples, the average median from all possible samples would lie between these two values.

A two-standard-error confidence interval may be determined by finding the values corresponding to 50 percent plus and minus twice the standard error determined in step 1. For about 95 out of 100 possible samples, the average median from all possible samples would lie between these two values.

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.9. The base of the distribution from which this median was determined is 142,500 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 142,500 is approximately 1.0 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.0 and 52.0.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with 3 persons (for purposes of calculating the median, the category of 3 persons is considered to be from 2.5 to 3.5 persons) contains the 48.0 percent derived in step 2. About 59,300 housing units or 41.6 percent fall below this interval, and 26,600 housing units or 18.7 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(48.0-41.6)}{18.7} = 2.8$$

Similarly, the interval for owner-occupied housing units with 3 persons contains the 52.0 percent derived in step 2. About 59,300 housing units or 41.6 percent fall below this interval, and 26,600 housing units or 18.7 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(52.0-41.6)}{18.7} = 3.1$$

Thus, the 95-percent confidence interval ranges from 2.8 to 2.7 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1979 Housing Inventory and for Estimated Number of 1975-1979 Lost Units for the Hartford, Conn., SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city
0	60	50	60
100	70	70	80
200	100	100	110
500	170	160	170
700	200	190	200
1,000	230	220	240
2,500	370	350	370
5,000	520	480	520
10,000	730	640	730
25,000	1,110	830	1,110
50,000	1,480	500	1,440
55,400	1,540	–	1,490
75,000	1,690	–	1,600
100,000	1,800	–	1,630
150,000	1,780	–	1,310
188,100	1,540	–	–
200,000	1,400	–	–
243,500	–	–	–

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.3 for the total SMSA, 1.4 for the central city, and 1.3 for the balance (not in central city) estimates. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero since these estimates were derived from census data which are not subject to sampling error. However, these estimates are subject to the nonsampling error associated with 1970 and 1980 censuses and with the interpolation procedures used to derive these estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1979 Housing Inventory and for Estimated Percentages of 1975-1979 Lost Housing Units for the Hartford, Conn., SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	36.1	36.1	36.1	36.1	36.1	37.6
200	22.0	22.0	22.0	22.0	23.0	26.6
500	10.2	10.2	10.2	10.2	14.6	16.8
700	7.5	7.5	7.5	8.5	12.3	14.2
1,000	5.4	5.4	5.4	7.1	10.3	11.9
2,500	2.2	2.2	3.3	4.5	6.5	7.5
5,000	1.1	1.1	2.3	3.2	4.6	5.3
10,000	0.6	0.7	1.6	2.3	3.3	3.8
25,000	0.2	0.5	1.0	1.4	2.1	2.4
50,000	0.11	0.3	0.7	1.0	1.5	1.7
75,000	0.08	0.3	0.6	0.8	1.2	1.4
100,000	0.06	0.2	0.5	0.7	1.0	1.2
150,000	0.04	0.2	0.4	0.6	0.8	1.0
200,000	0.03	0.2	0.4	0.5	0.7	0.8
250,000	0.02	0.15	0.3	0.5	0.7	0.8

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than 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 1.3 for the total SMSA, 1.4 for the central city, and 1.3 for the balance (not in central city).

Honolulu 1979

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-41
Annual Housing Survey App-41
Designation of sample housing units for the 1979 survey. App-41
Selection of the 1976 AHS-SMSA sample App-41
1976-1979 additions to the housing inventory App-43
Sample selection for the 1976 Coverage Improvement Program. App-43
Coverage improvement for deficiency 1 . . . . . App-43
Coverage improvement for deficiency 2 . . . . . App-43
Coverage improvement for deficiencies 3-6 . . . . . App-43
1970 Census of Population and Housing. App-44
ESTIMATION. App-44
1979 housing inventory App-44
1976-1979 lost housing units . . . App-45
1976 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES. App-46
Nonsampling errors. App-46
1970 census App-46
AHS-SMSA App-46
Coverage errors App-46
Rounding errors. App-46
Sampling errors for the AHS-SMSA sample App-47
Illustration of the use of the standard error tables App-47
Differences. App-48
Illustration of the computation of the standard error of a difference App-48
Medians App-48
Illustration of the computation of the 95-percent confidence interval of a median App-49
Standard error tables App-50

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 15 SMSA's in this report series (H-170-79) are based on data collected from the 1979 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 15 SMSA's, the data were collected during the 12-month period from April 1979 through March 1980 with one-twelfth of the sample housing units being visited each month.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1979 were interviewed previously in either 1975 or 1976 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1979 group of SMSA's, the largest SMSA from three of the four census regions of the United States is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All the remaining 12 SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the 1979 group are: Chicago, Ill.; Houston, Tex.; and Seattle-Everett, Wash.

The remaining SMSA's in the 1979 group are: Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.

In this SMSA, 4,768 housing units were eligible for interview. Of these sample housing units, 244 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 4,768 housing units eligible for interview, 387 units were visited but were not eligible for interview because they were condemned, unfit, demolished, converted to group quarters use, etc.

Designation of sample housing units for the 1979 survey—The sample housing units designated to be interviewed in the 1979 survey consisted of the following categories which are described in detail in the succeeding sections.

1. All sample housing units that were interviewed in the 1976 survey.
2. 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 1976 survey. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1979 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1976 survey. (This sample represented the housing units built in permit-issuing areas, since the 1976 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1976 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1976 survey.)
5. All sample housing units that were selected as part of the 1976 Coverage Improvement Program. (This sample represented most of the housing units which, until 1976, did not have a chance of selection.)

Selection of the 1976 AHS-SMSA sample—In 1970, the sample for the SMSA's which were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in

the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). In 1979, the following four SMSA's were 100-percent permit-issuing: Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; and Miami, Fla. The remaining 11 SMSA's contain a sample from the nonpermit universe.

Sampling operations, described in the following paragraphs, were performed separately within the central city and the balance of the SMSA for each of the sample frames. The overall sampling rate used to select the sample for each SMSA was determined by the size of the sample. Thus, for the three largest SMSA's, the overall sampling rate differed by central city and the balance of the SMSA, since the sample was divided equally between the central city and the balance of the SMSA. The remaining SMSA's had an overall sampling rate about the same for the sample selected from both the central city and the balance of the SMSA, since the sample was distributed proportionately between the central city and the balance of the SMSA according to the distribution of the total housing units in each sector.

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA 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 housing unit records were stratified by race of head (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 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household 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, for this SMSA, 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 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. 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 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 SMSA. 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.

The second frame from which this SMSA 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 this SMSA. Prior to sample selection, the list of permits was chronologically 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.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS 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 (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

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 are included.

1976-1979 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1976 survey, was selected to represent housing units built in permit-issuing areas since the 1976 survey. Sampling procedures were identical to those used in selecting the 1970-1976 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1976, to identify any housing units missed in the 1976 survey or any housing units added since the 1976 survey.

Sample selection for the 1976 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

1. New construction housing units from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes and trailers placed in parks either missed during the 1970 census or established since the 1970 census.
3. Housing units missed in the 1970 census.
4. Housing units converted to residential use that were nonresidential at the time of the 1970 census.
5. Housing units that have been moved onto their present site since the 1970 census.
6. Mobile homes and trailers placed outside parks since the 1970 census or vacant at the time of the 1970 census.

For each of the five SMSA's interviewed for the first time in 1975, a full implementation of the Coverage Improvement Program was conducted as a part of the 1979 AHS. For each of the 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with some updating and refining as a part of the 1979 AHS. The following discussion of coverage improvement procedures applies to both the 1976 and 1979 programs. For the 10 (1976) SMSA's, estimates of housing units added by a specific procedure reflect units added in 1976 as well as any additions that resulted from the updating and refining in 1979.

Coverage improvement for deficiency 1—A sample of new construction housing units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 1979 SMSA's. Two different procedures were used. For the SMSA's first interviewed in 1976, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. These new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units).

The first stage sample selection was a sample of permit offices, and the second stage a sample of the 1969 permits within each of the selected permit offices. (In the Honolulu, Hawaii, SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second stage.) For the third stage, structures of size three or more were divided into clusters of an expected size of four housing units and a sample of clusters was selected.

For the SMSA's first interviewed in 1975, housing units whose permits were issued before January 1970, but which were completed after April 1970, were identified from the Survey of Construction (SOC), a survey of building permits conducted monthly by the Bureau of the Census. These housing units were then sampled at one-third the rate of regular AHS units. These procedures added an estimated 2,385 new construction housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiency 2—In permit-issuing areas, a sample of mobile homes and trailers placed in parks that were missed by the census or established after the census was selected in two stages. First, a sample of tracts was selected and canvassed. All parks were listed and then matched back to the 1970 census to identify parks missed by the census and parks established after the census. Second, the parks were divided into clusters of an expected size of four sites and a sample of clusters was selected and interviewed. Each of the sample housing units represented the same number of units that the regular AHS sample housing units represented. Since research showed that there were no mobile home or trailer parks in the central cities of the Chicago, Ill.; Hartford, Conn.; and Portland, Oreg.-Wash.; SMSA's, this procedure was only implemented outside the central cities for these SMSA's. Based on a cost-benefit analysis, this procedure was updated as part of the 1979 AHS only for the Denver, Colo. and Las Vegas, Nev., SMSA's. This procedure added no housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed housing units were sampled by one of two procedures. The first procedure was designed to represent units from the following types of missed structures (structures that had no chance of selection for the AHS):

1. Structures missed in the 1970 census.
2. Structures that were completely nonresidential in the 1970 census but now contain residential housing units.
3. Mobile homes and trailers that had been placed outside parks since the 1970 census and have a utility hookup, or were on the site during the present survey but not occupied on April 1, 1970, or had no utility hookup but were occupied by persons with no usual residence elsewhere.
4. Housing units that had been moved onto their present site since the 1970 census.

Initially, a subsample of regular AHS units was selected at a rate of 1 in 24 for the Chicago, Ill.; Hartford, Conn.; Miami, Fla.; Milwaukee, Wis.; and Portland, Oreg.-Wash.; SMSA's, and at a rate of 1 in 22.2152 for the 10 remaining SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing units were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and housing units within

these structures were interviewed. In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. Although the Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Houston, Tex.; Omaha, Nebr.-Iowa; and Raleigh, N.C.; SMSA's, were interviewed for the first time in 1976 and had these sampling procedures implemented at that time, some 1976 coverage improvement assignments were not sent out to be interviewed due to time limitations. The sampling was completed as part of the 1979 AHS and this procedure added an estimated 1,557 housing units to the coverage of the housing inventory of this SMSA.

The second procedure was designed to represent missed housing units from structures represented in the AHS. These missed housing units were:

1. Housing units missed in the 1970 census.
2. Nonresidential space converted to residential use since the 1970 census in structures that contained some residential housing units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structures selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added no housing units to the coverage of the housing inventory for this SMSA.

1970 Census of Population and Housing—The estimates pertaining to the 1970 housing inventory (i.e., the housing inventory that existed at the time of the 1970 census) are based on either 20-, 15-, or 5-percent sample data collected in April 1970 for the Decennial Census of Population and Housing. A detailed description of the sample design employed for the 1970 census can be obtained in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

ESTIMATION

The 1979 AHS sample produced two types of estimates for each SMSA: Estimates pertaining to characteristics of the housing inventory at the time of the interview (i.e., the 1979 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1976 (i.e., 1976-1979 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1979 housing inventory—The AHS estimates of characteristics of the 1979 housing inventory were produced using a 2-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and a 3-stage ratio estimation procedure for the remaining 11 SMSA's. Prior to 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 the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of one or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes and trailers from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample housing units from both the nonpermit universe and the coverage improvement universe (if units were not included in any of the previous cells).

The following first-stage 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 permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of housing units enumerated in areas under the jurisdiction of permit-issuing offices. The denominators of the ratios were obtained from weighted estimates of all the AHS sample housing units within the corresponding ratio estimation categories using the existing weight (i.e., the basic weight times the noninterview factor). The computed ratio estimation factor was then applied to the existing weight for each sample housing unit within the corresponding ratio estimation category.

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 permit-issuing universe. Prior to the AHS sample selection within each SMSA, 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 sample. Since the number of housing units deleted from the AHS universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the Houston, Tex., SMSA, a second ratio estimation procedure was employed for new construction housing units outside the Houston central city. Due to rapid growth in the

nonpermit-issuing universe the sampling methods did not adequately represent the new construction outside the Houston central city. Because of this deficiency, separate factors were computed for the following cells:

1. Conventional housing units (non-mobile home or trailer) outside the central city, which were built between the 1970 census and the 1976 survey.
2. Mobile homes and trailers outside the central city with a model year between the 1970 census and 1976 survey.
3. Conventional housing units outside the central city built between the 1976 and 1979 surveys.
4. Mobile homes and trailers outside the central city with a model year between the 1976 and 1979 surveys.

In a specific cell, the ratio estimation factor was equal to the following:

$$\frac{\text{Weighted count of housing units in the cell} + \text{Estimate of the undercoverage in the cell}}{\text{Weighted count of housing units in the cell}}$$

The computed second-stage ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell. For the Baltimore, Md.; Buffalo, N.Y.; Chicago, Ill.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; and Seattle-Everett, Wash.; SMSA's, a second-stage ratio estimation procedure was employed to adjust the central city/balance distribution of the weighted sample estimate of new construction housing units built since the last survey in permit-issuing areas to an independently derived estimate of this distribution.

This ratio estimation factor was calculated separately for the central city and balance of each SMSA and was applied to all new construction housing units from permit-issuing areas within the corresponding sector (central city or balance of the SMSA). This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}{\text{(Sample estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}$$

The independent estimates of new construction were based upon the number of authorized building permits which were determined from the Survey of Construction (SOC). The sample estimates were obtained from the weighted estimate of the AHS-SMSA sample housing units after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

This second-stage ratio estimation procedure was not implemented in the Honolulu, Hawaii; Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's. In the Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, the central city permit offices, sources of SOC information on permits, had expanded their coverage of permits to include areas greater than those covered by the 1970 census central city definitions. Since the AHS-SMSA survey uses the 1970 definitions, use of this ratio estimation procedure would not correctly adjust the central city/balance distribution of new construction housing units. In the Honolulu, Hawaii, SMSA, this procedure could not be used because the permit office covers the entire SMSA and does not differentiate between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Neb.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and as a third-stage ratio estimation procedure for the remaining 11 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the October 1979 housing inventory of each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sectors. This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the October 1979 housing inventory for the corresponding sector of the SMSA)}}{\text{(AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA)}}$$

The numerator of this ratio was derived using 1970 and 1980 census counts. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight after the first-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and the existing weight after the second-stage ratio estimation procedure for the 11 remaining SMSA's.

The computed ratio estimation factors for the central city and balance of the SMSA's were then applied to the existing weight for all corresponding sample units and the resulting product was used as the final weight for tabulation purposes.

The effect of the total housing unit ratio estimation procedure, as well as the overall 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 SMSA as a whole, it can be expected that the sample estimates will be improved when the sample housing population, or different portions of it, are brought into agreement with known good estimates of the SMSA housing population.

1976-1979 lost housing units—The 1976-1979 lost housing unit (housing unit removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1976 housing inventory, as was described in the 1976 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1976-1979 lost housing units existed, by definition, in the 1976 housing inventory, there was a 1976 housing inventory weight associated with each 1976-1979 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1976-1979 lost housing units.

1976 estimation procedure—This report presents data on the housing characteristics of the 1976 housing inventory from the 1976 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1976.

Ratio estimation procedure of the 1970 Census of Population and Housing—This report presents data on the housing characteristics of the 1970 housing inventory from the 1970 Census of Population and Housing. The statistics based on 1970 census sample data employed a ratio estimation procedure which was applied separately for each of the three census samples. A detailed description of this ratio estimation procedure can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

RELIABILITY OF THE ESTIMATES

There are two types of possible errors associated with estimates based on data from sample surveys—sampling and nonsampling errors. The following is a description of the sampling and nonsampling errors associated with the AHS-SMSA sample and of the nonsampling errors associated with the 1970 census estimates. A description of the sampling errors associated with the sample estimates from the 1970 census can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

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; and other errors of collection, response, processing, 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 1970 Census of Population and Housing and the 1976 AHS-SMSA sample.

1970 census—A number of studies were conducted to measure two types of general errors associated with 1970 census estimates—"coverage" and "content" errors. The "coverage" errors determined how completely housing units were counted in the census and the extent to which occupancy status was erroneously reported. The "content" errors measured the accuracy of the data collected for enumerated housing units. These errors were measured by reinterviews, record checks, and other surveys.

The detailed results of these studies, as well as the methodology employed, are presented in the 1970 Census of Population and Housing Evaluation and Research Program Reports, Series PHC(E)-5, *The Coverage of Housing in the 1970 Census,* and PHC(E)-10, *Accuracy of Data for Selected Housing Characteristics as Measured by Reinterviews.*

AHS-SMSA—Results from the 1979 AHS-SMSA sample reinterview program were not available at the time this report was being prepared. However, a study was conducted for the 1976 AHS-SMSA sample. The results are presented in the Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1976."

Coverage errors—In errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional (non-mobile home or trailer) new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units. Due to 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, in which case, 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, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit.

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 estimated that the 1976 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 for the 1979 survey, the number of missed housing units may be considerably less for 1979.

The final ratio estimation procedure corrects for these deficiencies as far as the count of total housing is concerned; i.e., it adjusts the estimate of the total housing inventory to the best available estimate. However, biases of subtotals would still remain.

Rounding errors—For errors associated with processing, the rounding of estimates introduces another source of error in the data, the severity of which depends on 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-SMSA sample—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 each other. 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 due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by 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, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. 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.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two 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 (page App-50) are approximations to the standard errors of various estimates shown in this report for this SMSA. 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. Standard errors applicable to estimates of characteristics of the 1976 housing inventory can be found in the AHS Series H-170 reports for 1976.

Table I (page App-50) presents the standard errors applicable to estimates of characteristics of the 1979 housing inventory as well as estimates of characteristics of the 1976-1979 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 10,800 for the total SMSA, 8,160 for the central city of the SMSA, and 6,700 for the balance of the SMSA.

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 (page App-50) presents the standard errors of estimated percentages for the 1979 housing inventory as well as estimated percentages of the 1976-1979 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

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 as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is 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:

$$(100)\left(\frac{x}{y}\right)\sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}$$

where: x = the numerator of the ratio
y = the denominator of the ratio
σ_x = the standard error of the numerator
σ_y = the standard error of the denominator

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1979 there were 91,100 detached owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 1,780. 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
75,000	1,710
91,100	x
100,000	1,820

The entry for "x" is determined as follows by vertically interpolating between 1,710 and 1,820.

$$91{,}100-75{,}000 = 16{,}100$$
$$100{,}000-75{,}000 = 25{,}000$$

$$1{,}710 + \frac{16{,}100}{25{,}000}(1{,}820-1{,}710) = 1{,}780$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 89,320 to 92,880 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1979 owner-occupied housing units lies within a range computed in this way, would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 88,250 to 93,950 housing units with 90 percent confidence; and that the average estimate lies within the interval from 87,540 to 94,660 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 117,200 owner-occupied housing units, 18,300, or 15.6 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 15.6 percent is approximately 0.8 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	15.6	25 or 75
100,000.	0.7	a	1.0
117,200.		p	
150,000.	0.6	b	0.8

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

$$15.6-10.0 = 5.6$$
$$25.0-10.0 = 15.0$$

$$0.7 + \frac{5.6}{15.0}(1.0-0.7) = 0.8$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.6 and 0.8.

$$15.6-10.0 = 5.6$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{5.6}{15.0}(0.8-0.6) = 0.7$$

3. The entry for "p" is then determined by vertical interpolation between 0.8 and 0.7.

$$117{,}200-100{,}000 = 17{,}200$$
$$150{,}000-100{,}000 = 50{,}000$$

$$0.8 + \frac{17{,}200}{50{,}000}(0.7-0.8) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 14.8 to 16.4 percent; the 90-percent confidence interval if from 14.3 to 16.9 percent; and the 95-percent confidence interval is from 14.0 to 17.2 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. 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. Due to the overlap of the 1976 and 1979 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1979 characteristics

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1979 there were 63,200 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 44,900. Table I shows the standard error of 63,200 is approximately 1,610, and the standard error of 18,300 is approximately 950. Therefore, the standard error of the estimated difference of 44,900 is about 1,870.

$$1{,}870 = \sqrt{(1{,}610)^2 + (950)^2}$$

Consequently, the 68-percent confidence interval for the 44,900 difference is from 43,030 to 46,770 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 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 41,910 to 47,890 housing units, and the 95-percent confidence interval is from 41,160 to 48,640 housing units. Thus, we can conclude with 95 percent confidence that the number of 1979 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied housing units with two bedrooms, since the 95-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 deter-

mine an interval about the estimated median so that there is a stated degree of confidence 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. These two distribution intervals could be different.

For about 68 out of 100 possible samples, the average median from all possible samples would lie between these two values.

A two-standard-error confidence interval may be determined by finding the values corresponding to 50 percent plus and minus twice the standard error determined in step 1. For about 95 out of 100 possible samples, the average median from all possible samples would lie between these two values.

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 3.3. The base of the distribution from which this median was determined is 117,200 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 117,200 is approximately 1.1 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 47.8 and 52.2.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with 3 persons (for purposes of calculating the median, the category of 3 persons is considered to be from 2.5 to 3.5 persons) contains the 47.8 percent derived in step 2. About 41,800 housing units or 35.7 percent fall below this interval, and 20,100 housing units or 17.2 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(47.8-35.7)}{17.2} = 3.2$$

Similarly, the interval for owner-occupied housing units with 3 persons contains the 52.2 percent derived in step 2. About 41,800 housing units or 35.7 percent fall below this interval, and 20,100 housing units or 17.2 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(52.2-35.7)}{17.2} = 3.5$$

Thus, the 95-percent confidence interval ranges from 3.2 to 3.5 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1979 Housing Inventory and for Estimated Number of 1976-1979 Lost Units for the Honolulu, Hawaii, SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city
0	60	60	50
100	70	80	70
200	110	110	100
500	170	170	150
700	200	200	180
1,000	240	240	220
2,500	370	380	340
5,000	520	530	480
10,000	730	730	660
25,000	1,120	1,090	960
50,000	1,490	1,360	1,130
75,000	1,710	1,420	1,050
100,000	1,820	1,280	590
107,800	1,840	1,200	–
140,300	1,840	–	–
150,000	1,820	–	–
200,000	1,470	–	–
248,100	–	–	–

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2 for the total SMSA, 1.3 for the central city, and 1.2 for the balance (not in central city) estimates. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero since these estimates were derived from census data which are not subject to sampling error. However, these estimates are subject to the nonsampling errors associated with the 1970 and 1980 censuses and with the interpolation procedures used to derive these estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1979 Housing Inventory and for Estimated Percentages of 1976-1979 Lost Units for the Honolulu, Hawaii, SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	36.5	36.5	36.5	36.5	36.5	37.9
200	22.3	22.3	22.3	22.3	23.2	26.8
500	10.3	10.3	10.3	10.3	14.7	16.9
700	7.6	7.6	7.6	8.6	12.4	14.3
1,000	5.4	5.4	5.4	7.2	10.4	12.0
2,500	2.2	2.2	3.3	4.5	6.6	7.6
5,000	1.1	1.1	2.3	3.2	4.6	5.4
10,000	0.6	0.8	1.7	2.3	3.3	3.8
25,000	0.2	0.5	1.0	1.4	2.1	2.4
50,000	0.1	0.3	0.7	1.0	1.5	1.7
75,000	0.1	0.3	0.6	0.8	1.2	1.4
100,000	0.1	0.2	0.5	0.7	1.0	1.2
150,000	0.0	0.2	0.4	0.6	0.8	1.0
200,000	0.0	0.2	0.4	0.5	0.7	0.8
250,000	0.0	0.2	0.3	0.5	0.7	0.8

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than 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 1.2 for the total SMSA, 1.3 for the central city, and 1.2 for the balance (not in central city).

Houston 1979

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-41
Annual Housing Survey . . . App-41
Designation of sample housing units for the 1979 survey . App-41
Selection of the 1976 AHS-SMSA sample . App-41
1976-1979 additions to the housing inventory . App-43
Sample selection for the 1976 Coverage Improvement Program. App-43
Coverage improvement for deficiency 1 . App-43
Coverage improvement for deficiency 2 . App-43
Coverage improvement for deficiencies 3-6 . App-43
1970 Census of Population and Housing . App-44
ESTIMATION . App-44
1979 housing inventory . App-44
1976-1979 lost housing units . App-45
1976 estimation procedure . App-46
Ratio estimation procedure of the 1970 Census of Population and Housing . App-46
RELIABILITY OF THE ESTIMATES . App-46
Nonsampling errors . App-46
1970 census . App-46
AHS-SMSA . App-46
Coverage errors . App-46
Rounding errors . App-46
Sampling errors for the AHS-SMSA sample . App-47
Illustration of the use of the standard error tables . App-47
Differences . App-48
Illustration of the computation of the standard error of a difference . App-48
Medians . App-48
Illustration of the computation of the 95-percent confidence interval of a median . App-49
Standard error tables . App-49

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 15 SMSA's in this report series (H-170-79) are based on data collected from the 1979 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 15 SMSA's, the data were collected during the 12-month period from April 1979 through March 1980 with one-twelfth of the sample housing units being visited each month.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1979 were interviewed previously in either 1975 or 1976 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1979 group of SMSA's, the largest SMSA from three of the four census regions of the United States is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All the remaining 12 SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the 1979 group are: Chicago, Ill.; Houston, Tex.; and Seattle-Everett, Wash.

The remaining SMSA's in the 1979 group are: Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.

In this SMSA, 17,736 housing units were eligible for interview. Of these sample housing units, 842 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 17,736 housing units eligible for interview, 1,624 units were visited but were not eligible for interview because they were condemned, unfit, demolished, converted to group quarters use, etc.

Designation of sample housing units for the 1979 survey—The sample housing units designated to be interviewed in the 1979 survey consisted of the following categories which are described in detail in the succeeding sections.

1. All sample housing units that were interviewed in the 1976 survey.
2. 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 1976 survey. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1979 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1976 survey. (This sample represented the housing units built in permit-issuing areas, since the 1976 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1976 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1976 survey.)
5. All sample housing units that were selected as part of the 1976 Coverage Improvement Program. (This sample represented most of the housing units which, until 1976, did not have a chance of selection.)

Selection of the 1976 AHS-SMSA sample—In 1970, the sample for the SMSA's which were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in

the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (nonpermit universe). In 1979, the following four SMSA's were 100-percent permit-issuing: Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; and Miami, Fla. The remaining 11 SMSA's contain a sample from the nonpermit universe.

Sampling operations, described in the following paragraphs, were performed separately within the central city and the balance of the SMSA for each of the sample frames. The overall sampling rate used to select the sample for each SMSA was determined by the size of the sample. Thus, for the three largest SMSA's, the overall sampling rate differed by central city and the balance of the SMSA, since the sample was divided equally between the central city and the balance of the SMSA. The remaining SMSA's had an overall sampling rate about the same for the sample selected from both the central city and the balance of the SMSA, since the sample was distributed proportionately between the central city and the balance of the SMSA according to the distribution of the total housing units in each sector.

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA 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 housing unit records were stratified by race of head (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 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household 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, for this SMSA, 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 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. 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 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 SMSA. 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.

The second frame from which this SMSA 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 this SMSA. Prior to sample selection, the list of permits was chronologically 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.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS 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 (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

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 are included.

1976-1979 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1976 survey, was selected to represent housing units built in permit-issuing areas since the 1976 survey. Sampling procedures were identical to those used in selecting the 1970-1976 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1976, to identify any housing units missed in the 1976 survey or any housing units added since the 1976 survey.

Sample selection for the 1976 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

1. New construction housing units from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes and trailers placed in parks either missed during the 1970 census or established since the 1970 census.
3. Housing units missed in the 1970 census.
4. Housing units converted to residential use that were nonresidential at the time of the 1970 census.
5. Housing units that have been moved onto their present site since the 1970 census.
6. Mobile homes and trailers placed outside parks since the 1970 census or vacant at the time of the 1970 census.

For each of the five SMSA's interviewed for the first time in 1975, a full implementation of the Coverage Improvement Program was conducted as a part of the 1979 AHS. For each of the 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with some updating and refining as a part of the 1979 AHS. The following discussion of coverage improvement procedures applies to both the 1976 and 1979 programs. For the 10 (1976) SMSA's, estimates of housing units added by a specific procedure reflect units added in 1976 as well as any additions that resulted from the updating and refining in 1979.

Coverage improvement for deficiency 1—A sample of new construction housing units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 1979 SMSA's. Two different procedures were used. For the SMSA's first interviewed in 1976, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. These new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units).

The first stage sample selection was a sample of permit offices, and the second stage a sample of the 1969 permits within each of the selected permit offices. (In the Honolulu, Hawaii, SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second stage.) For the third stage, structures of size three or more were divided into clusters of an expected size of four housing units and a sample of clusters was selected.

For the SMSA's first interviewed in 1975, housing units whose permits were issued before January 1970, but which were completed after April 1970, were identified from the Survey of Construction (SOC), a survey of building permits conducted monthly by the Bureau of the Census. These housing units were then sampled at one-third the rate of regular AHS units. These procedures added an estimated 12,628 new construction housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiency 2—In permit-issuing areas, a sample of mobile homes and trailers placed in parks that were missed by the census or established after the census was selected in two stages. First, a sample of tracts was selected and canvassed. All parks were listed and then matched back to the 1970 census to identify parks missed by the census and parks established after the census. Second, the parks were divided into clusters of an expected size of four sites and a sample of clusters was selected and interviewed. Each of the sample housing units represented the same number of units that the regular AHS sample housing units represented. Since research showed that there were no mobile home or trailer parks in the central cities of the Chicago, Ill.; Hartford, Conn.; and Portland, Oreg.-Wash.; SMSA's, this procedure was only implemented outside the central cities for these SMSA's. Based on a cost-benefit analysis, this procedure was updated as part of the 1979 AHS only for the Denver, Colo. and Las Vegas, Nev., SMSA's. This procedure added 2,877 housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed housing units were sampled by one of two procedures. The first procedure was designed to represent units from the following types of missed structures (structures that had no chance of selection for the AHS):

1. Structures missed in the 1970 census.
2. Structures that were completely nonresidential in the 1970 census but now contain residential housing units.
3. Mobile homes and trailers that had been placed outside parks since the 1970 census and have a utility hookup, or were on the site during the present survey but not occupied on April 1, 1970, or had no utility hookup but were occupied by persons with no usual residence elsewhere.
4. Housing units that had been moved onto their present site since the 1970 census.

Initially, a subsample of regular AHS units was selected at a rate of 1 in 24 for the Chicago, Ill.; Hartford, Conn.; Miami, Fla.; Milwaukee, Wis.; and Portland, Oreg.-Wash.; SMSA's, and at a rate of 1 in 22.2152 for the 10 remaining SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing units were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and housing units within

these structures were interviewed. In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. Although the Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Houston, Tex.; Omaha, Nebr.-Iowa; and Raleigh, N.C.; SMSA's, were interviewed for the first time in 1976 and had these sampling procedures implemented at that time, some 1976 coverage improvement assignments were not sent out to be interviewed due to time limitations. The sampling was completed as part of the 1979 AHS and this procedure added an estimated 8,807 housing units to the coverage of the housing inventory of this SMSA.

The second procedure was designed to represent missed housing units from structures represented in the AHS. These missed housing units were:

1. Housing units missed in the 1970 census.
2. Nonresidential space converted to residential use since the 1970 census in structures that contained some residential housing units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structures selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 904 housing units to the coverage of the housing inventory for this SMSA.

1970 Census of Population and Housing—The estimates pertaining to the 1970 housing inventory (i.e., the housing inventory that existed at the time of the 1970 census) are based on either 20-, 15-, or 5-percent sample data collected in April 1970 for the Decennial Census of Population and Housing. A detailed description of the sample design employed for the 1970 census can be obtained in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

ESTIMATION

The 1979 AHS sample produced two types of estimates for each SMSA: Estimates pertaining to characteristics of the housing inventory at the time of the interview (i.e., the 1979 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1976 (i.e., 1976-1979 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1979 housing inventory—The AHS estimates of characteristics of the 1979 housing inventory were produced using a 2-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and a 3-stage ratio estimation procedure for the remaining 11 SMSA's. Prior to 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 the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of one or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes and trailers from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample housing units from both the nonpermit universe and the coverage improvement universe (if units were not included in any of the previous cells).

The following first-stage 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 permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of housing units enumerated in areas under the jurisdiction of permit-issuing offices. The denominators of the ratios were obtained from weighted estimates of all the AHS sample housing units within the corresponding ratio estimation categories using the existing weight (i.e., the basic weight times the noninterview factor). The computed ratio estimation factor was then applied to the existing weight for each sample housing unit within the corresponding ratio estimation category.

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 permit-issuing universe. Prior to the AHS sample selection within each SMSA, 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 sample. Since the number of housing units deleted from the AHS universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the Houston, Tex., SMSA, a second ratio estimation procedure was employed for new construction housing units outside the Houston central city. Due to rapid growth in the

nonpermit-issuing universe the sampling methods did not adequately represent the new construction outside the Houston central city. Because of this deficiency, separate factors were computed for the following cells:

1. Conventional housing units (non-mobile home or trailer) outside the central city, which were built between the 1970 census and the 1976 survey.
2. Mobile homes and trailers outside the central city with a model year between the 1970 census and 1976 survey.
3. Conventional housing units outside the central city built between the 1976 and 1979 surveys.
4. Mobile homes and trailers outside the central city with a model year between the 1976 and 1979 surveys.

In a specific cell, the ratio estimation factor was equal to the following:

$$\frac{\text{Weighted count of housing units in the cell} + \text{Estimate of the undercoverage in the cell}}{\text{Weighted count of housing units in the cell}}$$

The computed second-stage ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell. For the Baltimore, Md.; Buffalo, N.Y.; Chicago, Ill.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; and Seattle-Everett, Wash.; SMSA's, a second-stage ratio estimation procedure was employed to adjust the central city/balance distribution of the weighted sample estimate of new construction housing units built since the last survey in permit-issuing areas to an independently derived estimate of this distribution.

This ratio estimation factor was calculated separately for the central city and balance of each SMSA and was applied to all new construction housing units from permit-issuing areas within the corresponding sector (central city or balance of the SMSA). This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}{\text{(Sample estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}$$

The independent estimates of new construction were based upon the number of authorized building permits which were determined from the Survey of Construction (SOC). The sample estimates were obtained from the weighted estimate of the AHS-SMSA sample housing units after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

This second-stage ratio estimation procedure was not implemented in the Honolulu, Hawaii; Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's. In the Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, the central city permit offices, sources of SOC information on permits, had expanded their coverage of permits to include areas greater than those covered by the 1970 census central city definitions. Since the AHS-SMSA survey uses the 1970 definitions, use of this ratio estimation procedure would not correctly adjust the central city/balance distribution of new construction housing units. In the Honolulu, Hawaii, SMSA, this procedure could not be used because the permit office covers the entire SMSA and does not differentiate between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Neb.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and as a third-stage ratio estimation procedure for the remaining 11 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the October 1979 housing inventory of each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sectors. This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the October 1979 housing inventory for the corresponding sector of the SMSA)}}{\text{(AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA)}}$$

The numerator of this ratio was derived using 1970 and 1980 census counts. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight after the first-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and the existing weight after the second-stage ratio estimation procedure for the 11 remaining SMSA's.

The computed ratio estimation factors for the central city and balance of the SMSA's were then applied to the existing weight for all corresponding sample units and the resulting product was used as the final weight for tabulation purposes.

The effect of the total housing unit ratio estimation procedure, as well as the overall 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 SMSA as a whole, it can be expected that the sample estimates will be improved when the sample housing population, or different portions of it, are brought into agreement with known good estimates of the SMSA housing population.

1976-1979 lost housing units—The 1976-1979 lost housing unit (housing unit removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1976 housing inventory, as was described in the 1976 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1976-1979 lost housing units existed, by definition, in the 1976 housing inventory, there was a 1976 housing inventory weight associated with each 1976-1979 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1976-1979 lost housing units.

1976 estimation procedure—This report presents data on the housing characteristics of the 1976 housing inventory from the 1976 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1976.

Ratio estimation procedure of the 1970 Census of Population and Housing—This report presents data on the housing characteristics of the 1970 housing inventory from the 1970 Census of Population and Housing. The statistics based on 1970 census sample data employed a ratio estimation procedure which was applied separately for each of the three census samples. A detailed description of this ratio estimation procedure can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

RELIABILITY OF THE ESTIMATES

There are two types of possible errors associated with estimates based on data from sample surveys—sampling and nonsampling errors. The following is a description of the sampling and nonsampling errors associated with the AHS-SMSA sample and of the nonsampling errors associated with the 1970 census estimates. A description of the sampling errors associated with the sample estimates from the 1970 census can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

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; and other errors of collection, response, processing, 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 1970 Census of Population and Housing and the 1976 AHS-SMSA sample.

1970 census—A number of studies were conducted to measure two types of general errors associated with 1970 census estimates—"coverage" and "content" errors. The "coverage" errors determined how completely housing units were counted in the census and the extent to which occupancy status was erroneously reported. The "content" errors measured the accuracy of the data collected for enumerated housing units. These errors were measured by reinterviews, record checks, and other surveys.

The detailed results of these studies, as well as the methodology employed, are presented in the 1970 Census of Population and Housing Evaluation and Research Program Reports, Series PHC(E)-5, *The Coverage of Housing in the 1970 Census,* and PHC(E)-10, *Accuracy of Data for Selected Housing Characteristics as Measured by Reinterviews.*

AHS-SMSA—Results from the 1979 AHS-SMSA sample reinterview program were not available at the time this report was being prepared. However, a study was conducted for the 1976 AHS-SMSA sample. The results are presented in the Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1976."

Coverage errors—In errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional (non-mobile home or trailer) new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units. Due to 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, in which case, 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, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit.

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 estimated that the 1976 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 for the 1979 survey, the number of missed housing units may be considerably less for 1979.

The final ratio estimation procedure corrects for these deficiencies as far as the count of total housing is concerned; i.e., it adjusts the estimate of the total housing inventory to the best available estimate. However, biases of subtotals would still remain.

Rounding errors—For errors associated with processing, the rounding of estimates introduces another source of error in the data, the severity of which depends on 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-SMSA sample—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 each other. 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 due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by 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, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. 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.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two 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 (pages App-49 and App-50) are approximations to the standard errors of various estimates shown in this report for this SMSA. 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. Standard errors applicable to estimates of characteristics of the 1976 housing inventory can be found in the AHS Series H-170 reports for 1976.

Table I (page App-49) presents the standard errors applicable to estimates of characteristics of the 1979 housing inventory as well as estimates of characteristics of the 1976-1979 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 46,530 for the total SMSA, 20,610 for the central city of the SMSA, and 41,420 for the balance of the SMSA.

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.

Tables II and III (page App-50) presents the standard errors of estimated percentages for the 1979 housing inventory as well as estimated percentages of the 1976-1979 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in tables II and III.

Included in tables I, II, and III are estimates of standard errors for estimates of zero and zero percent. These estimates of standard errors are considered as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is 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:

$$(100)\left(\frac{x}{y}\right)\sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}$$

where: x = the numerator of the ratio
y = the denominator of the ratio
σ_x = the standard error of the numerator
σ_y = the standard error of the denominator

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1979 there were 606,100 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 4,982. 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
600,000	4,990
606,100	x
700,000	4,860

The entry for "x" is determined as follows by vertically interpolating between 4,990 and 4,860.

$$606{,}100-600{,}000 = 6{,}100$$
$$700{,}000-600{,}000 = 100{,}000$$

$$4{,}990 + \frac{6{,}100}{100{,}000}(4{,}860-4{,}990) = 4{,}982$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 601,118 to 611,082 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1979 owner-occupied housing units lies within a range computed in this way, would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 598,129 to 614,071 housing units with 90 percent confidence; and that the average estimate lies within the interval from 596,136 to 616,064 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 606,100 owner-occupied housing units, 137,500, or 22.7 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 22.7 percent is approximately 0.6 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	22.7	25 or 75
600,000.	0.4	a	0.6
606,100.		p	
700,000.	0.4	b	0.5

1. The entry for cell "a" is determined by horizontal interpolation between 0.4 and 0.6.

$$22.7-10.0 = 12.7$$
$$25.0-10.0 = 15.0$$

$$0.4 + \frac{12.7}{15.0}(0.6-0.4) = 0.6$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.4 and 0.5.

$$22.7-10.0 = 12.7$$
$$25.0-10.0 = 15.0$$

$$0.4 + \frac{12.7}{15.0}(0.5-0.4) = 0.5$$

3. The entry for "p" is then determined by vertical interpolation between 0.6 and 0.5.

$$606{,}100-600{,}000 = 6{,}100$$
$$700{,}000-600{,}000 = 100{,}000$$

$$0.6 + \frac{6{,}100}{100{,}000}(0.5-0.6) = 0.6$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 22.1 to 23.3 percent; the 90-percent confidence interval is from 21.7 to 23.7 percent; and the 95-percent confidence interval is from 21.5 to 23.9 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. 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. Due to the overlap of the 1976 and 1979 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1979 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1979 there were 340,200 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 202,700. Table I shows the standard error of 340,200 is approximately 4,559, and the standard error of 137,500 is approximately 3,253. Therefore, the standard error of the estimated difference of 202,700 is about 5,601.

$$5{,}601 = \sqrt{(4{,}559)^2 + (3{,}253)^2}$$

Consequently, the 68-percent confidence interval for the 202,700 difference is from 197,099 to 208,301 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 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 193,739 to 211,661 housing units, and the 95-percent confidence interval is from 191,498 to 213,902 housing units. Thus, we can conclude with 95 percent confidence that the number of 1979 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied housing units with two bedrooms, since the 95-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 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 tables II and III, 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. These two distribution intervals could be different.

For about 68 out of 100 possible samples, the average median from all possible samples would lie between these two values.

A two-standard-error confidence interval may be determined by finding the values corresponding to 50 percent plus and minus twice the standard error determined in step 1. For about 95 out of 100 possible samples, the average median from all possible samples would lie between these two values.

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.9. The base of the distribution from which this median was determined is 606,100 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 606,100 is approximately 0.6 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.8 and 51.2.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with 3 persons (for purposes of calculating the median, the category of 3 persons is considered to be from 2.5 to 3.5 persons) contains the 48.8 percent derived in step 2. About 253,200 housing units or 41.8 percent fall below this interval, and 120,500 housing units or 19.9 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(48.8-41.8)}{19.9} = 2.9$$

Similarly, the interval for owner-occupied housing units with 3 persons corresponds to the 51.2 percent derived in step 2. About 253,200 housing units or 41.8 percent fall below this interval, and 120,500 housing units or 19.9 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(51.2-41.8)}{19.9} = 3.0$$

Thus, the 95-percent confidence interval ranges from 2.9 to 3.0 persons. Although it appears that this confidence interval has the sample estimate as the lower limit, it actually is a reflection of the rounding error associated with the median (see the paragraph on rounding errors in the nonsampling errors section of this appendix).

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1979 Housing Inventory and for Estimated Number of 1976-1979 Lost Units for the Houston, Tex., SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city
0	90	100	60
100	90	100	80
200	130	140	110
500	210	220	170
700	250	270	200
1,000	300	320	240
2,500	470	500	380
5,000	660	710	540
10,000	940	990	760
25,000	1,470	1,550	1,180
50,000	2,050	2,140	1,630
75,000	2,490	2,560	1,950
100,000	2,840	2,880	2,190
150,000	3,390	3,330	2,540
200,000	3,810	3,610	2,750
250,000	4,150	3,750	2,850
300,000	4,410	3,770	2,860
400,000	4,780	3,450	2,610
500,000	4,970	2,470	1,830
564,600	5,000	650	—
568,800	5,000	—	—
600,000	4,990	—	—
700,000	4,860	—	—
800,000	4,560	—	—
900,000	4,050	—	—
1,000,000	3,220	—	—
1,100,000	1,690	—	—
1,133,400	—	—	—

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero since these estimates were derived from census data which are not subject to sampling error. However, these estimates are subject to the nonsampling errors associated with the 1970 and 1980 censuses and with the interpolation procedures used to derive these estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1979 Housing Inventory and for Estimated Percentages of 1976-1979 Lost Housing Units for the Houston, Tex., SMSA, and for the Central City of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
200	33.4	33.4	33.4	33.4	33.4	35.4
500	16.7	16.7	16.7	16.7	19.4	22.4
700	12.6	12.6	12.6	12.6	16.4	18.9
1,000	9.1	9.1	9.1	9.5	13.7	15.8
2,500	3.9	3.9	4.4	6.0	8.7	10.0
5,000	2.0	2.0	3.1	4.3	6.1	7.1
10,000	1.0	1.0	2.2	3.0	4.3	5.0
25,000	0.4	0.6	1.4	1.9	2.7	3.2
50,000	0.2	0.4	1.0	1.3	1.9	2.2
75,000	0.13	0.4	0.8	1.1	1.6	1.8
100,000	0.10	0.3	0.7	1.0	1.4	1.6
150,000	0.07	0.3	0.6	0.8	1.1	1.3
200,000	0.05	0.2	0.5	0.7	1.0	1.1
250,000	0.04	0.2	0.4	0.6	0.9	1.0
300,000	0.03	0.2	0.4	0.5	0.8	0.9
400,000	0.03	0.2	0.3	0.5	0.7	0.8
500,000	0.02	0.14	0.3	0.4	0.6	0.7
600,000	0.02	0.13	0.3	0.4	0.6	0.6
700,000	0.01	0.12	0.3	0.4	0.5	0.6
800,000	0.01	0.11	0.2	0.3	0.5	0.6
900,000	0.01	0.11	0.2	0.3	0.5	0.5
1,000,000	0.01	0.10	0.2	0.3	0.4	0.5
1,100,000	0.01	0.10	0.2	0.3	0.4	0.5
1,200,000	0.01	0.09	0.2	0.3	0.4	0.5

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than 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 1.2.

TABLE III. Standard Errors for Estimated Percentages of Housing Units in the 1979 Housing Inventory and for Estimated Percentages of 1976-1979 Lost Housing Units for the Balance (Not in Central City) of the Houston, Tex., SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	36.9	36.9	36.9	36.9	36.9	38.2
200	22.6	22.6	22.6	22.6	23.4	27.0
500	10.5	10.5	10.5	10.5	14.8	17.1
700	7.7	7.7	7.7	8.7	12.5	14.4
1,000	5.5	5.5	5.5	7.2	10.5	12.1
2,500	2.3	2.3	3.3	4.6	6.6	7.6
5,000	1.2	1.2	2.4	3.2	4.7	5.4
10,000	0.6	0.8	1.7	2.3	3.3	3.8
25,000	0.2	0.5	1.1	1.4	2.1	2.4
50,000	0.12	0.3	0.7	1.0	1.5	1.7
75,000	0.08	0.3	0.6	0.8	1.2	1.4
100,000	0.06	0.2	0.5	0.7	1.0	1.2
150,000	0.04	0.2	0.4	0.6	0.9	1.0
200,000	0.03	0.2	0.4	0.5	0.7	0.9
250,000	0.02	0.15	0.3	0.5	0.7	0.8
300,000	0.02	0.14	0.3	0.4	0.6	0.7
400,000	0.01	0.12	0.3	0.4	0.5	0.6
500,000	0.01	0.11	0.2	0.3	0.5	0.5
600,000	0.01	0.10	0.2	0.3	0.4	0.5

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than 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 1.2.

Las Vegas 1979

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-41
Annual Housing Survey App-41
Designation of sample housing units for the 1979 survey. App-41
Selection of the 1976 AHS-SMSA sample App-41
1976-1979 additions to the housing inventory App-43
Sample selection for the 1976 Coverage Improvement Program. App-43
Coverage improvement for deficiency 1 . . . . . App-43
Coverage improvement for deficiency 2 . . . . . App-43
Coverage improvement for deficiencies 3-6 . . . . . App-43
1970 Census of Population and Housing. App-44
ESTIMATION. App-44
1979 housing inventory App-44
1976-1979 lost housing units . . . App-45
1976 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES. App-46
Nonsampling errors. App-46
1970 census App-46
AHS-SMSA App-46
Coverage errors App-46
Rounding errors. App-46
Sampling errors for the AHS-SMSA sample App-47
Illustration of the use of the standard error tables App-47
Differences. App-48
Illustration of the computation of the standard error of a difference App-48
Medians App-48
Illustration of the computation of the 95-percent confidence interval of a median App-49
Standard error tables App-50

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 15 SMSA's in this report series (H-170-79) are based on data collected from the 1979 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 15 SMSA's, the data were collected during the 12-month period from April 1979 through March 1980 with one-twelfth of the sample housing units being visited each month.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1979 were interviewed previously in either 1975 or 1976 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1979 group of SMSA's, the largest SMSA from three of the four census regions of the United States is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All the remaining 12 SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the 1979 group are: Chicago, Ill.; Houston, Tex.; and Seattle-Everett, Wash.

The remaining SMSA's in the 1979 group are: Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.

In this SMSA, 6,351 housing units were eligible for interview. Of these sample housing units, 469 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 6,351 housing units eligible for interview, 632 units were visited but were not eligible for interview because they were condemned, unfit, demolished, converted to group quarters use, etc.

Designation of sample housing units for the 1979 survey—The sample housing units designated to be interviewed in the 1979 survey consisted of the following categories which are described in detail in the succeeding sections.

1. All sample housing units that were interviewed in the 1976 survey.
2. 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 1976 survey. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1979 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1976 survey. (This sample represented the housing units built in permit-issuing areas, since the 1976 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1976 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1976 survey.)
5. All sample housing units that were selected as part of the 1976 Coverage Improvement Program. (This sample represented most of the housing units which, until 1976, did not have a chance of selection.)

Selection of the 1976 AHS-SMSA sample—In 1970, the sample for the SMSA's which were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in

the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (nonpermit universe). In 1979, the following four SMSA's were 100-percent permit-issuing: Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; and Miami, Fla. The remaining 11 SMSA's contain a sample from the nonpermit universe.

Sampling operations, described in the following paragraphs, were performed separately within the central city and the balance of the SMSA for each of the sample frames. The overall sampling rate used to select the sample for each SMSA was determined by the size of the sample. Thus, for the three largest SMSA's, the overall sampling rate differed by central city and the balance of the SMSA, since the sample was divided equally between the central city and the balance of the SMSA. The remaining SMSA's had an overall sampling rate about the same for the sample selected from both the central city and the balance of the SMSA, since the sample was distributed proportionately between the central city and the balance of the SMSA according to the distribution of the total housing units in each sector.

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA 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 housing unit records were stratified by race of head (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 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household 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, for this SMSA, 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 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. 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 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 SMSA. 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.

The second frame from which this SMSA 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 this SMSA. Prior to sample selection, the list of permits was chronologically 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.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS 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 (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

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 are included.

1976-1979 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1976 survey, was selected to represent housing units built in permit-issuing areas since the 1976 survey. Sampling procedures were identical to those used in selecting the 1970-1976 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1976, to identify any housing units missed in the 1976 survey or any housing units added since the 1976 survey.

Sample selection for the 1976 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

1. New construction housing units from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes and trailers placed in parks either missed during the 1970 census or established since the 1970 census.
3. Housing units missed in the 1970 census.
4. Housing units converted to residential use that were nonresidential at the time of the 1970 census.
5. Housing units that have been moved onto their present site since the 1970 census.
6. Mobile homes and trailers placed outside parks since the 1970 census or vacant at the time of the 1970 census.

For each of the five SMSA's interviewed for the first time in 1975, a full implementation of the Coverage Improvement Program was conducted as a part of the 1979 AHS. For each of the 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with some updating and refining as a part of the 1979 AHS. The following discussion of coverage improvement procedures applies to both the 1976 and 1979 programs. For the 10 (1976) SMSA's, estimates of housing units added by a specific procedure reflect units added in 1976 as well as any additions that resulted from the updating and refining in 1979.

Coverage improvement for deficiency 1—A sample of new construction housing units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 1979 SMSA's. Two different procedures were used. For the SMSA's first interviewed in 1976, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. These new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units).

The first stage sample selection was a sample of permit offices, and the second stage a sample of the 1969 permits within each of the selected permit offices. (In the Honolulu, Hawaii, SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second stage.) For the third stage, structures of size three or more were divided into clusters of an expected size of four housing units and a sample of clusters was selected.

For the SMSA's first interviewed in 1975, housing units whose permits were issued before January 1970, but which were completed after April 1970, were identified from the Survey of Construction (SOC), a survey of building permits conducted monthly by the Bureau of the Census. These housing units were then sampled at one-third the rate of regular AHS units. These procedures added an estimated 1,090 new construction housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiency 2—In permit-issuing areas, a sample of mobile homes and trailers placed in parks that were missed by the census or established after the census was selected in two stages. First, a sample of tracts was selected and canvassed. All parks were listed and then matched back to the 1970 census to identify parks missed by the census and parks established after the census. Second, the parks were divided into clusters of an expected size of four sites and a sample of clusters was selected and interviewed. Each of the sample housing units represented the same number of units that the regular AHS sample housing units represented. Since research showed that there were no mobile home or trailer parks in the central cities of the Chicago, Ill.; Hartford, Conn.; and Portland, Oreg.-Wash.; SMSA's, this procedure was only implemented outside the central cities for these SMSA's. Based on a cost-benefit analysis, this procedure was updated as part of the 1979 AHS only for the Denver, Colo. and Las Vegas, Nev., SMSA's. This procedure added 5,933 housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed housing units were sampled by one of two procedures. The first procedure was designed to represent units from the following types of missed structures (structures that had no chance of selection for the AHS):

1. Structures missed in the 1970 census.
2. Structures that were completely nonresidential in the 1970 census but now contain residential housing units.
3. Mobile homes and trailers that had been placed outside parks since the 1970 census and have a utility hookup, or were on the site during the present survey but not occupied on April 1, 1970, or had no utility hookup but were occupied by persons with no usual residence elsewhere.
4. Housing units that had been moved onto their present site since the 1970 census.

Initially, a subsample of regular AHS units was selected at a rate of 1 in 24 for the Chicago, Ill.; Hartford, Conn.; Miami, Fla.; Milwaukee, Wis.; and Portland, Oreg.-Wash.; SMSA's, and at a rate of 1 in 22.2152 for the 10 remaining SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing units were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and housing units within

these structures were interviewed. In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. Although the Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Houston, Tex.; Omaha, Nebr.-Iowa; and Raleigh, N.C.; SMSA's, were interviewed for the first time in 1976 and had these sampling procedures implemented at that time, some 1976 coverage improvement assignments were not sent out to be interviewed due to time limitations. The sampling was completed as part of the 1979 AHS and this procedure added an estimated 226 housing units to the coverage of the housing inventory of this SMSA.

The second procedure was designed to represent missed housing units from structures represented in the AHS. These missed housing units were:

1. Housing units missed in the 1970 census.
2. Nonresidential space converted to residential use since the 1970 census in structures that contained some residential housing units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structures selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 55 housing units to the coverage of the housing inventory for this SMSA.

1970 Census of Population and Housing—The estimates pertaining to the 1970 housing inventory (i.e., the housing inventory that existed at the time of the 1970 census) are based on either 20-, 15-, or 5-percent sample data collected in April 1970 for the Decennial Census of Population and Housing. A detailed description of the sample design employed for the 1970 census can be obtained in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties*, Part 1.

ESTIMATION

The 1979 AHS sample produced two types of estimates for each SMSA: Estimates pertaining to characteristics of the housing inventory at the time of the interview (i.e., the 1979 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1976 (i.e., 1976-1979 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1979 housing inventory—The AHS estimates of characteristics of the 1979 housing inventory were produced using a 2-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and a 3-stage ratio estimation procedure for the remaining 11 SMSA's. Prior to 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 the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of one or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes and trailers from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample housing units from both the nonpermit universe and the coverage improvement universe (if units were not included in any of the previous cells).

The following first-stage 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 permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of housing units enumerated in areas under the jurisdiction of permit-issuing offices. The denominators of the ratios were obtained from weighted estimates of all the AHS sample housing units within the corresponding ratio estimation categories using the existing weight (i.e., the basic weight times the noninterview factor). The computed ratio estimation factor was then applied to the existing weight for each sample housing unit within the corresponding ratio estimation category.

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 permit-issuing universe. Prior to the AHS sample selection within each SMSA, 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 sample. Since the number of housing units deleted from the AHS universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the Houston, Tex., SMSA, a second ratio estimation procedure was employed for new construction housing units outside the Houston central city. Due to rapid growth in the

nonpermit-issuing universe the sampling methods did not adequately represent the new construction outside the Houston central city. Because of this deficiency, separate factors were computed for the following cells:

1. Conventional housing units (non-mobile home or trailer) outside the central city, which were built between the 1970 census and the 1976 survey.
2. Mobile homes and trailers outside the central city with a model year between the 1970 census and 1976 survey.
3. Conventional housing units outside the central city built between the 1976 and 1979 surveys.
4. Mobile homes and trailers outside the central city with a model year between the 1976 and 1979 surveys.

In a specific cell, the ratio estimation factor was equal to the following:

$$\frac{\text{Weighted count of housing units in the cell} + \text{Estimate of the undercoverage in the cell}}{\text{Weighted count of housing units in the cell}}$$

The computed second-stage ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell. For the Baltimore, Md.; Buffalo, N.Y.; Chicago, Ill.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; and Seattle-Everett, Wash.; SMSA's, a second-stage ratio estimation procedure was employed to adjust the central city/balance distribution of the weighted sample estimate of new construction housing units built since the last survey in permit-issuing areas to an independently derived estimate of this distribution.

This ratio estimation factor was calculated separately for the central city and balance of each SMSA and was applied to all new construction housing units from permit-issuing areas within the corresponding sector (central city or balance of the SMSA). This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}{\text{(Sample estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}$$

The independent estimates of new construction were based upon the number of authorized building permits which were determined from the Survey of Construction (SOC). The sample estimates were obtained from the weighted estimate of the AHS-SMSA sample housing units after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

This second-stage ratio estimation procedure was not implemented in the Honolulu, Hawaii; Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's. In the Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, the central city permit offices, sources of SOC information on permits, had expanded their coverage of permits to include areas greater than those covered by the 1970 census central city definitions. Since the AHS-SMSA survey uses the 1970 definitions, use of this ratio estimation procedure would not correctly adjust the central city/balance distribution of new construction housing units. In the Honolulu, Hawaii, SMSA, this procedure could not be used because the permit office covers the entire SMSA and does not differentiate between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Neb.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and as a third-stage ratio estimation procedure for the remaining 11 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the October 1979 housing inventory of each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sectors. This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the October 1979 housing inventory for the corresponding sector of the SMSA)}}{\text{(AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA)}}$$

The numerator of this ratio was derived using 1970 and 1980 census counts. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight after the first-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and the existing weight after the second-stage ratio estimation procedure for the 11 remaining SMSA's.

The computed ratio estimation factors for the central city and balance of the SMSA's were then applied to the existing weight for all corresponding sample units and the resulting product was used as the final weight for tabulation purposes.

The effect of the total housing unit ratio estimation procedure, as well as the overall 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 SMSA as a whole, it can be expected that the sample estimates will be improved when the sample housing population, or different portions of it, are brought into agreement with known good estimates of the SMSA housing population.

1976-1979 lost housing units—The 1976-1979 lost housing unit (housing unit removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1976 housing inventory, as was described in the 1976 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1976-1979 lost housing units existed, by definition, in the 1976 housing inventory, there was a 1976 housing inventory weight associated with each 1976-1979 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1976-1979 lost housing units.

1976 estimation procedure—This report presents data on the housing characteristics of the 1976 housing inventory from the 1976 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1976.

Ratio estimation procedure of the 1970 Census of Population and Housing—This report presents data on the housing characteristics of the 1970 housing inventory from the 1970 Census of Population and Housing. The statistics based on 1970 census sample data employed a ratio estimation procedure which was applied separately for each of the three census samples. A detailed description of this ratio estimation procedure can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

RELIABILITY OF THE ESTIMATES

There are two types of possible errors associated with estimates based on data from sample surveys—sampling and nonsampling errors. The following is a description of the sampling and nonsampling errors associated with the AHS-SMSA sample and of the nonsampling errors associated with the 1970 census estimates. A description of the sampling errors associated with the sample estimates from the 1970 census can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

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; and other errors of collection, response, processing, 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 1970 Census of Population and Housing and the 1976 AHS-SMSA sample.

1970 census—A number of studies were conducted to measure two types of general errors associated with 1970 census estimates—"coverage" and "content" errors. The "coverage" errors determined how completely housing units were counted in the census and the extent to which occupancy status was erroneously reported. The "content" errors measured the accuracy of the data collected for enumerated housing units. These errors were measured by reinterviews, record checks, and other surveys.

The detailed results of these studies, as well as the methodology employed, are presented in the 1970 Census of Population and Housing Evaluation and Research Program Reports, Series PHC(E)-5, *The Coverage of Housing in the 1970 Census,* and PHC(E)-10, *Accuracy of Data for Selected Housing Characteristics as Measured by Reinterviews.*

AHS-SMSA—Results from the 1979 AHS-SMSA sample reinterview program were not available at the time this report was being prepared. However, a study was conducted for the 1976 AHS-SMSA sample. The results are presented in the Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1976."

Coverage errors—In errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional (non-mobile home or trailer) new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units. Due to 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, in which case, 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, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit.

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 estimated that the 1976 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 for the 1979 survey, the number of missed housing units may be considerably less for 1979.

The final ratio estimation procedure corrects for these deficiencies as far as the count of total housing is concerned; i.e., it adjusts the estimate of the total housing inventory to the best available estimate. However, biases of subtotals would still remain.

Rounding errors—For errors associated with processing, the rounding of estimates introduces another source of error in the data, the severity of which depends on 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-SMSA sample—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 each other. 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 due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by 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, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. 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.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two 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 (page App-50) are approximations to the standard errors of various estimates shown in this report for this SMSA. 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. Standard errors applicable to estimates of characteristics of the 1976 housing inventory can be found in the AHS Series H-170 reports for 1976.

Table I (page App-50) presents the standard errors applicable to estimates of characteristics of the 1979 housing inventory as well as estimates of characteristics of the 1976-1979 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 7,810 for the total SMSA, 5,000 for the central city of the SMSA, and 5,890 for the balance of the SMSA.

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 (page App-50) presents the standard errors of estimated percentages for the 1979 housing inventory as well as estimated percentages of the 1976-1979 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

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 as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is 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:

$$(100)\left(\frac{x}{y}\right)\sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}$$

where: x = the numerator of the ratio
y = the denominator of the ratio
σ_x = the standard error of the numerator
σ_y = the standard error of the denominator

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1979 there were 104,200 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 1,270. 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
100,000	1,280
104,200	x
119,800	1,230

The entry for "x" is determined as follows by vertically interpolating between 1,280 and 1,230.

$$104{,}200-100{,}000 = 4{,}200$$
$$119{,}800-100{,}000 = 19{,}800$$

$$1{,}280 + \frac{4{,}200}{19{,}800}(1{,}230-1{,}280) = 1{,}270$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 102,930 to 105,470 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1979 owner-occupied housing units lies within a range computed in this way, would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 102,170 to 106,230 housing units with 90 percent confidence; and that the average estimate lies within the interval from 101,660 to 106,740 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 104,200 owner-occupied housing units, 20,600, or 19.8 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 19.8 percent is approximately 0.7 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	19.8	25 or 75
100,000.	0.6	a	0.8
104,200.		p	
150,000.	0.5	b	0.7

1. The entry for cell "a" is determined by horizontal interpolation between 0.6 and 0.8.

$$19.8-10.0 = 9.8$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{9.8}{15.0}(0.8-0.6) = 0.7$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.5 and 0.7.

$$19.8-10.0 = 9.8$$
$$25.0-10.0 = 15.0$$

$$0.5 + \frac{9.8}{15.0}(0.7-0.5) = 0.6$$

3. The entry for "p" is then determined by vertical interpolation between 0.7 and 0.6.

$$104{,}200-100{,}000 = 4{,}200$$
$$150{,}000-100{,}000 = 50{,}000$$

$$0.7 + \frac{4{,}200}{50{,}000}(0.6-0.7) = 0.7$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 19.1 to 20.5 percent; the 90-percent confidence interval is from 18.7 to 20.9 percent; and the 95-percent confidence interval is from 18.4 to 21.2 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. 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. Due to the overlap of the 1976 and 1979 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1979 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1979 there were 53,700 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 33,100. Table I shows the standard error of 53,700 is approximately 1,160, and the standard error of 20,600 is approximately 790. Therefore, the standard error of the estimated difference of 33,100 is about 1,400.

$$1{,}400 = \sqrt{(1{,}160)^2 + (790)^2}$$

Consequently, the 68-percent confidence interval for the 33,100 difference is from 31,700 to 34,500 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 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 30,860 to 35,340 housing units, and the 95-percent confidence interval is from 30,300 to 35,900 housing units. Thus, we can conclude with 95 percent confidence that the number of 1979 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied housing units with two bedrooms, since the 95-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 deter-

mine an interval about the estimated median so that there is a stated degree of confidence 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. These two distribution intervals could be different.

For about 68 out of 100 possible samples, the average median from all possible samples would lie between these two values.

A two-standard-error confidence interval may be determined by finding the values corresponding to 50 percent plus and minus twice the standard error determined in step 1. For about 95 out of 100 possible samples, the average median from all possible samples would lie between these two values.

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.7. The base of the distribution from which this median was determined is 104,200 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 104,200 is approximately 0.9 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.2 and 51.8.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with 3 persons (for purposes of calculating the median, the category of 3 persons is considered to be from 2.5 to 3.5 persons) contains the 48.2 percent derived in step 2. About 48,600 housing units or 46.6 percent fall below this interval, and 20,400 housing units or 19.6 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(48.2-46.6)}{19.6} = 2.6$$

Similarly, the interval for owner-occupied housing units with 3 persons contains the 51.8 percent derived in step 2. About 48,600 housing units or 46.6 percent fall below this interval, and 20,400 housing units or 19.6 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(51.2-46.6)}{19.6} = 2.7$$

Thus, the 95-percent confidence interval ranges from 2.6 to 2.7 persons. Although it appears that this confidence interval has the sample estimate as the upper limit, it actually is a reflection of the rounding error associated with this median (see the paragraph on rounding errors in the nonsampling errors section of this appendix).

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1979 Housing Inventory and for Estimated Number of 1976-1979 Lost Units for the Las Vegas, Nev., SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city
0	40	30	30
100	60	60	60
200	80	80	80
500	130	130	130
700	160	150	160
1,000	190	180	190
2,500	300	280	290
5,000	420	380	410
10,000	580	520	570
25,000	880	700	830
50,000	1,140	620	1,010
65,900	1,230	—	1,020
75,000	1,260	—	990
100,000	1,280	—	760
119,800	1,230	—	—
150,000	1,010	—	—
185,700	—	—	—

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.1. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero since these estimates were derived from census data which are not subject to sampling error. However, these estimates are subject to the nonsampling errors associated with the 1970 and 1980 censuses and with the interpolation procedures used to derive these estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1979 Housing Inventory and for Estimated Percentages of 1976-1979 Lost Housing Units for the Las Vegas, Nev., SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
200	15.1	15.1	15.1	15.1	18.2	21.1
500	6.6	6.6	6.6	8.0	11.5	13.3
700	4.8	4.8	4.9	6.8	9.8	11.3
1,000	3.4	3.4	4.1	5.7	8.2	9.4
2,500	1.4	1.4	2.6	3.6	5.2	6.0
5,000	0.7	0.8	1.8	2.5	3.6	4.2
10,000	0.4	0.6	1.3	1.8	2.6	3.0
25,000	0.14	0.4	0.8	1.1	1.6	1.9
50,000	0.07	0.3	0.6	0.8	1.2	1.3
75,000	0.05	0.2	0.5	0.7	0.9	1.1
100,000	0.04	0.2	0.4	0.6	0.8	0.9
150,000	0.02	0.15	0.3	0.5	0.7	0.8
200,000	0.02	0.13	0.3	0.4	0.6	0.7

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than 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.

Miami 1979

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-41
Annual Housing Survey . . . App-41
Designation of sample housing units for the 1979 survey . . App-41
Selection of the 1975 AHS-SMSA sample . . . App-41
1975-1979 additions to the housing inventory . . . App-43
Sample selection for the 1979 Coverage Improvement Program . App-43
Coverage improvement for deficiency 1 . . . App-43
Coverage improvement for deficiency 2 . . . App-43
Coverage improvement for deficiencies 3-6 . . . App-43
1970 Census of Population and Housing . . . App-44
ESTIMATION . . . App-44
1979 housing inventory . . . App-44
1975-1979 lost housing units . . . App-45
1975 estimation procedure . . . App-46
Ratio estimation procedure of the 1970 Census of Population and Housing . . . App-46
RELIABILITY OF THE ESTIMATES . . . App-46
Nonsampling errors . . . App-46
1970 census . . . App-46
AHS-SMSA . . . App-46
Coverage errors . . . App-46
Rounding errors . . . App-46
Sampling errors for the AHS-SMSA sample . . . App-47
Illustration of the use of the standard error tables . . . App-47
Differences . . . App-48
Illustration of the computation of the standard error of a difference . . . App-48
Medians . . . App-48
Illustration of the computation of the 95-percent confidence interval of a median . . . App-49
Standard error tables . . . App-50

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 15 SMSA's in this report series (H-170-79) are based on data collected from the 1979 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 15 SMSA's, the data were collected during the 12-month period from April 1979 through March 1980 with one-twelfth of the sample housing units being visited each month.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1979 were interviewed previously in either 1975 or 1976 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1979 group of SMSA's, the largest SMSA from three of the four census regions of the United States is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All the remaining 12 SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the 1979 group are: Chicago, Ill.; Houston, Tex.; and Seattle-Everett, Wash.

The remaining SMSA's in the 1979 group are: Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.

In this SMSA, 5,052 housing units were eligible for interview. Of these sample housing units, 174 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 5,052 housing units eligible for interview, 523 units were visited but were not eligible for interview because they were condemned, unfit, demolished, converted to group quarters use, etc.

Designation of sample housing units for the 1979 survey—The sample housing units designated to be interviewed in the 1979 survey consisted of the following categories which are described in detail in the succeeding sections.

1. All sample housing units that were interviewed in the 1975 survey.
2. 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 1975 survey. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1979 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1975 survey. (This sample represented the housing units built in permit-issuing areas, since the 1975 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1975 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1975 survey.)
5. All sample housing units that were selected as part of the 1979 Coverage Improvement Program. (This sample represented most of the housing units which, until 1979, did not have a chance of selection.)

Selection of the 1975 AHS-SMSA sample—In 1970, the sample for the SMSA's which were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in

the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). In 1979, the following four SMSA's were 100-percent permit-issuing: Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; and Miami, Fla. The remaining 11 SMSA's contain a sample from the nonpermit universe.

Sampling operations, described in the following paragraphs, were performed separately within the central city and the balance of the SMSA for each of the sample frames. The overall sampling rate used to select the sample for each SMSA was determined by the size of the sample. Thus, for the three largest SMSA's, the overall sampling rate differed by central city and the balance of the SMSA, since the sample was divided equally between the central city and the balance of the SMSA. The remaining SMSA's had an overall sampling rate about the same for the sample selected from both the central city and the balance of the SMSA, since the sample was distributed proportionately between the central city and the balance of the SMSA according to the distribution of the total housing units in each sector.

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA 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 housing unit records were stratified by race of head (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 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household 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, for this SMSA, 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 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. 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 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 SMSA. 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.

The second frame from which this SMSA 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 this SMSA. Prior to sample selection, the list of permits was chronologically 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.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS 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 (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

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 are included.

1975-1979 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1975 survey, was selected to represent housing units built in permit-issuing areas since the 1975 survey. Sampling procedures were identical to those used in selecting the 1970-1975 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1975, to identify any housing units missed in the 1975 survey or any housing units added since the 1975 survey.

Sample selection for the 1979 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

1. New construction housing units from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes and trailers placed in parks either missed during the 1970 census or established since the 1970 census.
3. Housing units missed in the 1970 census.
4. Housing units converted to residential use that were nonresidential at the time of the 1970 census.
5. Housing units that have been moved onto their present site since the 1970 census.
6. Mobile homes and trailers placed outside parks since the 1970 census or vacant at the time of the 1970 census.

For each of the five SMSA's interviewed for the first time in 1975, a full implementation of the Coverage Improvement Program was conducted as a part of the 1979 AHS. For each of the 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with some updating and refining as a part of the 1979 AHS. The following discussion of coverage improvement procedures applies to both the 1976 and 1979 programs. For the 10 (1976) SMSA's, estimates of housing units added by a specific procedure reflect units added in 1976 as well as any additions that resulted from the updating and refining in 1979.

Coverage improvement for deficiency 1—A sample of new construction housing units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 1979 SMSA's. Two different procedures were used. For the SMSA's first interviewed in 1976, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. These new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units).

The first stage sample selection was a sample of permit offices, and the second stage a sample of the 1969 permits within each of the selected permit offices. (In the Honolulu, Hawaii, SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second stage.) For the third stage, structures of size three or more were divided into clusters of an expected size of four housing units and a sample of clusters was selected.

For the SMSA's first interviewed in 1975, housing units whose permits were issued before January 1970, but which were completed after April 1970, were identified from the Survey of Construction (SOC), a survey of building permits conducted monthly by the Bureau of the Census. These housing units were then sampled at one-third the rate of regular AHS units. These procedures added an estimated 18,492 new construction housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiency 2—In permit-issuing areas, a sample of mobile homes and trailers placed in parks that were missed by the census or established after the census was selected in two stages. First, a sample of tracts was selected and canvassed. All parks were listed and then matched back to the 1970 census to identify parks missed by the census and parks established after the census. Second, the parks were divided into clusters of an expected size of four sites and a sample of clusters was selected and interviewed. Each of the sample housing units represented the same number of units that the regular AHS sample housing units represented. Since research showed that there were no mobile home or trailer parks in the central cities of the Chicago, Ill.; Hartford, Conn.; and Portland, Oreg.-Wash.; SMSA's, this procedure was only implemented outside the central cities for these SMSA's. Based on a cost-benefit analysis, this procedure was updated as part of the 1979 AHS only for the Denver, Colo. and Las Vegas, Nev., SMSA's. This procedure added 1,468 housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed housing units were sampled by one of two procedures. The first procedure was designed to represent units from the following types of missed structures (structures that had no chance of selection for the AHS):

1. Structures missed in the 1970 census.
2. Structures that were completely nonresidential in the 1970 census but now contain residential housing units.
3. Mobile homes and trailers that had been placed outside parks since the 1970 census and have a utility hookup, or were on the site during the present survey but not occupied on April 1, 1970, or had no utility hookup but were occupied by persons with no usual residence elsewhere.
4. Housing units that had been moved onto their present site since the 1970 census.

Initially, a subsample of regular AHS units was selected at a rate of 1 in 24 for the Chicago, Ill.; Hartford, Conn.; Miami, Fla.; Milwaukee, Wis.; and Portland, Oreg.-Wash.; SMSA's, and at a rate of 1 in 22.2152 for the 10 remaining SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing units were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and housing units within

these structures were interviewed. In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. Although the Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Houston, Tex.; Omaha, Nebr.-Iowa; and Raleigh, N.C.; SMSA's, were interviewed for the first time in 1976 and had these sampling procedures implemented at that time, some 1976 coverage improvement assignments were not sent out to be interviewed due to time limitations. The sampling was completed as part of the 1979 AHS and this procedure added an estimated 14,654 housing units to the coverage of the housing inventory of this SMSA.

The second procedure was designed to represent missed housing units from structures represented in the AHS. These missed housing units were:

1. Housing units missed in the 1970 census.
2. Nonresidential space converted to residential use since the 1970 census in structures that contained some residential housing units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structures selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 1,975 housing units to the coverage of the housing inventory for this SMSA.

1970 Census of Population and Housing—The estimates pertaining to the 1970 housing inventory (i.e., the housing inventory that existed at the time of the 1970 census) are based on either 20-, 15-, or 5-percent sample data collected in April 1970 for the Decennial Census of Population and Housing. A detailed description of the sample design employed for the 1970 census can be obtained in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties*, Part 1.

ESTIMATION

The 1979 AHS sample produced two types of estimates for each SMSA: Estimates pertaining to characteristics of the housing inventory at the time of the interview (i.e., the 1979 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1975 (i.e., 1975-1979 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1979 housing inventory—The AHS estimates of characteristics of the 1979 housing inventory were produced using a 2-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and a 3-stage ratio estimation procedure for the remaining 11 SMSA's. Prior to 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 the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of one or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes and trailers from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample housing units from both the nonpermit universe and the coverage improvement universe (if units were not included in any of the previous cells).

The following first-stage 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 permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of housing units enumerated in areas under the jurisdiction of permit-issuing offices. The denominators of the ratios were obtained from weighted estimates of all the AHS sample housing units within the corresponding ratio estimation categories using the existing weight (i.e., the basic weight times the noninterview factor). The computed ratio estimation factor was then applied to the existing weight for each sample housing unit within the corresponding ratio estimation category.

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 permit-issuing universe. Prior to the AHS sample selection within each SMSA, 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 sample. Since the number of housing units deleted from the AHS universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the Houston, Tex., SMSA, a second ratio estimation procedure was employed for new construction housing units outside the Houston central city. Due to rapid growth in the

nonpermit-issuing universe the sampling methods did not adequately represent the new construction outside the Houston central city. Because of this deficiency, separate factors were computed for the following cells:

1. Conventional housing units (non-mobile home or trailer) outside the central city, which were built between the 1970 census and the 1976 survey.
2. Mobile homes and trailers outside the central city with a model year between the 1970 census and 1976 survey.
3. Conventional housing units outside the central city built between the 1976 and 1979 surveys.
4. Mobile homes and trailers outside the central city with a model year between the 1976 and 1979 surveys.

In a specific cell, the ratio estimation factor was equal to the following:

$$\frac{\text{Weighted count of housing units in the cell} + \text{Estimate of the undercoverage in the cell}}{\text{Weighted count of housing units in the cell}}$$

The computed second-stage ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell. For the Baltimore, Md.; Buffalo, N.Y.; Chicago, Ill.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; and Seattle-Everett, Wash.; SMSA's, a second-stage ratio estimation procedure was employed to adjust the central city/balance distribution of the weighted sample estimate of new construction housing units built since the last survey in permit-issuing areas to an independently derived estimate of this distribution.

This ratio estimation factor was calculated separately for the central city and balance of each SMSA and was applied to all new construction housing units from permit-issuing areas within the corresponding sector (central city or balance of the SMSA). This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}{\text{(Sample estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}$$

The independent estimates of new construction were based upon the number of authorized building permits which were determined from the Survey of Construction (SOC). The sample estimates were obtained from the weighted estimate of the AHS-SMSA sample housing units after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

This second-stage ratio estimation procedure was not implemented in the Honolulu, Hawaii; Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's. In the Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C., SMSA's, the central city permit offices, sources of SOC information on permits, had expanded their coverage of permits to include areas greater than those covered by the 1970 census central city definitions. Since the AHS-SMSA survey uses the 1970 definitions, use of this ratio estimation procedure would not correctly adjust the central city/balance distribution of new construction housing units. In the Honolulu, Hawaii, SMSA, this procedure could not be used because the permit office covers the entire SMSA and does not differentiate between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Neb.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and as a third-stage ratio estimation procedure for the remaining 11 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the October 1979 housing inventory of each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sectors. This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the October 1979 housing inventory for the corresponding sector of the SMSA)}}{\text{(AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA)}}$$

The numerator of this ratio was derived using 1970 and 1980 census counts. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight after the first-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and the existing weight after the second-stage ratio estimation procedure for the 11 remaining SMSA's.

The computed ratio estimation factors for the central city and balance of the SMSA's were then applied to the existing weight for all corresponding sample units and the resulting product was used as the final weight for tabulation purposes.

The effect of the total housing unit ratio estimation procedure, as well as the overall 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 SMSA as a whole, it can be expected that the sample estimates will be improved when the sample housing population, or different portions of it, are brought into agreement with known good estimates of the SMSA housing population.

1975-1979 lost housing units—The 1975-1979 lost housing unit (housing unit removed from the inventory) estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1975 housing inventory, as was described in the 1975 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas*. Since the 1975-1979 lost housing units existed, by definition, in the 1975 housing inventory, there was a 1975 housing inventory weight associated with each 1975-1979 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1975-1979 lost housing units.

1975 estimation procedure—This report presents data on the housing characteristics of the 1975 housing inventory from the 1975 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1975.

Ratio estimation procedure of the 1970 Census of Population and Housing—This report presents data on the housing characteristics of the 1970 housing inventory from the 1970 Census of Population and Housing. The statistics based on 1970 census sample data employed a ratio estimation procedure which was applied separately for each of the three census samples. A detailed description of this ratio estimation procedure can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

RELIABILITY OF THE ESTIMATES

There are two types of possible errors associated with estimates based on data from sample surveys—sampling and nonsampling errors. The following is a description of the sampling and nonsampling errors associated with the AHS-SMSA sample and of the nonsampling errors associated with the 1970 census estimates. A description of the sampling errors associated with the sample estimates from the 1970 census can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

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; and other errors of collection, response, processing, 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 1970 Census of Population and Housing and the 1975 AHS-SMSA sample.

1970 census—A number of studies were conducted to measure two types of general errors associated with 1970 census estimates—"coverage" and "content" errors. The "coverage" errors determined how completely housing units were counted in the census and the extent to which occupancy status was erroneously reported. The "content" errors measured the accuracy of the data collected for enumerated housing units. These errors were measured by reinterviews, record checks, and other surveys.

The detailed results of these studies, as well as the methodology employed, are presented in the 1970 Census of Population and Housing Evaluation and Research Program Reports, Series PHC(E)-5, *The Coverage of Housing in the 1970 Census,* and PHC(E)-10, *Accuracy of Data for Selected Housing Characteristics as Measured by Reinterviews.*

AHS-SMSA—Results from the 1979 AHS-SMSA sample reinterview program were not available at the time this report was being prepared. However, a study was conducted for the 1975 AHS-SMSA sample. The results are presented in the Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1975."

Coverage errors—In errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional (non-mobile home or trailer) new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units. Due to 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, in which case, 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, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit.

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 estimated that the 1975 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 for the 1979 survey, the number of missed housing units may be considerably less for 1979.

The final ratio estimation procedure corrects for these deficiencies as far as the count of total housing is concerned; i.e., it adjusts the estimate of the total housing inventory to the best available estimate. However, biases of subtotals would still remain.

Rounding errors—For errors associated with processing, the rounding of estimates introduces another source of error in the data, the severity of which depends on 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-SMSA sample—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 each other. 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 due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by 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, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. 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.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two 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 (page App-50) are approximations to the standard errors of various estimates shown in this report for this SMSA. 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. Standard errors applicable to estimates of characteristics of the 1975 housing inventory can be found in the AHS Series H-170 reports for 1975.

Table I (page App-50) presents the standard errors applicable to estimates of characteristics of the 1979 housing inventory as well as estimates of characteristics of the 1975-1979 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 25,780 for the total SMSA, 10,610 for the central city of the SMSA, and 23,400 for the balance of the SMSA.

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 (page App-50) presents the standard errors of estimated percentages for the 1979 housing inventory as well as estimated percentages of the 1975-1979 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

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 as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is 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:

$$(100)\left(\frac{x}{y}\right)\sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}$$

where: x = the numerator of the ratio
y = the denominator of the ratio
σ_x = the standard error of the numerator
σ_y = the standard error of the denominator

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1979 there were 319,400 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 4,960. 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
300,000	4,980
319,400	x
400,000	4,870

The entry for "x" is determined as follows by vertically interpolating between 4,980 and 4,870.

$$319{,}400-300{,}000 = 19{,}400$$
$$400{,}000-300{,}000 = 100{,}000$$

$$4{,}980 + \frac{19{,}400}{100{,}000}(4{,}870-4{,}980) = 4{,}960$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 314,440 to 324,360 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1979 owner-occupied housing units lies within a range computed in this way, would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 311,460 to 327,340 housing units with 90 percent confidence; and that the average estimate lies within the interval from 309,480 to 329,320 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 319,400 owner-occupied housing untis, 98,900, or 31.0 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 31.0 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. The entry for "p" is the one sought.

Base of percentage	Estimated percentage		
	25 or 75	31.0	50
300,000.	1.0	a	1.1
319,400.		p	
400,000.	0.8	b	1.0

1. The entry for cell "a" is determined by horizontal interpolation between 1.0 and 1.1.

$$31.0-25.0 = 6.0$$
$$50.0-25.0 = 25.0$$

$$1.0 + \frac{6.0}{25.0}(1.1-1.0) = 1.0$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.8 and 1.0.

$$31.0-25.0 = 6.0$$
$$50.0-25.0 = 25.0$$

$$0.8 + \frac{6.0}{25.0}(1.0-0.8) = 0.8$$

3. The entry for "p" is then determined by vertical interpolation between 1.0 and 0.8.

$$319{,}400-300{,}000 = 19{,}400$$
$$400{,}000-300{,}000 = 100{,}000$$

$$1.0 + \frac{19{,}400}{100{,}000}(0.8-1.0) = 1.0$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 30.0 to 32.0 percent; the 90-percent confidence interval is from 29.4 to 32.6 percent; and the 95-percent confidence interval is from 29.0 to 33.0 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. 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. Due to the overlap of the 1975 and 1979 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1975 and 1979 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1979 there were 145,200 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 46,300. Table I shows the standard error of 145,200 is approximately 4,150, and the standard error of 98,900 is approximately 3,580. Therefore, the standard error of the estimated difference of 46,300 is about 5,480.

$$5{,}480 = \sqrt{(4{,}150)^2 + (3{,}580)^2}$$

Consequently, the 68-percent confidence interval for the 46,300 difference is from 40,820 to 51,780 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 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 37,530 to 55,070 housing units, and the 95-percent confidence interval is from 35,340 to 57,260 housing units. Thus, we can conclude with 95 percent confidence that the number of 1979 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied housing units with two bedrooms, since the 95-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 deter-

mine an interval about the estimated median so that there is a stated degree of confidence that the average median from all possible sample 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. These two distribution intervals could be different.

For about 68 out of 100 possible samples, the average median from all possible samples would lie between these two values.

A two-standard-error confidence interval may be determined by finding the values corresponding to 50 percent plus and minus twice the standard error determined in step 1. For about 95 out of 100 possible samples, the average median from all possible samples would lie between these two values.

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.5. The base of the distribution from which this median was determined is 319,400 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 319,400 is approximately 1.1 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 47.8 and 52.2.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with 2 persons (for purposes of calculating the median, the category of 2 persons is considered to be from 1.5 to 2.5 persons) contains the 47.8 percent derived in step 2. About 52,400 housing units or 16.4 percent fall below this interval, and 104,500 housing units or 32.7 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$1.5 + (2.5-1.5)\,\frac{(47.8-16.4)}{32.7} = 2.5$$

Similarly, the interval for owner-occupied housing units with 3 persons contains the 52.2 percent derived in step 2. About 156,900 housing units or 49.1 percent fall below this interval, and 58,300 housing units or 18.3 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(52.2-49.1)}{18.3} = 2.7$$

Thus, the 95-percent confidence interval ranges from 2.5 to 2.7 persons. Although it appears that this confidence interval has the sample estimate as the lower limit, it actually is a reflection of the rounding error associated with the median (see the paragraph on rounding errors in the nonsampling errors section of this appendix).

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1979 Housing Inventory and for Estimated Number of 1975-1979 Lost Units for the Miami, Fla., SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city
0	150	150	150
100	150	150	150
200	170	170	170
500	280	270	280
700	330	320	330
1,000	390	380	390
2,500	620	600	620
5,000	870	840	870
10,000	1,230	1,160	1,220
25,000	1,920	1,730	1,900
50,000	2,650	2,180	2,620
75,000	3,180	2,290	3,120
100,000	3,600	2,120	3,500
144,700	4,150	—	3,970
150,000	4,200	—	4,020
200,000	4,600	—	4,300
250,000	4,860	—	4,410
300,000	4,980	—	4,340
400,000	4,870	—	3,630
500,000	4,250	—	1,230
510,100	4,150	—	—
600,000	2,770	—	—
654,800	—	—	—

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.3 for the total SMSA, 1.2 for the central city, and 1.2 for the balance (not in central city) estimates. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero since these estimates were derived from census data which are not subject to sampling error. However, these estimates are subject to the nonsampling errors associated with 1970 and 1980 censuses and with the interpolation procedures used to derive these estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1979 Housing Inventory and for Estimated Percentages of 1975-1979 Lost Housing Units for the Miami, Fla., SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
200	43.3	43.3	43.3	43.3	43.3	43.7
500	23.4	23.4	23.4	23.4	23.9	27.6
700	17.9	17.9	17.9	17.9	20.2	23.3
1,000	13.2	13.2	13.2	13.2	16.9	19.5
2,500	5.8	5.8	5.8	7.4	10.7	12.4
5,000	3.0	3.0	3.8	5.2	7.6	8.7
10,000	1.5	1.5	2.7	3.7	5.3	6.2
25,000	0.6	0.8	1.7	2.3	3.4	3.9
50,000	0.3	0.5	1.2	1.7	2.4	2.8
75,000	0.2	0.4	1.0	1.4	2.0	2.3
100,000	0.15	0.4	0.9	1.2	1.7	2.0
150,000	0.10	0.3	0.7	1.0	1.4	1.6
200,000	0.08	0.3	0.6	0.8	1.2	1.4
250,000	0.06	0.2	0.5	0.7	1.1	1.2
300,000	0.05	0.2	0.5	0.7	1.0	1.1
400,000	0.04	0.2	0.4	0.6	0.8	1.0
500,000	0.03	0.2	0.4	0.5	0.8	0.9
600,000	0.03	0.2	0.3	0.5	0.7	0.8
700,000	0.02	0.15	0.3	0.4	0.6	0.7

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than 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 1.3 for the total SMSA, 1.2 for the central city, and 1.2 for the balance (not in central city).

Milwaukee 1979

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-41
Annual Housing Survey App-41
Designation of sample housing units for the 1979 survey. App-41
Selection of the 1975 AHS-SMSA sample App-41
1975-1979 additions to the housing inventory App-43
Sample selection for the 1979 Coverage Improvement Program. App-43
Coverage improvement for deficiency 1 . . . . . App-43
Coverage improvement for deficiency 2 . . . . . App-43
Coverage improvement for deficiencies 3-6 . . . . . App-43
1970 Census of Population and Housing. App-44
ESTIMATION. App-44
1979 housing inventory App-44
1975-1979 lost housing units . . . App-45
1975 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES. App-46
Nonsampling errors. App-46
1970 census App-46
AHS-SMSA App-46
Coverage errors App-46
Rounding errors. App-46
Sampling errors for the AHS-SMSA sample App-47
Illustration of the use of the standard error tables App-47
Differences. App-48
Illustration of the computation of the standard error of a difference App-48
Medians App-48
Illustration of the computation of the 95-percent confidence interval of a median App-49
Standard error tables App-50

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 15 SMSA's in this report series (H-170-79) are based on data collected from the 1979 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 15 SMSA's, the data were collected during the 12-month period from April 1979 through March 1980 with one-twelfth of the sample housing units being visited each month.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1979 were interviewed previously in either 1975 or 1976 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1979 group of SMSA's, the largest SMSA from three of the four census regions of the United States is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All the remaining 12 SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the 1979 group are: Chicago, Ill.; Houston, Tex.; and Seattle-Everett, Wash.

The remaining SMSA's in the 1979 group are: Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.

In this SMSA, 5,289 housing units were eligible for interview. Of these sample housing units, 123 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 5,289 housing units eligible for interview, 252 units were visited but were not eligible for interview because they were condemned, unfit, demolished, converted to group quarters use, etc.

Designation of sample housing units for the 1979 survey—The sample housing units designated to be interviewed in the 1979 survey consisted of the following categories which are described in detail in the succeeding sections.

1. All sample housing units that were interviewed in the 1975 survey.
2. 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 1975 survey. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1979 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1975 survey. (This sample represented the housing units built in permit-issuing areas, since the 1975 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1975 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1975 survey.)
5. All sample housing units that were selected as part of the 1979 Coverage Improvement Program. (This sample represented most of the housing units which, until 1979, did not have a chance of selection.)

Selection of the 1975 AHS-SMSA sample—In 1970, the sample for the SMSA's which were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in

the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). In 1979, the following four SMSA's were 100-percent permit-issuing: Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; and Miami, Fla. The remaining 11 SMSA's contain a sample from the nonpermit universe.

Sampling operations, described in the following paragraphs, were performed separately within the central city and the balance of the SMSA for each of the sample frames. The overall sampling rate used to select the sample for each SMSA was determined by the size of the sample. Thus, for the three largest SMSA's, the overall sampling rate differed by central city and the balance of the SMSA, since the sample was divided equally between the central city and the balance of the SMSA. The remaining SMSA's had an overall sampling rate about the same for the sample selected from both the central city and the balance of the SMSA, since the sample was distributed proportionately between the central city and the balance of the SMSA according to the distribution of the total housing units in each sector.

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA 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 housing unit records were stratified by race of head (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 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household 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, for this SMSA, 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 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. 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 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 SMSA. 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.

The second frame from which this SMSA 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 this SMSA. Prior to sample selection, the list of permits was chronologically 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.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS 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 (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

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 are included.

1975-1979 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1975 survey, was selected to represent housing units built in permit-issuing areas since the 1975 survey. Sampling procedures were identical to those used in selecting the 1970-1975 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1975, to identify any housing units missed in the 1975 survey or any housing units added since the 1975 survey.

Sample selection for the 1979 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

1. New construction housing units from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes and trailers placed in parks either missed during the 1970 census or established since the 1970 census.
3. Housing units missed in the 1970 census.
4. Housing units converted to residential use that were nonresidential at the time of the 1970 census.
5. Housing units that have been moved onto their present site since the 1970 census.
6. Mobile homes and trailers placed outside parks since the 1970 census or vacant at the time of the 1970 census.

For each of the five SMSA's interviewed for the first time in 1975, a full implementation of the Coverage Improvement Program was conducted as a part of the 1979 AHS. For each of the 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with some updating and refining as a part of the 1979 AHS. The following discussion of coverage improvement procedures applies to both the 1976 and 1979 programs. For the 10 (1976) SMSA's, estimates of housing units added by a specific procedure reflect units added in 1976 as well as any additions that resulted from the updating and refining in 1979.

Coverage improvement for deficiency 1—A sample of new construction housing units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 1979 SMSA's. Two different procedures were used. For the SMSA's first interviewed in 1976, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. These new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units).

The first stage sample selection was a sample of permit offices, and the second stage a sample of the 1969 permits within each of the selected permit offices. (In the Honolulu, Hawaii, SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second stage.) For the third stage, structures of size three or more were divided into clusters of an expected size of four housing units and a sample of clusters was selected.

For the SMSA's first interviewed in 1975, housing units whose permits were issued before January 1970, but which were completed after April 1970, were identified from the Survey of Construction (SOC), a survey of building permits conducted monthly by the Bureau of the Census. These housing units were then sampled at one-third the rate of regular AHS units. These procedures added an estimated 4,077 new construction housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiency 2—In permit-issuing areas, a sample of mobile homes and trailers placed in parks that were missed by the census or established after the census was selected in two stages. First, a sample of tracts was selected and canvassed. All parks were listed and then matched back to the 1970 census to identify parks missed by the census and parks established after the census. Second, the parks were divided into clusters of an expected size of four sites and a sample of clusters was selected and interviewed. Each of the sample housing units represented the same number of units that the regular AHS sample housing units represented. Since research showed that there were no mobile home or trailer parks in the central cities of the Chicago, Ill.; Hartford, Conn.; and Portland, Oreg.-Wash.; SMSA's, this procedure was only implemented outside the central cities for these SMSA's. Based on a cost-benefit analysis, this procedure was updated as part of the 1979 AHS only for the Denver, Colo. and Las Vegas, Nev., SMSA's. This procedure added 1,160 housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed housing units were sampled by one of two procedures. The first procedure was designed to represent units from the following types of missed structures (structures that had no chance of selection for the AHS):

1. Structures missed in the 1970 census.
2. Structures that were completely nonresidential in the 1970 census but now contain residential housing units.
3. Mobile homes and trailers that had been placed outside parks since the 1970 census and have a utility hookup, or were on the site during the present survey but not occupied on April 1, 1970, or had no utility hookup but were occupied by persons with no usual residence elsewhere.
4. Housing units that had been moved onto their present site since the 1970 census.

Initially, a subsample of regular AHS units was selected at a rate of 1 in 24 for the Chicago, Ill.; Hartford, Conn.; Miami, Fla.; Milwaukee, Wis.; and Portland, Oreg.-Wash.; SMSA's, and at a rate of 1 in 22.2152 for the 10 remaining SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing units were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and housing units within these structures were interviewed. In cases where the interviewer

workload would have been too great, a representative subsample of units within these structures was selected. Although the Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Houston, Tex.; Omaha, Nebr.-Iowa; and Raleigh, N.C.; SMSA's, were interviewed for the first time in 1976 and had these sampling procedures implemented at that time, some 1976 coverage improvement assignments were not sent out to be interviewed due to time limitations. The sampling was completed as part of the 1979 AHS and this procedure added an estimated 4,474 housing units to the coverage of the housing inventory of this SMSA.

The second procedure was designed to represent missed housing units from structures represented in the AHS. These missed housing units were:

1. Housing units missed in the 1970 census.
2. Nonresidential space converted to residential use since the 1970 census in structures that contained some residential housing units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structures selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 650 housing units to the coverage of the housing inventory for this SMSA.

1970 Census of Population and Housing—The estimates pertaining to the 1970 housing inventory (i.e., the housing inventory that existed at the time of the 1970 census) are based on either 20-, 15-, or 5-percent sample data collected in April 1970 for the Decennial Census of Population and Housing. A detailed description of the sample design employed for the 1970 census can be obtained in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

ESTIMATION

The 1979 AHS sample produced two types of estimates for each SMSA: Estimates pertaining to characteristics of the housing inventory at the time of the interview (i.e., the 1979 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1975 (i.e., 1975-1979 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1979 housing inventory—The AHS estimates of characteristics of the 1979 housing inventory were produced using a 2-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and a 3-stage ratio estimation procedure for the remaining 11 SMSA's. Prior to 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 the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of one or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes and trailers from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample housing units from both the nonpermit universe and the coverage improvement universe (if units were not included in any of the previous cells).

The following first-stage 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 permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of housing units enumerated in areas under the jurisdiction of permit-issuing offices. The denominators of the ratios were obtained from weighted estimates of all the AHS sample housing units within the corresponding ratio estimation categories using the existing weight (i.e., the basic weight times the noninterview factor). The computed ratio estimation factor was then applied to the existing weight for each sample housing unit within the corresponding ratio estimation category.

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 permit-issuing universe. Prior to the AHS sample selection within each SMSA, 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 sample. Since the number of housing units deleted from the AHS universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the Houston, Tex., SMSA, a second ratio estimation procedure was employed for new construction housing units outside the Houston central city. Due to rapid growth in the

nonpermit-issuing universe the sampling methods did not adequately represent the new construction outside the Houston central city. Because of this deficiency, separate factors were computed for the following cells:

1. Conventional housing units (non-mobile home or trailer) outside the central city, which were built between the 1970 census and the 1976 survey.
2. Mobile homes and trailers outside the central city with a model year between the 1970 census and 1976 survey.
3. Conventional housing units outside the central city built between the 1976 and 1979 surveys.
4. Mobile homes and trailers outside the central city with a model year between the 1976 and 1979 surveys.

In a specific cell, the ratio estimation factor was equal to the following:

$$\frac{\text{Weighted count of housing units in the cell} + \text{Estimate of the undercoverage in the cell}}{\text{Weighted count of housing units in the cell}}$$

The computed second-stage ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell. For the Baltimore, Md.; Buffalo, N.Y.; Chicago, Ill.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; and Seattle-Everett, Wash.; SMSA's, a second-stage ratio estimation procedure was employed to adjust the central city/balance distribution of the weighted sample estimate of new construction housing units built since the last survey in permit-issuing areas to an independently derived estimate of this distribution.

This ratio estimation factor was calculated separately for the central city and balance of each SMSA and was applied to all new construction housing units from permit-issuing areas within the corresponding sector (central city or balance of the SMSA). This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}{\text{(Sample estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}$$

The independent estimates of new construction were based upon the number of authorized building permits which were determined from the Survey of Construction (SOC). The sample estimates were obtained from the weighted estimate of the AHS-SMSA sample housing units after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

This second-stage ratio estimation procedure was not implemented in the Honolulu, Hawaii; Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's. In the Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, the central city permit offices, sources of SOC information on permits, had expanded their coverage of permits to include areas greater than those covered by the 1970 census central city definitions. Since the AHS-SMSA survey uses the 1970 definitions, use of this ratio estimation procedure would not correctly adjust the central city/balance distribution of new construction housing units. In the Honolulu, Hawaii, SMSA, this procedure could not be used because the permit office covers the entire SMSA and does not differentiate between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Neb.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and as a third-stage ratio estimation procedure for the remaining 11 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the October 1979 housing inventory of each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sectors. This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the October 1979 housing inventory for the corresponding sector of the SMSA)}}{\text{(AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA)}}$$

The numerator of this ratio was derived using 1970 and 1980 census counts. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight after the first-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and the existing weight after the second-stage ratio estimation procedure for the 11 remaining SMSA's.

The computed ratio estimation factors for the central city and balance of the SMSA's were then applied to the existing weight for all corresponding sample units and the resulting product was used as the final weight for tabulation purposes.

The effect of the total housing unit ratio estimation procedure, as well as the overall 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 SMSA as a whole, it can be expected that the sample estimates will be improved when the sample housing population, or different portions of it, are brought into agreement with known good estimates of the SMSA housing population.

1975-1979 lost housing units—The 1975-1979 lost housing unit (housing unit removed from the inventory) estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1975 housing inventory, as was described in the 1975 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1975-1979 lost housing units existed, by definition, in the 1975 housing inventory, there was a 1975 housing inventory weight associated with each 1975-1979 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1975-1979 lost housing units.

1975 estimation procedure—This report presents data on the housing characteristics of the 1975 housing inventory from the 1975 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1975.

Ratio estimation procedure of the 1970 Census of Population and Housing—This report presents data on the housing characteristics of the 1970 housing inventory from the 1970 Census of Population and Housing. The statistics based on 1970 census sample data employed a ratio estimation procedure which was applied separately for each of the three census samples. A detailed description of this ratio estimation procedure can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

RELIABILITY OF THE ESTIMATES

There are two types of possible errors associated with estimates based on data from sample surveys—sampling and nonsampling errors. The following is a description of the sampling and nonsampling errors associated with the AHS-SMSA sample and of the nonsampling errors associated with the 1970 census estimates. A description of the sampling errors associated with the sample estimates from the 1970 census can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

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; and other errors of collection, response, processing, 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 1970 Census of Population and Housing and the 1975 AHS-SMSA sample.

1970 census—A number of studies were conducted to measure two types of general errors associated with 1970 census estimates—"coverage" and "content" errors. The "coverage" errors determined how completely housing units were counted in the census and the extent to which occupancy status was erroneously reported. The "content" errors measured the accuracy of the data collected for enumerated housing units. These errors were measured by reinterviews, record checks, and other surveys.

The detailed results of these studies, as well as the methodology employed, are presented in the 1970 Census of Population and Housing Evaluation and Research Program Reports, Series PHC(E)-5, *The Coverage of Housing in the 1970 Census,* and PHC(E)-10, *Accuracy of Data for Selected Housing Characteristics as Measured by Reinterviews.*

AHS-SMSA—Results from the 1979 AHS-SMSA sample reinterview program were not available at the time this report was being prepared. However, a study was conducted for the 1975 AHS-SMSA sample. The results are presented in the Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1975."

Coverage errors—In errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional (non-mobile home or trailer) new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units. Due to 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, in which case, 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, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit.

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 estimated that the 1975 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 for the 1979 survey, the number of missed housing units may be considerably less for 1979.

The final ratio estimation procedure corrects for these deficiencies as far as the count of total housing is concerned; i.e., it adjusts the estimate of the total housing inventory to the best available estimate. However, biases of subtotals would still remain.

Rounding errors—For errors associated with processing, the rounding of estimates introduces another source of error in the data, the severity of which depends on 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-SMSA sample—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 each other. 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 due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by 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, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. 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.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two 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 (page App-50) are approximations to the standard errors of various estimates shown in this report for this SMSA. 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. Standard errors applicable to estimates of characteristics of the 1975 housing inventory can be found in the AHS Series H-170 reports for 1975.

Table I (page App-50) presents the standard errors applicable to estimates of characteristics of the 1979 housing inventory as well as estimates of characteristics of the 1975-1979 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 17,280 for the total SMSA, 10,620 for the central city of the SMSA, and 13,560 for the balance of the SMSA.

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 (page App-50) presents the standard errors of estimated percentages for the 1979 housing inventory as well as estimated percentages of the 1975-1979 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

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 as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is 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:

$$(100)\left(\frac{x}{y}\right)\sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}$$

where: x = the numerator of the ratio
y = the denominator of the ratio
σ_x = the standard error of the numerator
σ_y = the standard error of the denominator

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1979 there were 304,900 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 3,720. 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
300,000	3,750
304,900	x
400,000	3,190

The entry for "x" is determined as follows by vertically interpolating between 3,750 and 3,190.

$$304{,}900-300{,}000 = 4{,}900$$
$$400{,}000-300{,}000 = 100{,}000$$

$$3{,}750 + \frac{4{,}900}{100{,}000}(3{,}190-3{,}750) = 3{,}720$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 301,180 to 308,620 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1979 owner-occupied housing units lies within a range computed in this way, would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 298,950 to 310,850 housing units with 90 percent confidence; and that the average estimate lies within the interval from 297,460 to 312,340 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 304,900 owner-occupied housing units, 68,900, or 22.6 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 22.6 percent is approximately 0.8 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	22.6	25 or 75
300,000.	0.6	a	0.8
304,900.		p	
400,000.	0.5	b	0.7

1. The entry for cell "a" is determined by horizontal interpolation between 0.6 and 0.8.

$$22.6-10.0 = 12.6$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{12.6}{15.0}(0.8-0.6) = 0.8$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.5 and 0.7.

$$22.6-10.0 = 12.6$$
$$25.0-10.0 = 15.0$$

$$0.5 + \frac{12.6}{15.0}(0.7-0.5) = 0.7$$

3. The entry for "p" is then determined by vertical interpolation between 0.8 and 0.7.

$$304{,}900-300{,}000 = 4{,}900$$
$$400{,}000-300{,}000 = 100{,}000$$

$$0.8 + \frac{4{,}900}{100{,}000}(0.7-0.8) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 21.8 to 23.4 percent; the 90-percent confidence interval is from 21.3 to 23.9 percent; and the 95-percent confidence interval is from 21.0 to 24.2 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. 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. Due to the overlap of the 1975 and 1979 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1975 and 1979 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1979 there were 163,900 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 95,000. Table I shows that standard error of 163,900 is approximately 3,520, and the standard error of 68,900 is approximately 2,580. Therefore, the standard error of the estimated difference of 95,000 is about 4,360.

$$4{,}360 = \sqrt{(3{,}520)^2 + (2{,}580)^2}$$

Consequently, the 68-percent confidence interval for the 95,000 difference is from 90,640 to 99,360 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 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 88,020 to 101,980 housing units, and the 95-percent confidence interval is from 86,280 to 103,720 housing units. Thus, we can conclude with 95 percent confidence that the number of 1979 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied housing units with two bedrooms, since the 95-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 deter-

mine an interval about the estimated median so that there is a stated degree of confidence 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. These two distribution intervals could be different.

For about 68 out of 100 possible samples, the average median from all possible samples would lie between these two values.

A two-standard-error confidence interval may be determined by finding the values corresponding to 50 percent plus and minus twice the standard error determined in step 1. For about 95 out of 100 possible samples, the average median from all possible samples would lie between these two values.

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.9. The base of the distribution from which this median was determined is 304,900 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 304,900 is approximately 1.0 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.0 and 52.0.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with 3 persons (for purposes of calculating the median, the category of 3 persons is considered to be from 2.5 to 3.5 persons) contains the 48.0 percent derived in step 2. About 126,400 housing units or 41.5 percent fall below this interval, and 57,700 housing units or 18.9 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about 2.8.

$$2.5 + (3.5-2.5)\,\frac{(48.0-41.5)}{18.9} = 2.8$$

Similarly, the interval for owner-occupied housing units with 3 persons contains the 52.0 percent derived in step 2. About 126,400 housing units or 41.5 percent fall below this interval, and 57,700 housing units or 18.9 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about 3.1.

$$2.5 + (3.5-2.5)\,\frac{(52.0-41.5)}{18.9} = 3.1$$

Thus, the 95-percent confidence interval ranges from 2.8 to 3.1 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1979 Housing Inventory and for Estimated Number of 1975-1979 Lost Units for the Milwaukee, Wis., SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city
0	110	110	110
100	110	110	110
200	150	150	150
500	240	230	240
700	280	270	280
1,000	330	330	340
2,500	530	510	530
5,000	740	720	750
10,000	1,050	1,010	1,050
25,000	1,630	1,550	1,600
50,000	2,250	2,070	2,150
75,000	2,680	2,370	2,470
100,000	3,000	2,540	2,660
150,000	3,450	2,550	2,720
200,000	3,700	2,120	2,360
250,000	3,800	570	1,260
253,100	3,800	—	1,120
264,700	3,800	—	—
300,000	3,750	—	—
400,000	3,190	—	—
500,000	1,390	—	—
517,800	—	—	—

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2 for the total SMSA, 1.3 for the central city, and 1.2 for the balance (not in central city) estimates. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero since these estimates were derived from census data which are not subject to sampling errors. However, these estimates are subject to the nonsampling errors associated with 1970 and 1980 censuses and with the interpolation procedures used to derive these estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1979 Housing Inventory and for Estimated Percentages of 1975-1979 Lost Housing Units for the Milwaukee, Wis., SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
200	36.2	36.2	36.2	36.2	36.2	37.7
500	18.5	18.5	18.5	18.5	20.7	23.8
700	14.0	14.0	14.0	14.0	17.5	20.2
1,000	10.2	10.2	10.2	10.2	14.6	16.9
2,500	4.4	4.4	4.6	6.4	9.2	10.7
5,000	2.2	2.2	3.3	4.5	6.5	7.5
10,000	1.1	1.1	2.3	3.2	4.6	5.3
25,000	0.5	0.7	1.5	2.0	2.9	3.4
50,000	0.2	0.5	1.0	1.4	2.1	2.4
75,000	0.15	0.4	0.8	1.2	1.7	1.9
100,000	0.11	0.3	0.7	1.0	1.5	1.7
150,000	0.08	0.3	0.6	0.8	1.2	1.4
200,000	0.06	0.2	0.5	0.7	1.0	1.2
250,000	0.05	0.2	0.5	0.6	0.9	1.1
300,000	0.04	0.2	0.4	0.6	0.8	1.0
400,000	0.03	0.2	0.4	0.5	0.7	0.8
500,000	0.02	0.15	0.3	0.5	0.7	0.8
600,000	0.02	0.14	0.3	0.4	0.6	0.7

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than 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 1.2 for the total SMSA, 1.3 for the central city, and 1.2 for the balance (not in central city).

Omaha 1979

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-41
Annual Housing Survey App-41
Designation of sample housing units for the 1979 survey. App-41
Selection of the 1976 AHS-SMSA sample App-41
1976-1979 additions to the housing inventory App-43
Sample selection for the 1976 Coverage Improvement Program. App-43
Coverage improvement for deficiency 1 . . . . . App-43
Coverage improvement for deficiency 2 . . . . . App-43
Coverage improvement for deficiencies 3-6 . . . . . App-43
1970 Census of Population and Housing. App-44
ESTIMATION. App-44
1979 housing inventory App-44
1976-1979 lost housing units . . . App-45
1976 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES. App-46
Nonsampling errors. App-46
1970 census App-46
AHS-SMSA App-46
Coverage errors App-46
Rounding errors. App-46
Sampling errors for the AHS-SMSA sample App-47
Illustration of the use of the standard error tables App-47
Differences. App-48
Illustration of the computation of the standard error of a difference App-48
Medians App-48
Illustration of the computation of the 95-percent confidence interval of a median App-49
Standard error tables App-50

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 15 SMSA's in this report series (H-170-79) are based on data collected from the 1979 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 15 SMSA's, the data were collected during the 12-month period from April 1979 through March 1980 with one-twelfth of the sample housing units being visited each month.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1979 were interviewed previously in either 1975 or 1976 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1979 group of SMSA's, the largest SMSA from three of the four census regions of the United States is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All the remaining 12 SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the 1979 group are: Chicago, Ill.; Houston, Tex.; and Seattle-Everett, Wash.

The remaining SMSA's in the 1979 group are: Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.

In this SMSA, 4,947 housing units were eligible for interview. Of these sample housing units, 256 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 4,947 housing units eligible for interview, 376 units were visited but were not eligible for interview because they were condemned, unfit, demolished, converted to group quarters use, etc.

Designation of sample housing units for the 1979 survey—The sample housing units designated to be interviewed in the 1979 survey consisted of the following categories which are described in detail in the succeeding sections.

1. All sample housing units that were interviewed in the 1976 survey.
2. 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 1976 survey. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1979 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1976 survey. (This sample represented the housing units built in permit-issuing areas, since the 1976 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1976 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1976 survey.)
5. All sample housing units that were selected as part of the 1976 Coverage Improvement Program. (This sample represented most of the housing units which, until 1976, did not have a chance of selection.)

Selection of the 1976 AHS-SMSA sample—In 1970, the sample for the SMSA's which were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in

the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (nonpermit universe). In 1979, the following four SMSA's were 100-percent permit-issuing: Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; and Miami, Fla. The remaining 11 SMSA's contain a sample from the nonpermit universe.

Sampling operations, described in the following paragraphs, were performed separately within the central city and the balance of the SMSA for each of the sample frames. The overall sampling rate used to select the sample for each SMSA was determined by the size of the sample. Thus, for the three largest SMSA's, the overall sampling rate differed by central city and the balance of the SMSA, since the sample was divided equally between the central city and the balance of the SMSA. The remaining SMSA's had an overall sampling rate about the same for the sample selected from both the central city and the balance of the SMSA, since the sample was distributed proportionately between the central city and the balance of the SMSA according to the distribution of the total housing units in each sector.

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA 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 housing unit records were stratified by race of head (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 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household 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, for this SMSA, 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 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. 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 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 SMSA. 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.

The second frame from which this SMSA 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 this SMSA. Prior to sample selection, the list of permits was chronologically 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.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS 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 (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

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 are included

1976-1979 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1976 survey, was selected to represent housing units built in permit-issuing areas since the 1976 survey. Sampling procedures were identical to those used in selecting the 1970-1976 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1976, to identify any housing units missed in the 1976 survey or any housing units added since the 1976 survey.

Sample selection for the 1976 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

1. New construction housing units from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes and trailers placed in parks either missed during the 1970 census or established since the 1970 census.
3. Housing units missed in the 1970 census.
4. Housing units converted to residential use that were nonresidential at the time of the 1970 census.
5. Housing units that have been moved onto their present site since the 1970 census.
6. Mobile homes and trailers placed outside parks since the 1970 census or vacant at the time of the 1970 census.

For each of the five SMSA's interviewed for the first time in 1975, a full implementation of the Coverage Improvement Program was conducted as a part of the 1979 AHS. For each of the 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with some updating and refining as a part of the 1979 AHS. The following discussion of coverage improvement procedures applies to both the 1976 and 1979 programs. For the 10 (1976) SMSA's, estimates of housing units added by a specific procedure reflect units added in 1976 as well as any additions that resulted from the updating and refining in 1979.

Coverage improvement for deficiency 1—A sample of new construction housing units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 1979 SMSA's. Two different procedures were used. For the SMSA's first interviewed in 1976, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. These new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units).

The first stage sample selection was a sample of permit offices, and the second stage a sample of the 1969 permits within each of the selected permit offices. (In the Honolulu, Hawaii, SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second stage.) For the third stage, structures of size three or more were divided into clusters of an expected size of four housing units and a sample of clusters was selected.

For the SMSA's first interviewed in 1975, housing units whose permits were issued before January 1970, but which were completed after April 1970, were identified from the Survey of Construction (SOC), a survey of building permits conducted monthly by the Bureau of the Census. These housing units were then sampled at one-third the rate of regular AHS units. These procedures added an estimated 1,613 new construction housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiency 2—In permit-issuing areas, a sample of mobile homes and trailers placed in parks that were missed by the census or established after the census was selected in two stages. First, a sample of tracts was selected and canvassed. All parks were listed and then matched back to the 1970 census to identify parks missed by the census and parks established after the census. Second, the parks were divided into clusters of an expected size of four sites and a sample of clusters was selected and interviewed. Each of the sample housing units represented the same number of units that the regular AHS sample housing units represented. Since research showed that there were no mobile home or trailer parks in the central cities of the Chicago, Ill.; Hartford, Conn.; and Portland, Oreg.-Wash.; SMSA's, this procedure was only implemented outside the central cities for these SMSA's. Based on a cost-benefit analysis, this procedure was updated as part of the 1979 AHS only for the Denver, Colo. and Las Vegas, Nev., SMSA's. This procedure added 298 housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed housing units were sampled by one of two procedures. The first procedure was designed to represent units from the following types of missed structures (structures that had no chance of selection for the AHS):

1. Structures missed in the 1970 census.
2. Structures that were completely nonresidential in the 1970 census but now contain residential housing units.
3. Mobile homes and trailers that had been placed outside parks since the 1970 census and have a utility hookup, or were on the site during the present survey but not occupied on April 1, 1970, or had no utility hookup but were occupied by persons with no usual residence elsewhere.
4. Housing units that had been moved onto their present site since the 1970 census.

Initially, a subsample of regular AHS units was selected at a rate of 1 in 24 for the Chicago, Ill.; Hartford, Conn.; Miami, Fla.; Milwaukee, Wis.; and Portland, Oreg.-Wash.; SMSA's, and at a rate of 1 in 22.2152 for the 10 remaining SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing units were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and housing units within

these structures were interviewed. In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. Although the Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Houston, Tex.; Omaha, Nebr.-Iowa; and Raleigh, N.C.; SMSA's, were interviewed for the first time in 1976 and had these sampling procedures implemented at that time, some 1976 coverage improvement assignments were not sent out to be interviewed due to time limitations. The sampling was completed as part of the 1979 AHS and this procedure added an estimated 3,619 housing units to the coverage of the housing inventory of this SMSA.

The second procedure was designed to represent missed housing units from structures represented in the AHS. These missed housing units were:

1. Housing units missed in the 1970 census.
2. Nonresidential space converted to residential use since the 1970 census in structures that contained some residential housing units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structures selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 371 housing units to the coverage of the housing inventory for this SMSA.

1970 Census of Population and Housing—The estimates pertaining to the 1970 housing inventory (i.e., the housing inventory that existed at the time of the 1970 census) are based on either 20-, 15-, or 5-percent sample data collected in April 1970 for the Decennial Census of Population and Housing. A detailed description of the sample design employed for the 1970 census can be obtained in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

ESTIMATION

The 1979 AHS sample produced two types of estimates for each SMSA: Estimates pertaining to characteristics of the housing inventory at the time of the interview (i.e., the 1979 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1976 (i.e., 1976-1979 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1979 housing inventory—The AHS estimates of characteristics of the 1979 housing inventory were produced using a 2-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and a 3-stage ratio estimation procedure for the remaining 11 SMSA's. Prior to 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 the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of one or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes and trailers from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample housing units from both the nonpermit universe and the coverage improvement universe (if units were not included in any of the previous cells).

The following first-stage 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 permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of housing units enumerated in areas under the jurisdiction of permit-issuing offices. The denominators of the ratios were obtained from weighted estimates of all the AHS sample housing units within the corresponding ratio estimation categories using the existing weight (i.e., the basic weight times the noninterview factor). The computed ratio estimation factor was then applied to the existing weight for each sample housing unit within the corresponding ratio estimation category.

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 permit-issuing universe. Prior to the AHS sample selection within each SMSA, 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 sample. Since the number of housing units deleted from the AHS universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the Houston, Tex., SMSA, a second ratio estimation procedure was employed for new construction housing units outside the Houston central city. Due to rapid growth in the

nonpermit-issuing universe the sampling methods did not adequately represent the new construction outside the Houston central city. Because of this deficiency, separate factors were computed for the following cells:

1. Conventional housing units (non-mobile home or trailer) outside the central city, which were built between the 1970 census and the 1976 survey.
2. Mobile homes and trailers outside the central city with a model year between the 1970 census and 1976 survey.
3. Conventional housing units outside the central city built between the 1976 and 1979 surveys.
4. Mobile homes and trailers outside the central city with a model year between the 1976 and 1979 surveys.

In a specific cell, the ratio estimation factor was equal to the following:

$$\frac{\text{Weighted count of housing units in the cell} + \text{Estimate of the undercoverage in the cell}}{\text{Weighted count of housing units in the cell}}$$

The computed second-stage ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell. For the Baltimore, Md.; Buffalo, N.Y.; Chicago, Ill.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; and Seattle-Everett, Wash.; SMSA's, a second-stage ratio estimation procedure was employed to adjust the central city/balance distribution of the weighted sample estimate of new construction housing units built since the last survey in permit-issuing areas to an independently derived estimate of this distribution.

This ratio estimation factor was calculated separately for the central city and balance of each SMSA and was applied to all new construction housing units from permit-issuing areas within the corresponding sector (central city or balance of the SMSA). This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}{\text{(Sample estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}$$

The independent estimates of new construction were based upon the number of authorized building permits which were determined from the Survey of Construction (SOC). The sample estimates were obtained from the weighted estimate of the AHS-SMSA sample housing units after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

This second-stage ratio estimation procedure was not implemented in the Honolulu, Hawaii; Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's. In the Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, the central city permit offices, sources of SOC information on permits, had expanded their coverage of permits to include areas greater than those covered by the 1970 census central city definitions. Since the AHS-SMSA survey uses the 1970 definitions, use of this ratio estimation procedure would not correctly adjust the central city/balance distribution of new construction housing units. In the Honolulu, Hawaii, SMSA, this procedure could not be used because the permit office covers the entire SMSA and does not differentiate between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Neb.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and as a third-stage ratio estimation procedure for the remaining 11 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the October 1979 housing inventory of each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sectors. This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the October 1979 housing inventory for the corresponding sector of the SMSA)}}{\text{(AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA)}}$$

The numerator of this ratio was derived using 1970 and 1980 census counts. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight after the first-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and the existing weight after the second-stage ratio estimation procedure for the 11 remaining SMSA's.

The computed ratio estimation factors for the central city and balance of the SMSA's were then applied to the existing weight for all corresponding sample units and the resulting product was used as the final weight for tabulation purposes.

The effect of the total housing unit ratio estimation procedure, as well as the overall 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 SMSA as a whole, it can be expected that the sample estimates will be improved when the sample housing population, or different portions of it, are brought into agreement with known good estimates of the SMSA housing population.

1976-1979 lost housing units—The 1976-1979 lost housing unit (housing unit removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1976 housing inventory, as was described in the 1976 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1976-1979 lost housing units existed, by definition, in the 1976 housing inventory, there was a 1976 housing inventory weight associated with each 1976-1979 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1976-1979 lost housing units.

1976 estimation procedure—This report presents data on the housing characteristics of the 1976 housing inventory from the 1976 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1976.

Ratio estimation procedure of the 1970 Census of Population and Housing—This report presents data on the housing characteristics of the 1970 housing inventory from the 1970 Census of Population and Housing. The statistics based on 1970 census sample data employed a ratio estimation procedure which was applied separately for each of the three census samples. A detailed description of this ratio estimation procedure can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

RELIABILITY OF THE ESTIMATES

There are two types of possible errors associated with estimates based on data from sample surveys—sampling and nonsampling errors. The following is a description of the sampling and nonsampling errors associated with the AHS-SMSA sample and of the nonsampling errors associated with the 1970 census estimates. A description of the sampling errors associated with the sample estimates from the 1970 census can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

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; and other errors of collection, response, processing, 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 1970 Census of Population and Housing and the 1976 AHS-SMSA sample.

1970 census—A number of studies were conducted to measure two types of general errors associated with 1970 census estimates—"coverage" and "content" errors. The "coverage" errors determined how completely housing units were counted in the census and the extent to which occupancy status was erroneously reported. The "content" errors measured the accuracy of the data collected for enumerated housing units. These errors were measured by reinterviews, record checks, and other surveys.

The detailed results of these studies, as well as the methodology employed, are presented in the 1970 Census of Population and Housing Evaluation and Research Program Reports, Series PHC(E)-5, *The Coverage of Housing in the 1970 Census,* and PHC(E)-10, *Accuracy of Data for Selected Housing Characteristics as Measured by Reinterviews.*

AHS-SMSA—Results from the 1979 AHS-SMSA sample reinterview program were not available at the time this report was being prepared. However, a study was conducted for the 1976 AHS-SMSA sample. The results are presented in the Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1976."

Coverage errors—In errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional (non-mobile home or trailer) new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units. Due to 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, in which case, 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, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit.

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 estimated that the 1976 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 for the 1979 survey, the number of missed housing units may be considerably less for 1979.

The final ratio estimation procedure corrects for these deficiencies as far as the count of total housing is concerned; i.e., it adjusts the estimate of the total housing inventory to the best available estimate. However, biases of subtotals would still remain.

Rounding errors—For errors associated with processing, the rounding of estimates introduces another source of error in the data, the severity of which depends on 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-SMSA sample—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 each other. 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 due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by 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, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. 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.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two 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 (page App-50) are approximations to the standard errors of various estimates shown in this report for this SMSA. 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. Standard errors applicable to estimates of characteristics of the 1976 housing inventory can be found in the AHS Series H-170 reports for 1976.

Table I (page App-50) presents the standard errors applicable to estimates of characteristics of the 1979 housing inventory as well as estimates of characteristics of the 1976-1979 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 8,470 for the total SMSA, 4,330 for the central city of the SMSA, and 7,190 for the balance of the SMSA.

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 (page App-50) presents the standard errors of estimated percentages for the 1979 housing inventory as well as estimated percentages of the 1976-1979 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

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 as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is 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:

$$(100)\left(\frac{x}{y}\right)\sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}$$

where: x = the numerator of the ratio
y = the denominator of the ratio
σ_x = the standard error of the numerator
σ_y = the standard error of the denominator

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1979 there were 131,500 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 1,550. 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
125,300	1,580
131,500	x
150,000	1,460

The entry for "x" is determined as follows by vertically interpolating between 1,580 and 1,460.

$$131{,}500-125{,}300 = 6{,}200$$
$$150{,}000-125{,}300 = 24{,}700$$

$$1{,}580 + \frac{6{,}200}{24{,}700}(1{,}460-1{,}580) = 1{,}550$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 129,950 to 133,050 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples of 1979 owner-occupied housing units lies within a range computed in this way, would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 129,020 to 133,980 housing units with 90 percent confidence; and that the average estimate lies within the interval from 128,400 to 134,600 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 131,500 owner-occupied housing units, 33,300, or 25.3 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 25.3 percent is approximately 0.8 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	25 or 75	25.3	50
100,000.	0.9	a	1.1
131,500.		p	
150,000.	0.8	b	0.9

1. The entry for cell "a" is determined by horizontal interpolation between 0.9 and 1.1.

$$25.3-2.5 = 0.3$$
$$50.0-25.0 = 25.0$$

$$0.9 + \frac{0.3}{25.0}(1.1-0.9) = 0.9$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.8 and 0.9.

$$25.3-25.0 = 0.3$$
$$50.0-25.0 = 25.0$$

$$0.8 + \frac{0.3}{25.0}(0.9-0.8) = 0.8$$

3. The entry for "p" is then determined by vertical interpolation between 0.9 and 0.8.

$$131{,}500-100{,}000 = 31{,}500$$
$$150{,}000-100{,}000 = 50{,}000$$

$$0.9 + \frac{31{,}500}{50{,}000}(0.8-0.9) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 24.5 to 26.1 percent; the 90-percent confidence interval is from 24.0 to 26.6 percent; and the 95-percent confidence interval is from 23.7 to 26.9 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. 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. Due to the overlap of the 1976 and 1979 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1979 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1979 there were 69,600 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 36,300. Table I shows the standard error of 69,600 is approximately 1,490, and the standard error of 33,300 is approximately 1,140. Therefore, the standard error of the estimated difference of 36,300 is about 1,880.

$$1{,}880 = \sqrt{(1{,}490)^2 + (1{,}140)^2}$$

Consequently, the 68-percent confidence interval for the 36,300 difference is from 34,420 to 38,180 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 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 33,290 to 39,310 housing units, and the 95-percent confidence interval is from 32,540 to 40,060 housing units. Thus, we can conclude with 95 percent confidence that the number of 1979 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied housing units with two bedrooms, since the 95-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 deter-

mine an interval about the estimated median so that there is a stated degree of confidence 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. These two distribution intervals could be different.

For about 68 out of 100 possible samples, the average median from all possible samples would lie between these two values.

A two-standard-error confidence interval may be determined by finding the values corresponding to 50 percent plus and minus twice the standard error determined in step 1. For about 95 out of 100 possible samples, the average median from all possible samples would lie between these two values.

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.8. The base of the distribution from which this median was determined is 131,500 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 131,500 is approximately 1.0 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.0 and 52.0.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with 3 persons (for purposes of calculating the median, the category of 3 persons is considered to be from 2.5 to 3.5 persons) contains the 48.0 percent derived in step 2. About 57,500 housing units or 43.7 percent fall below this interval, and 23,900 housing units or 18.2 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(48.0-43.7)}{18.2} = 2.7$$

Similarly, the interval for owner-occupied housing units with 3 persons contains the 52.0 percent derived in step 2. About 57,500 housing units or 43.7 percent fall below this interval, and 23,900 housing units or 18.2 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(52.0-43.7)}{18.2} = 3.0$$

Thus, the 95-percent confidence interval ranges from 2.7 to 3.0 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1979 Housing Inventory and for Estimated Number of 1976-1979 Lost Units for the Omaha, Nebr.-Iowa, SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city
0	50	50	50
100	70	70	70
200	100	100	100
500	150	150	150
700	180	180	180
1,000	220	220	220
2,500	340	340	340
5,000	480	480	470
10,000	680	660	650
25,000	1,030	980	920
50,000	1,350	1,200	1,020
75,000	1,530	1,200	740
88,400	1,580	1,120	–
100,000	1,600	980	–
125,300	1,580	–	–
150,000	1,460	–	–
200,000	780	–	–
213,700	–	–	–

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.1 for the total SMSA, 1.3 for the central city, and 1.1 for the balance (not in central city) estimates. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero since these estimates were derived from census data which are not subject to sampling error. However, these estimates are subject to the nonsampling errors associated with the 1970 and 1980 censuses and with the interpolation procedures used to derive these estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1979 Housing Inventory and for Estimated Percentages of 1976-1979 Lost Housing Units for the Omaha, Nebr.-Iowa, SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
200	19.4	19.4	19.4	19.4	21.2	24.5
500	8.8	8.8	8.8	9.3	13.4	15.5
700	6.4	6.4	6.4	7.9	11.3	13.1
1,000	4.6	4.6	4.8	6.6	9.5	11.0
2,500	1.9	1.9	3.0	4.2	6.0	6.9
5,000	1.0	1.0	2.1	2.9	4.2	4.9
10,000	0.5	0.7	1.5	2.1	3.0	3.5
25,000	0.2	0.4	1.0	1.3	1.9	2.2
50,000	0.1	0.3	0.7	0.9	1.3	1.5
75,000	0.1	0.3	0.6	0.8	1.1	1.3
100,000	0.0	0.2	0.5	0.7	0.9	1.1
150,000	0.0	0.2	0.4	0.5	0.8	0.9
200,000	0.0	0.2	0.3	0.5	0.7	0.8
250,000	0.0	0.1	0.3	0.4	0.6	0.7

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than 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 for the total SMSA, 1.3 for the central city, and 1.1 for the balance (not in central city).

Portland 1979

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-41
Annual Housing Survey App-41
Designation of sample housing units for the 1979 survey. App-41
Selection of the 1975 AHS-SMSA sample App-41
1975-1979 additions to the housing inventory App-43
Sample selection for the 1979 Coverage Improvement Program. App-43
Coverage improvement for deficiency 1 . . . . . App-43
Coverage improvement for deficiency 2 . . . . . App-43
Coverage improvement for deficiencies 3-6 . . . . . App-43
1970 Census of Population and Housing. App-44
ESTIMATION. App-44
1979 housing inventory App-44
1975-1979 lost housing units . . . App-45
1975 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES. App-46
Nonsampling errors. App-46
1970 census App-46
AHS-SMSA App-46
Coverage errors App-46
Rounding errors. App-46
Sampling errors for the AHS-SMSA sample App-47
Illustration of the use of the standard error tables App-47
Differences. App-48
Illustration of the computation of the standard error of a difference App-48
Medians App-48
Illustration of the computation of the 95-percent confidence interval of a median App-49
Standard error tables App-50

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 15 SMSA's in this report series (H-170-79) are based on data collected from the 1979 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 15 SMSA's, the data were collected during the 12-month period from April 1979 through March 1980 with one-twelfth of the sample housing units being visited each month.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1979 were interviewed previously in either 1975 or 1976 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1979 group of SMSA's, the largest SMSA from three of the four census regions of the United States is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All the remaining 12 SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the 1979 group are: Chicago, Ill.; Houston, Tex.; and Seattle-Everett, Wash.

The remaining SMSA's in the 1979 group are: Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.

In this SMSA, 5,743 housing units were eligible for interview. Of these sample housing units, 260 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 5,743 housing units eligible for interview, 354 units were visited but were not eligible for interview because they were condemned, unfit, demolished, converted to group quarters use, etc.

Designation of sample housing units for the 1979 survey—The sample housing units designated to be interviewed in the 1979 survey consisted of the following categories which are described in detail in the succeeding sections.

1. All sample housing units that were interviewed in the 1975 survey.
2. 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 1975 survey. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1979 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1975 survey. (This sample represented the housing units built in permit-issuing areas since the 1975 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1975 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1975 survey.)
5. All sample housing units that were selected as part of the 1979 Coverage Improvement Program. (This sample represented most of the housing units which, until 1979, did not have a chance of selection.)

Selection of the 1975 AHS-SMSA sample—In 1970, the sample for the SMSA's which were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in

the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (nonpermit universe). In 1979, the following four SMSA's were 100-percent permit-issuing: Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; and Miami, Fla. The remaining 11 SMSA's contain a sample from the nonpermit universe.

Sampling operations, described in the following paragraphs, were performed separately within the central city and the balance of the SMSA for each of the sample frames. The overall sampling rate used to select the sample for each SMSA was determined by the size of the sample. Thus, for the three largest SMSA's, the overall sampling rate differed by central city and the balance of the SMSA, since the sample was divided equally between the central city and the balance of the SMSA. The remaining SMSA's had an overall sampling rate about the same for the sample selected from both the central city and the balance of the SMSA, since the sample was distributed proportionately between the central city and the balance of the SMSA according to the distribution of the total housing units in each sector.

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA 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 housing unit records were stratified by race of head (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 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household 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, for this SMSA, 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 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. 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 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 SMSA. 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.

The second frame from which this SMSA 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 this SMSA. Prior to sample selection, the list of permits was chronologically 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.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS 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 (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

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 are included.

1975-1979 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1975 survey, was selected to represent housing units built in permit-issuing areas since the 1975 survey. Sampling procedures were identical to those used in selecting the 1970-1975 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1975, to identify any housing units missed in the 1975 survey or any housing units added since the 1975 survey.

Sample selection for the 1979 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

1. New construction housing units from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes and trailers placed in parks either missed during the 1970 census or established since the 1970 census.
3. Housing units missed in the 1970 census.
4. Housing units converted to residential use that were nonresidential at the time of the 1970 census.
5. Housing units that have been moved onto their present site since the 1970 census.
6. Mobile homes and trailers placed outside parks since the 1970 census or vacant at the time of the 1970 census.

For each of the five SMSA's interviewed for the first time in 1975, a full implementation of the Coverage Improvement Program was conducted as a part of the 1979 AHS. For each of the 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with some updating and refining as a part of the 1979 AHS. The following discussion of coverage improvement procedures applies to both the 1976 and 1979 programs. For the 10 (1976) SMSA's, estimates of housing units added by a specific procedure reflect units added in 1976 as well as any additions that resulted from the updating and refining in 1979.

Coverage improvement for deficiency 1—A sample of new construction housing units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 1979 SMSA's. Two different procedures were used. For the SMSA's first interviewed in 1976, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. These new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units).

The first stage sample selection was a sample of permit offices, and the second stage a sample of the 1969 permits within each of the selected permit offices. (In the Honolulu, Hawaii, SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second stage.) For the third stage, structures of size three or more were divided into clusters of an expected size of four housing units and a sample of clusters was selected.

For the SMSA's first interviewed in 1975, housing units whose permits were issued before January 1970, but which were completed after April 1970, were identified from the Survey of Construction (SOC), a survey of building permits conducted monthly by the Bureau of the Census. These housing units were then sampled at one-third the rate of regular AHS units. These procedures added an estimated 5,233 new construction housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiency 2—In permit-issuing areas, a sample of mobile homes and trailers placed in parks that were missed by the census or established after the census was selected in two stages. First, a sample of tracts was selected and canvassed. All parks were listed and then matched back to the 1970 census to identify parks missed by the census and parks established after the census. Second, the parks were divided into clusters of an expected size of four sites and a sample of clusters was selected and interviewed. Each of the sample housing units represented the same number of units that the regular AHS sample housing units represented. Since research showed that there were no mobile home or trailer parks in the central cities of the Chicago, Ill.; Hartford, Conn.; and Portland, Oreg.-Wash.; SMSA's, this procedure was only implemented outside the central cities for these SMSA's. Based on a cost-benefit analysis, this procedure was updated as part of the 1979 AHS only for the Denver, Colo. and Las Vegas, Nev., SMSA's. This procedure added 3,907 housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed housing units were sampled by one of two procedures. The first procedure was designed to represent units from the following types of missed structures (structures that had no chance of selection for the AHS):

1. Structures missed in the 1970 census.
2. Structures that were completely nonresidential in the 1970 census but now contain residential housing units.
3. Mobile homes and trailers that had been placed outside parks since the 1970 census and have a utility hookup, or were on the site during the present survey but not occupied on April 1, 1970, or had no utility hookup but were occupied by persons with no usual residence elsewhere.
4. Housing units that had been moved onto their present site since the 1970 census.

Initially, a subsample of regular AHS units was selected at a rate of 1 in 24 for the Chicago, Ill.; Hartford, Conn.; Miami, Fla.; Milwaukee, Wis.; and Portland, Oreg.-Wash., SMSA's, and at a rate of 1 in 22.2152 for the 10 remaining SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing units were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and housing units within

these structures were interviewed. In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. Although the Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Houston, Tex.; Omaha, Nebr.-Iowa; and Raleigh, N.C., SMSA's, were interviewed for the first time in 1976 and had these sampling procedures implemented at that time, some 1976 coverage improvement assignments were not sent out to be interviewed due to time limitations. The sampling was completed as part of the 1979 AHS and this procedure added an estimated 12,343 housing units to the coverage of the housing inventory of this SMSA.

The second procedure was designed to represent missed housing units from structures represented in the AHS. These missed housing units were:

1. Housing units missed in the 1970 census.
2. Nonresidential space converted to residential use since the 1970 census in structures that contained some residential housing units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structures selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 779 housing units to the coverage of the housing inventory for this SMSA.

1970 Census of Population and Housing—The estimates pertaining to the 1970 housing inventory (i.e., the housing inventory that existed at the time of the 1970 census) are based on either 20-, 15-, or 5-percent sample data collected in April 1970 for the Decennial Census of Population and Housing. A detailed description of the sample design employed for the 1970 census can be obtained in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

ESTIMATION

The 1979 AHS sample produced two types of estimates for each SMSA: Estimates pertaining to characteristics of the housing inventory at the time of the interview (i.e., the 1979 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1975 (i.e., 1975-1979 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1979 housing inventory—The AHS estimates of characteristics of the 1979 housing inventory were produced using a 2-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C., SMSA's, and a 3-stage ratio estimation procedure for the remaining 11 SMSA's. Prior to 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 the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of one or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes and trailers from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample housing units from both the nonpermit universe and the coverage improvement universe (if units were not included in any of the previous cells).

The following first-stage 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 permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of housing units enumerated in areas under the jurisdiction of permit-issuing offices. The denominators of the ratios were obtained from weighted estimates of all the AHS sample housing units within the corresponding ratio estimation categories using the existing weight (i.e., the basic weight times the noninterview factor). The computed ratio estimation factor was then applied to the existing weight for each sample housing unit within the corresponding ratio estimation category.

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 permit-issuing universe. Prior to the AHS sample selection within each SMSA, 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 sample. Since the number of housing units deleted from the AHS universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the Houston, Tex., SMSA, a second ratio estimation procedure was employed for new construction housing units outside the Houston central city. Due to rapid growth in the

nonpermit-issuing universe the sampling methods did not adequately represent the new construction outside the Houston central city. Because of this deficiency, separate factors were computed for the following cells:

1. Conventional housing units (non-mobile home or trailer) outside the central city, which were built between the 1970 census and the 1976 survey.
2. Mobile homes and trailers outside the central city with a model year between the 1970 census and 1976 survey.
3. Conventional housing units outside the central city built between the 1976 and 1979 surveys.
4. Mobile homes and trailers outside the central city with a model year between the 1976 and 1979 surveys.

In a specific cell, the ratio estimation factor was equal to the following:

$$\frac{\text{Weighted count of housing units in the cell} + \text{Estimate of the undercoverage in the cell}}{\text{Weighted count of housing units in the cell}}$$

The computed second-stage ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell. For the Baltimore, Md.; Buffalo, N.Y.; Chicago, Ill.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; and Seattle-Everett, Wash., SMSA's, a second-stage ratio estimation procedure was employed to adjust the central city/balance distribution of the weighted sample estimate of new construction housing units built since the last survey in permit-issuing areas to an independently derived estimate of this distribution.

This ratio estimation factor was calculated separately for the central city and balance of each SMSA and was applied to all new construction housing units from permit-issuing areas within the corresponding sector (central city or balance of the SMSA). This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}{\text{(Sample estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}$$

The independent estimates of new construction were based upon the number of authorized building permits which were determined from the Survey of Construction (SOC). The sample estimates were obtained from the weighted estimate of the AHS-SMSA sample housing units after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

This second-stage ratio estimation procedure was not implemented in the Honolulu, Hawaii; Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C., SMSA's. In the Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C., SMSA's, the central city permit offices, sources of SOC information on permits, had expanded their coverage of permits to include areas greater than those covered by the 1970 census central city definitions. Since the AHS-SMSA survey uses the 1970 definitions, use of this ratio estimation procedure would not correctly adjust the central city/balance distribution of new construction housing units. In the Honolulu, Hawaii, SMSA, this procedure could not be used because the permit office covers the entire SMSA and does not differentiate between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Neb.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C., SMSA's, and as a third-stage ratio estimation procedure for the remaining 11 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the October 1979 housing inventory of each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sectors. This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the October 1979 housing inventory for the corresponding sector of the SMSA)}}{\text{(AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA)}}$$

The numerator of this ratio was derived using 1970 and 1980 census counts. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight after the first-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C., SMSA's, and the existing weight after the second-stage ratio estimation procedure for the 11 remaining SMSA's.

The computed ratio estimation factors for the central city and balance of the SMSA's were then applied to the existing weight for all corresponding sample units and the resulting product was used as the final weight for tabulation purposes.

The effect of the total housing unit ratio estimation procedure, as well as the overall 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 SMSA as a whole, it can be expected that the sample estimates will be improved when the sample housing population, or different portions of it, are brought into agreement with known good estimates of the SMSA housing population.

1975-1979 lost housing units—The 1975-1979 lost housing unit (housing unit removed from the inventory) estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1975 housing inventory, as was described in the 1975 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1975-1979 lost housing units existed, by definition, in the 1975 housing inventory, there was a 1975 housing inventory weight associated with each 1975-1979 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1975-1979 lost housing units.

1975 estimation procedure—This report presents data on the housing characteristics of the 1975 housing inventory from the 1975 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1975.

Ratio estimation procedure of the 1970 Census of Population and Housing—This report presents data on the housing characteristics of the 1970 housing inventory from the 1970 Census of Population and Housing. The statistics based on 1970 census sample data employed a ratio estimation procedure which was applied separately for each of the three census samples. A detailed description of this ratio estimation procedure can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

RELIABILITY OF THE ESTIMATES

There are two types of possible errors associated with estimates based on data from sample surveys—sampling and nonsampling errors. The following is a description of the sampling and nonsampling errors associated with the AHS-SMSA sample and of the nonsampling errors associated with the 1970 census estimates. A description of the sampling errors associated with the sample estimates from the 1970 census can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

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; and other errors of collection, response, processing, 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 1970 Census of Population and Housing and the 1975 AHS-SMSA sample.

1970 census—A number of studies were conducted to measure two types of general errors associated with 1970 census estimates—"coverage" and "content" errors. The "coverage" errors determined how completely housing units were counted in the census and the extent to which occupancy status was erroneously reported. The "content" errors measured the accuracy of the data collected for enumerated housing units. These errors were measured by reinterviews, record checks, and other surveys.

The detailed results of these studies, as well as the methodology employed, are presented in the 1970 Census of Population and Housing Evaluation and Research Program Reports, Series PHC(E)-5, *The Coverage of Housing in the 1970 Census,* and PHC(E)-10, *Accuracy of Data for Selected Housing Characteristics as Measured by Reinterviews.*

AHS-SMSA—Results from the 1979 AHS-SMSA sample reinterview program were not available at the time this report was being prepared. However, a study was conducted for the 1975 AHS-SMSA sample. The results are presented in the Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1975."

Coverage errors—In errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional (non-mobile home or trailer) new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units. Due to 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, in which case, 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, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit.

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 estimated that the 1975 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 for the 1979 survey, the number of missed housing units may be considerably less for 1979.

The final ratio estimation procedure corrects for these deficiencies as far as the count of total housing is concerned; i.e., it adjusts the estimate of the total housing inventory to the best available estimate. However, biases of subtotals would still remain.

Rounding errors—For errors associated with processing, the rounding of estimates introduces another source of error in the data, the severity of which depends on 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-SMSA sample—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 each other. 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 due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by 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, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. 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.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two 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 (page App-50) are approximations to the standard errors of various estimates shown in this report for this SMSA. 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. Standard errors applicable to estimates of characteristics of the 1975 housing inventory can be found in the AHS Series H-170 reports for 1975.

Table I (page App-50) presents the standard errors applicable to estimates of characteristics of the 1979 housing inventory as well as estimates of characteristics of the 1975-1979 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 13,350 for the total SMSA, 6,890 for the central city of the SMSA, and 11,360 for the balance of the SMSA.

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 (page App-50) presents the standard errors of estimated percentages for the 1979 housing inventory as well as estimated percentages of the 1975-1979 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

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 as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is 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:

$$(100)\left(\frac{x}{y}\right)\sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}$$

where: x = the numerator of the ratio
y = the denominator of the ratio
σ_x = the standard error of the numerator
σ_y = the standard error of the denominator

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1979 there were 297,800 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 3,520. 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
250,000	3,590
297,800	x
300,000	3,520

The entry for "x" is determined as follows by vertically interpolating between 3,590 and 3,520.

$$297{,}800-250{,}000 = 47{,}800$$
$$300{,}000-250{,}000 = 50{,}000$$

$$3{,}590 + \frac{47{,}800}{50{,}000}(3{,}520-3{,}590) = 3{,}520$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 294,280 to 301,320 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1979 owner-occupied housing units lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 292,170 to 303,430 housing units with 90 percent confidence; and that the average estimate lies within the interval from 290,760 to 304,840 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 297,800 owner-occupied housing units, 77,600, or 26.1 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 26.1 percent is approximately 0.8 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	25 or 75	26.1	50
250,000.	0.9	a	1.0
297,800.		p	
300,000.	0.8	b	0.9

1. The entry for cell "a" is determined by horizontal interpolation between 0.9 and 1.0.

$$26.1-25.0 = 1.1$$
$$50.0-25.0 = 25.0$$

$$0.9 + \frac{1.1}{25.0}(1.0-0.9) = 0.9$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.8 and 0.9.

$$26.1-25.0 = 1.1$$
$$50.0-25.0 = 25.0$$

$$0.8 + \frac{1.1}{25.0}(0.9-0.8) = 0.8$$

3. The entry for "p" is then determined by vertical interpolation between 0.9 and 0.8.

$$297{,}800-250{,}000 = 47{,}800$$
$$300{,}000-250{,}000 = 50{,}000$$

$$0.9 + \frac{47{,}800}{50{,}000}(0.8-0.9) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 25.3 to 26.9 percent; the 90-percent confidence interval is from 24.8 to 27.4 percent; and the 95-percent confidence interval is from 24.5 to 27.7 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. 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. Due to the overlap of the 1975 and 1979 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1975 and 1979 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1979 there were 144,500 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 66,900. Table I shows the standard error of 77,600 is approximately 2,600, and the standard error of 144,500 is approximately 3,250. Therefore, the standard error of the estimated difference of 66,900 is about 4,160.

$$4{,}160 = \sqrt{(2{,}600)^2 + (3{,}250)^2}$$

Consequently, the 68-percent confidence interval for the 66,900 difference is from 62,740 to 71,060 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 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 60,240 to 73,560 housing units, and the 95-percent confidence interval is from 58,580 to 75,220 housing units. Thus, we can conclude with 95 percent confidence that the number of 1979 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied housing units with two bedrooms, since the 95-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 deter-

mine an interval about the estimated median so that there is a stated degree of confidence 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. These two distribution intervals could be different.

For about 68 out of 100 possible samples, the average median from all possible samples would lie between these two values.

A two-standard-error confidence interval may be determined by finding the values corresponding to 50 percent plus and minus twice the standard error determined in step 1. For about 95 out of 100 possible samples, the average median from all possible samples would lie between these two values.

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.5. The base of the distribution from which this median was determined is 297,800 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 297,800 is approximately 0.9 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.2 and 51.8.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) contains the 48.2 percent derived in step 2. About 46,200 housing units or 15.5 percent fall below this interval, and 106,500 housing units or 35.8 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$1.5 + (2.5-1.5)\,\frac{(48.2-15.5)}{35.8} = 2.4$$

Similarly, the interval for owner-occupied housing units with three persons contains the 51.8 percent derived in step 2. About 152,700 housing units or 51.3 percent fall below this interval, and 50,600 housing units or 17.0 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(51.8-51.3)}{17.0} = 2.5$$

Thus, the 95-percent confidence interval ranges from 2.4 to 2.5 persons. Although it appears that this confidence interval has the sample estimate as the lower limit, it actually is a reflection of the rounding error associated with the median (see the paragraph on rounding errors in the nonsampling errors section of this appendix).

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1979 Housing Inventory and for Estimated Number of 1975-1979 Lost Units for the Portland, Oreg.-Wash., SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city
0	100	100	100
100	100	100	100
200	140	140	140
500	230	230	220
700	270	270	260
1,000	320	320	310
2,500	510	500	500
5,000	720	700	700
10,000	1,010	980	980
25,000	1,570	1,470	1,510
50,000	2,160	1,890	2,050
75,000	2,570	2,050	2,400
100,000	2,880	1,990	2,640
150,000	3,300	1,120	2,870
163,400	3,370	—	2,880
200,000	3,520	—	2,830
250,000	3,590	—	2,510
300,000	3,520	—	1,770
335,500	3,370	—	—
400,000	2,870	—	—
498,900	—	—	—

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2 for the total SMSA, 1.3 for the central city, and 1.2 for the balance (not in central city) estimates. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero since these estimates were derived from census data which are not subject to sampling error. However, these estimates are subject to the nonsampling errors associated with the 1970 and 1980 censuses and with the interpolation procedure used to derive these estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1979 Housing Inventory and for Estimated Percentages of 1975-1979 Lost Housing Units for the Portland, Oreg.-Wash., SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
200	34.1	34.1	34.1	34.1	34.1	36.0
500	17.2	17.2	17.2	17.2	19.7	22.8
700	12.9	12.9	12.9	12.9	16.7	19.2
1,000	9.4	9.4	9.4	9.7	13.9	16.1
2,500	4.0	4.0	4.4	6.1	8.8	10.2
5,000	2.0	2.0	3.1	4.3	6.2	7.2
10,000	1.0	1.0	2.2	3.1	4.4	5.1
25,000	0.4	0.6	1.4	1.9	2.8	3.2
50,000	0.2	0.5	1.0	1.4	2.0	2.3
75,000	0.14	0.4	0.8	1.1	1.6	1.9
100,000	0.10	0.3	0.7	1.0	1.4	1.6
150,000	0.07	0.3	0.6	0.8	1.1	1.3
200,000	0.05	0.2	0.5	0.7	1.0	1.1
250,000	0.04	0.2	0.4	0.6	0.9	1.0
300,000	0.03	0.2	0.4	0.6	0.8	0.9
400,000	0.03	0.2	0.4	0.5	0.7	0.8
500,000	0.02	0.14	0.3	0.4	0.6	0.7

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than 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 1.2 for the total SMSA, 1.3 for the central city, and 1.2 for the balance (not in central city).

Raleigh 1979

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-41
Annual Housing Survey App-41
Designation of sample housing units for the 1979 survey. . . . App-41
Selection of the 1976 AHS-SMSA sample App-41
1976-1979 additions to the housing inventory App-43
Sample selection for the 1976 Coverage Improvement Program. App-43
Coverage improvement for deficiency 1 . . . . App-43
Coverage improvement for deficiency 2 . . . . App-43
Coverage improvement for deficiencies 3-6 . . . . App-43
1970 Census of Population and Housing. . . . App-44
ESTIMATION. . . . App-44
1979 housing inventory App-44
1976-1979 lost housing units . . . App-45
1976 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES. . . . App-46
Nonsampling errors. . . . App-46
1970 census App-46
AHS-SMSA App-46
Coverage errors App-46
Rounding errors. . . . App-46
Sampling errors for the AHS-SMSA sample App-47
Illustration of the use of the standard error tables App-47
Differences. . . . App-48
Illustration of the computation of the standard error of a difference App-48
Medians App-48
Illustration of the computation of the 95-percent confidence interval of a median App-49
Standard error tables App-50

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 15 SMSA's in this report series (H-170-79) are based on data collected from the 1979 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 15 SMSA's, the data were collected during the 12-month period from April 1979 through March 1980 with one-twelfth of the sample housing units being visited each month.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1979 were interviewed previously in either 1975 or 1976 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1979 group of SMSA's, the largest SMSA from three of the four census regions of the United States is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All the remaining 12 SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the 1979 group are: Chicago, Ill.; Houston, Tex.; and Seattle-Everett, Wash.

The remaining SMSA's in the 1979 group are: Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.

In this SMSA, 5,620 housing units were eligible for interview. Of these sample housing units, 186 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 5,620 housing units eligible for interview, 619 units were visited but were not eligible for interview because they were condemned, unfit, demolished, converted to group quarters use, etc.

Designation of sample housing units for the 1979 survey—The sample housing units designated to be interviewed in the 1979 survey consisted of the following categories which are described in detail in the succeeding sections.

1. All sample housing units that were interviewed in the 1976 survey.
2. 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 1976 survey. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1979 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1976 survey. (This sample represented the housing units built in permit-issuing areas, since the 1976 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1976 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1976 survey.)
5. All sample housing units that were selected as part of the 1976 Coverage Improvement Program. (This sample represented most of the housing units which, until 1976, did not have a chance of selection.)

Selection of the 1976 AHS-SMSA sample—In 1970, the sample for the SMSA's which were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in

the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (nonpermit universe). In 1979, the following four SMSA's were 100-percent permit-issuing: Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; and Miami, Fla. The remaining 11 SMSA's contain a sample from the nonpermit universe.

Sampling operations, described in the following paragraphs, were performed separately within the central city and the balance of the SMSA for each of the sample frames. The overall sampling rate used to select the sample for each SMSA was determined by the size of the sample. Thus, for the three largest SMSA's, the overall sampling rate differed by central city and the balance of the SMSA, since the sample was divided equally between the central city and the balance of the SMSA. The remaining SMSA's had an overall sampling rate about the same for the sample selected from both the central city and the balance of the SMSA, since the sample was distributed proportionately between the central city and the balance of the SMSA according to the distribution of the total housing units in each sector.

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA 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 housing unit records were stratified by race of head (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 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household 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, for this SMSA, 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 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. 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 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 SMSA. 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.

The second frame from which this SMSA 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 this SMSA. Prior to sample selection, the list of permits was chronologically 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.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS 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 (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

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 are included.

1976-1979 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1976 survey, was selected to represent housing units built in permit-issuing areas since the 1976 survey. Sampling procedures were identical to those used in selecting the 1970-1976 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1976, to identify any housing units missed in the 1976 survey or any housing units added since the 1976 survey.

Sample selection for the 1976 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

1. New construction housing units from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes and trailers placed in parks either missed during the 1970 census or established since the 1970 census.
3. Housing units missed in the 1970 census.
4. Housing units converted to residential use that were nonresidential at the time of the 1970 census.
5. Housing units that have been moved onto their present site since the 1970 census.
6. Mobile homes and trailers placed outside parks since the 1970 census or vacant at the time of the 1970 census.

For each of the five SMSA's interviewed for the first time in 1975, a full implementation of the Coverage Improvement Program was conducted as a part of the 1979 AHS. For each of the 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with some updating and refining as a part of the 1979 AHS. The following discussion of coverage improvement procedures applies to both the 1976 and 1979 programs. For the 10 (1976) SMSA's, estimates of housing units added by a specific procedure reflect units added in 1976 as well as any additions that resulted from the updating and refining in 1979.

Coverage improvement for deficiency 1—A sample of new construction housing units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 1979 SMSA's. Two different procedures were used. For the SMSA's first interviewed in 1976, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. These new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units).

The first stage sample selection was a sample of permit offices, and the second stage a sample of the 1969 permits within each of the selected permit offices. (In the Honolulu, Hawaii, SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second stage.) For the third stage, structures of size three or more were divided into clusters of an expected size of four housing units and a sample of clusters was selected.

For the SMSA's first interviewed in 1975, housing units whose permits were issued before January 1970, but which were completed after April 1970, were identified from the Survey of Construction (SOC), a survey of building permits conducted monthly by the Bureau of the Census. These housing units were then sampled at one-third the rate of regular AHS units. These procedures added an estimated 360 new construction housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiency 2—In permit-issuing areas, a sample of mobile homes and trailers placed in parks that were missed by the census or established after the census was selected in two stages. First, a sample of tracts was selected and canvassed. All parks were listed and then matched back to the 1970 census to identify parks missed by the census and parks established after the census. Second, the parks were divided into clusters of an expected size of four sites and a sample of clusters was selected and interviewed. Each of the sample housing units represented the same number of units that the regular AHS sample housing units represented. Since research showed that there were no mobile home or trailer parks in the central cities of the Chicago, Ill.; Hartford, Conn.; and Portland, Oreg.-Wash.; SMSA's, this procedure was only implemented outside the central cities for these SMSA's. Based on a cost-benefit analysis, this procedure was updated as part of the 1979 AHS only for the Denver, Colo. and Las Vegas, Nev., SMSA's. This procedure added no housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed housing units were sampled by one of two procedures. The first procedure was designed to represent units from the following types of missed structures (structures that had no chance of selection for the AHS):

1. Structures missed in the 1970 census.
2. Structures that were completely nonresidential in the 1970 census but now contain residential housing units.
3. Mobile homes and trailers that had been placed outside parks since the 1970 census and have a utility hookup, or were on the site during the present survey but not occupied on April 1, 1970, or had no utility hookup but were occupied by persons with no usual residence elsewhere.
4. Housing units that had been moved onto their present site since the 1970 census.

Initially, a subsample of regular AHS units was selected at a rate of 1 in 24 for the Chicago, Ill.; Hartford, Conn.; Miami, Fla.; Milwaukee, Wis.; and Portland, Oreg.-Wash.; SMSA's, and at a rate of 1 in 22.2152 for the 10 remaining SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing units were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and housing units within

these structures were interviewed. In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. Although the Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Houston, Tex.; Omaha, Nebr.-Iowa; and Raleigh, N.C.; SMSA's, were interviewed for the first time in 1976 and had these sampling procedures implemented at that time, some 1976 coverage improvement assignments were not sent out to be interviewed due to time limitations. The sampling was completed as part of the 1979 AHS and this procedure added an estimated 2,206 housing units to the coverage of the housing inventory of this SMSA.

The second procedure was designed to represent missed housing units from structures represented in the AHS. These missed housing units were:

1. Housing units missed in the 1970 census.
2. Nonresidential space converted to residential use since the 1970 census in structures that contained some residential housing units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structures selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 377 housing units to the coverage of the housing inventory for this SMSA.

1970 Census of Population and Housing—The estimates pertaining to the 1970 housing inventory (i.e., the housing inventory that existed at the time of the 1970 census) are based on either 20-, 15-, or 5-percent sample data collected in April 1970 for the Decennial Census of Population and Housing. A detailed description of the sample design employed for the 1970 census can be obtained in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

ESTIMATION

The 1979 AHS sample produced two types of estimates for each SMSA: Estimates pertaining to characteristics of the housing inventory at the time of the interview (i.e., the 1979 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1976 (i.e., 1976-1979 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1979 housing inventory—The AHS estimates of characteristics of the 1979 housing inventory were produced using a 2-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and a 3-stage ratio estimation procedure for the remaining 11 SMSA's. Prior to 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 the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of one or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes and trailers from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample housing units from both the nonpermit universe and the coverage improvement universe (if units were not included in any of the previous cells).

The following first-stage 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 permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of housing units enumerated in areas under the jurisdiction of permit-issuing offices. The denominators of the ratios were obtained from weighted estimates of all the AHS sample housing units within the corresponding ratio estimation categories using the existing weight (i.e., the basic weight times the noninterview factor). The computed ratio estimation factor was then applied to the existing weight for each sample housing unit within the corresponding ratio estimation category.

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 permit-issuing universe. Prior to the AHS sample selection within each SMSA, 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 sample. Since the number of housing units deleted from the AHS universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the Houston, Tex., SMSA, a second ratio estimation procedure was employed for new construction housing units outside the Houston central city. Due to rapid growth in the

nonpermit-issuing universe the sampling methods did not adequately represent the new construction outside the Houston central city. Because of this deficiency, separate factors were computed for the following cells:

1. Conventional housing units (non-mobile home or trailer) outside the central city, which were built between the 1970 census and the 1976 survey.
2. Mobile homes and trailers outside the central city with a model year between the 1970 census and 1976 survey.
3. Conventional housing units outside the central city built between the 1976 and 1979 surveys.
4. Mobile homes and trailers outside the central city with a model year between the 1976 and 1979 surveys.

In a specific cell, the ratio estimation factor was equal to the following:

$$\frac{\text{Weighted count of housing units in the cell} + \text{Estimate of the undercoverage in the cell}}{\text{Weighted count of housing units in the cell}}$$

The computed second-stage ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell. For the Baltimore, Md.; Buffalo, N.Y.; Chicago, Ill.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; and Seattle-Everett, Wash.; SMSA's, a second-stage ratio estimation procedure was employed to adjust the central city/balance distribution of the weighted sample estimate of new construction housing units built since the last survey in permit-issuing areas to an independently derived estimate of this distribution.

This ratio estimation factor was calculated separately for the central city and balance of each SMSA and was applied to all new construction housing units from permit-issuing areas within the corresponding sector (central city or balance of the SMSA). This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}{\text{(Sample estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}$$

The independent estimates of new construction were based upon the number of authorized building permits which were determined from the Survey of Construction (SOC). The sample estimates were obtained from the weighted estimate of the AHS-SMSA sample housing units after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

This second-stage ratio estimation procedure was not implemented in the Honolulu, Hawaii; Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's. In the Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C., SMSA's, the central city permit offices, sources of SOC information on permits, had expanded their coverage of permits to include areas greater than those covered by the 1970 census central city definitions. Since the AHS-SMSA survey uses the 1970 definitions, use of this ratio estimation procedure would not correctly adjust the central city/balance distribution of new construction housing units. In the Honolulu, Hawaii, SMSA, this procedure could not be used because the permit office covers the entire SMSA and does not differentiate between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Neb.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and as a third-stage ratio estimation procedure for the remaining 11 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the October 1979 housing inventory of each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sectors. This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the October 1979 housing inventory for the corresponding sector of the SMSA)}}{\text{(AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA)}}$$

The numerator of this ratio was derived using 1970 and 1980 census counts. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight after the first-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and the existing weight after the second-stage ratio estimation procedure for the 11 remaining SMSA's.

The computed ratio estimation factors for the central city and balance of the SMSA's were then applied to the existing weight for all corresponding sample units and the resulting product was used as the final weight for tabulation purposes.

The effect of the total housing unit ratio estimation procedure, as well as the overall 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 SMSA as a whole, it can be expected that the sample estimates will be improved when the sample housing population, or different portions of it, are brought into agreement with known good estimates of the SMSA housing population.

1976-1979 lost housing units—The 1976-1979 lost housing unit (housing unit removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1976 housing inventory, as was described in the 1976 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1976-1979 lost housing units existed, by definition, in the 1976 housing inventory, there was a 1976 housing inventory weight associated with each 1976-1979 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1976-1979 lost housing units.

1976 estimation procedure—This report presents data on the housing characteristics of the 1976 housing inventory from the 1976 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1976.

Ratio estimation procedure of the 1970 Census of Population and Housing—This report presents data on the housing characteristics of the 1970 housing inventory from the 1970 Census of Population and Housing. The statistics based on 1970 census sample data employed a ratio estimation procedure which was applied separately for each of the three census samples. A detailed description of this ratio estimation procedure can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

RELIABILITY OF THE ESTIMATES

There are two types of possible errors associated with estimates based on data from sample surveys—sampling and nonsampling errors. The following is a description of the sampling and nonsampling errors associated with the AHS-SMSA sample and of the nonsampling errors associated with the 1970 census estimates. A description of the sampling errors associated with the sample estimates from the 1970 census can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

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; and other errors of collection, response, processing, 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 1970 Census of Population and Housing and the 1976 AHS-SMSA sample.

1970 census—A number of studies were conducted to measure two types of general errors associated with 1970 census estimates—"coverage" and "content" errors. The "coverage" errors determined how completely housing units were counted in the census and the extent to which occupancy status was erroneously reported. The "content" errors measured the accuracy of the data collected for enumerated housing units. These errors were measured by reinterviews, record checks, and other surveys.

The detailed results of these studies, as well as the methodology employed, are presented in the 1970 Census of Population and Housing Evaluation and Research Program Reports, Series PHC(E)-5, *The Coverage of Housing in the 1970 Census,* and PHC(E)-10, *Accuracy of Data for Selected Housing Characteristics as Measured by Reinterviews.*

AHS-SMSA—Results from the 1979 AHS-SMSA sample reinterview program were not available at the time this report was being prepared. However, a study was conducted for the 1976 AHS-SMSA sample. The results are presented in the Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1976."

Coverage errors—In errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional (non-mobile home or trailer) new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units. Due to 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, in which case, 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, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit.

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 estimated that the 1976 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 for the 1979 survey, the number of missed housing units may be considerably less for 1979.

The final ratio estimation procedure corrects for these deficiencies as far as the count of total housing is concerned; i.e., it adjusts the estimate of the total housing inventory to the best available estimate. However, biases of subtotals would still remain.

Rounding errors—For errors associated with processing, the rounding of estimates introduces another source of error in the data, the severity of which depends on 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-SMSA sample—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 each other. 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 due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by 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, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. 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.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two 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 (page App-50) are approximations to the standard errors of various estimates shown in this report for this SMSA. 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. Standard errors applicable to estimates of characteristics of the 1976 housing inventory can be found in the AHS Series H-170 reports for 1976.

Table I (page App-50) presents the standard errors applicable to estimates of characteristics of the 1979 housing inventory as well as estimates of characteristics of the 1976-1979 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 4,670 for the total SMSA, 3,340 for the central city of the SMSA, and 3,170 for the balance of the SMSA.

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 (page App-50) presents the standard errors of estimated percentages for the 1979 housing inventory as well as estimated percentages of the 1976-1979 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

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 as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is 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:

$$(100)\left(\frac{x}{y}\right)\sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}$$

where: x = the numerator of the ratio
y = the denominator of the ratio
σ_x = the standard error of the numerator
σ_y = the standard error of the denominator

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1979 there were 64,500 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 820. 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
63,200	830
64,500	x
75,000	780

The entry for "x" determined as follows by vertically interpolating between 830 and 780.

$$64{,}500-63{,}200 = 1{,}300$$
$$75{,}000-63{,}200 = 11{,}800$$

$$830 + \frac{1{,}300}{11{,}800}(780-830) = 820$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 63,680 to 65,320 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1979 owner-occupied housing units lies within a range computed in this way, would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 63,188 to 65,812 housing units with 90 percent confidence; and that the average estimate lies within the interval from 62,860 to 66,140 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 64,500 owner-occupied housing units, 12,600, or 19.5 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 19.5 percent is approximately 0.8 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	19.5	25 or 75
50,000	0.7	a	1.0
64,500		p	
75,000	0.5	b	0.8

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

$$19.5-10.0 = 9.5$$
$$25.0-10.0 = 15.0$$

$$0.7 + \frac{9.5}{15.0}(1.0-0.7) = 0.9$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.5 and 0.8.

$$19.5-10.0 = 9.5$$
$$25.0-10.0 = 15.0$$

$$0.5 + \frac{9.5}{15.0}(0.8-0.5) = 0.7$$

3. The entry for "p" is then determined by vertical interpolation between 0.9 and 0.7.

$$64{,}500-50{,}000 = 14{,}500$$
$$75{,}000-50{,}000 = 25{,}000$$

$$0.9 + \frac{14{,}500}{25{,}000}(0.7-0.9) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 18.7 to 20.3 percent; the 90-percent confidence interval is from 18.2 to 20.8 percent; and the 95-percent confidence interval is from 17.9 to 21.1 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. 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. Due to the overlap of the 1976 and 1979 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1979 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1979 there were 34,900 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 22,300. Table I shows the standard error of 12,600 is approximately 520, and the standard error of 34,900 is approximately 760. Therefore, the standard error of the estimated difference of 22,300 is about 920.

$$920 = \sqrt{(520)^2 + (760)^2}$$

Consequently, the 68-percent confidence interval for the 22,300 difference is from 21,380 to 23,220 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 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 20,830 to 23,770 housing units, and the 95-percent confidence interval is from 20,460 to 24,140 housing units. Thus, we can conclude with 95 percent confidence that the number of 1979 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied housing units with two bedrooms, since the 95-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 deter-

mine an interval about the estimated median so that there is a stated degree of confidence 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. These two distribution intervals could be different.

For about 68 out of 100 possible samples, the average median from all possible samples would lie between these two values.

A two-standard-error confidence interval may be determined by finding the values corresponding to 50 percent plus and minus twice the standard error determined in step 1. For about 95 out of 100 possible samples, the average median from all possible samples would lie between these two values.

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.8. The base of the distribution from which this median was determined is 64,500 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 64,500 is approximately 1.0 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.0 and 52.0.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with 3 persons (for purposes of calculating the median, the category of 3 persons is considered to be from 2.5 to 3.5 persons) contains the 48.0 percent derived in step 2. About 27,300 housing units or 42.3 percent fall below this interval, and 14,000 housing units or 21.7 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\frac{(48.0-42.3)}{21.7} = 2.8$$

Similarly, the interval for owner-occupied housing units with 3 persons corresponds to the 52.0 percent derived in step 2. About 27,300 housing units or 42.3 percent fall below this interval, and 14,000 housing units or 21.7 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\frac{(52.0-42.3)}{21.7} = 2.9$$

Thus, the 95-percent confidence interval ranges from 2.8 to 2.9 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1979 Housing Inventory and for Estimated Number of 1976-1979 Lost Units for the Raleigh, N.C., SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city
0	30	20	20
100	50	50	50
200	70	70	70
500	110	110	110
700	130	130	130
1,000	160	150	150
2,500	250	240	240
5,000	350	330	330
10,000	480	440	450
25,000	700	540	600
48,000	830	–	530
50,000	830	–	500
63,200	830	–	–
75,000	780	–	–
100,000	500	–	–
111,300	–	–	–

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.0 for the total SMSA, 1.1 for the central city, and 1.0 for the balance (not in central city) estimates. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero since these estimates were derived from census data which are not subject to sampling error. However, these estimates are subject to the nonsampling errors associated with the 1970 and 1980 censuses and with the interpolation procedures used to derived these estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1979 Housing Inventory and for Estimated Percentages of 1976-1979 Lost Housing Units for the Raleigh, N.C., SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	20.0	20.0	20.0	20.0	21.7	25.0
200	11.1	11.1	11.1	11.1	15.3	17.7
500	4.8	4.8	4.9	6.7	9.7	11.2
700	3.4	3.4	4.1	5.7	8.2	9.5
1,000	2.4	2.4	3.4	4.7	6.8	7.9
2,500	1.0	1.0	2.2	3.0	4.3	5.0
5,000	0.5	0.7	1.5	2.1	3.1	3.5
10,000	0.2	0.5	1.1	1.5	2.2	2.5
25,000	0.10	0.3	0.7	0.9	1.4	1.6
50,000	0.05	0.2	0.5	0.7	1.0	1.1
75,000	0.03	0.2	0.4	0.5	0.8	0.9
100,000	0.03	0.2	0.3	0.5	0.7	0.8
150,000	0.02	0.13	0.3	0.4	0.6	0.6

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than 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 1.0 for the total SMSA, 1.1 for the central city, and 1.0 for the balance (not in central city).

Seattle-Everett 1979

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-41
Annual Housing Survey . . . App-41
Designation of sample housing units for the 1979 survey . . App-41
Selection of the 1976 AHS-SMSA sample . . App-41
1976-1979 additions to the housing inventory . . App-43
Sample selection for the 1976 Coverage Improvement Program . App-43
Coverage improvement for deficiency 1 . . App-43
Coverage improvement for deficiency 2 . . App-43
Coverage improvement for deficiencies 3-6 . . App-43
1970 Census of Population and Housing . . App-44
ESTIMATION . . App-44
1979 housing inventory . . App-44
1976-1979 lost housing units . . App-45
1976 estimation procedure . . App-46
Ratio estimation procedure of the 1970 Census of Population and Housing . . App-46
RELIABILITY OF THE ESTIMATES . . App-46
Nonsampling errors . . App-46
1970 census . . App-46
AHS-SMSA . . App-46
Coverage errors . . App-46
Rounding errors . . App-46
Sampling errors for the AHS-SMSA sample . . App-47
Illustration of the use of the standard error tables . . App-47
Differences . . App-48
Illustration of the computation of the standard error of a difference . . App-48
Medians . . App-48
Illustration of the computation of the 95-percent confidence interval of a median . . App-49
Standard error tables . . App-49

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 15 SMSA's in this report series (H-170-79) are based on data collected from the 1979 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 15 SMSA's, the data were collected during the 12-month period from April 1979 through March 1980 with one-twelfth of the sample housing units being visited each month.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1979 were interviewed previously in either 1975 or 1976 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1979 group of SMSA's, the largest SMSA from three of the four census regions of the United States is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All the remaining 12 SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the 1979 group are: Chicago, Ill.; Houston, Tex.; and Seattle-Everett, Wash.

The remaining SMSA's in the 1979 group are: Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.

In this SMSA, 16,485 housing units were eligible for interview. Of these sample housing units, 797 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 16,485 housing units eligible for interview, 1,171 units were visited but were not eligible for interview because they were condemned, unfit, demolished, converted to group quarters use, etc.

Designation of sample housing units for the 1979 survey—The sample housing units designated to be interviewed in the 1979 survey consisted of the following categories which are described in detail in the succeeding sections.

1. All sample housing units that were interviewed in the 1976 survey.
2. 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 1976 survey. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1979 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1976 survey. (This sample represented the housing units built in permit-issuing areas, since the 1976 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1976 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1976 survey.)
5. All sample housing units that were selected as part of the 1976 Coverage Improvement Program. (This sample represented most of the housing units which, until 1979, did not have a chance of selection.)

Selection of the 1976 AHS-SMSA sample—In 1970, the sample for the SMSA's which were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in

the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). In 1979, the following four SMSA's were 100-percent permit-issuing: Hartford, Conn.; Honolulu, Hawaii; Las Vegas, Nev.; and Miami, Fla. The remaining 11 SMSA's contain a sample from the nonpermit universe.

Sampling operations, described in the following paragraphs, were performed separately within the central city and the balance of the SMSA for each of the sample frames. The overall sampling rate used to select the sample for each SMSA was determined by the size of the sample. Thus, for the three largest SMSA's, the overall sampling rate differed by central city and the balance of the SMSA, since the sample was divided equally between the central city and the balance of the SMSA. The remaining SMSA's had an overall sampling rate about the same for the sample selected from both the central city and the balance of the SMSA, since the sample was distributed proportionately between the central city and the balance of the SMSA according to the distribution of the total housing units in each sector.

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA 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 housing unit records were stratified by race of head (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 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household 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, for this SMSA, 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 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. 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 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 SMSA. 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.

The second frame from which this SMSA 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 this SMSA. Prior to sample selection, the list of permits was chronologically 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.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS 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 (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

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 are included.

1976-1979 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1976 survey, was selected to represent housing units built in permit-issuing areas since the 1976 survey. Sampling procedures were identical to those used in selecting the 1970-1976 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1976, to identify any housing units missed in the 1976 survey or any housing units added since the 1976 survey.

Sample selection for the 1976 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

1. New construction housing units from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes and trailers placed in parks either missed during the 1970 census or established since the 1970 census.
3. Housing units missed in the 1970 census.
4. Housing units converted to residential use that were nonresidential at the time of the 1970 census.
5. Housing units that have been moved onto their present site since the 1970 census.
6. Mobile homes and trailers placed outside parks since the 1970 census or vacant at the time of the 1970 census.

For each of the five SMSA's interviewed for the first time in 1975, a full implementation of the Coverage Improvement Program was conducted as a part of the 1979 AHS. For each of the 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with some updating and refining as a part of the 1979 AHS. The following discussion of coverage improvement procedures applies to both the 1976 and 1979 programs. For the 10 (1976) SMSA's, estimates of housing units added by a specific procedure reflect units added in 1976 as well as any additions that resulted from the updating and refining in 1979.

Coverage improvement for deficiency 1—A sample of new construction housing units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 1979 SMSA's. Two different procedures were used. For the SMSA's first interviewed in 1976, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. These new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units).

The first stage sample selection was a sample of permit offices, and the second stage a sample of the 1969 permits within each of the selected permit offices. (In the Honolulu, Hawaii, SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second stage.) For the third stage, structures of size three or more were divided into clusters of an expected size of four housing units and a sample of clusters was selected.

For the SMSA's first interviewed in 1975, housing units whose permits were issued before January 1970, but which were completed after April 1970, were identified from the Survey of Construction (SOC), a survey of building permits conducted monthly by the Bureau of the Census. These housing units were then sampled at one-third the rate of regular AHS units. These procedures added an estimated 5,087 new construction housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiency 2—In permit-issuing areas, a sample of mobile homes and trailers placed in parks that were missed by the census or established after the census was selected in two stages. First, a sample of tracts was selected and canvassed. All parks were listed and then matched back to the 1970 census to identify parks missed by the census and parks established after the census. Second, the parks were divided into clusters of an expected size of four sites and a sample of clusters was selected and interviewed. Each of the sample housing units represented the same number of units that the regular AHS sample housing units represented. Since research showed that there were no mobile home or trailer parks in the central cities of the Chicago, Ill.; Hartford, Conn.; and Portland, Oreg.-Wash.; SMSA's, this procedure was only implemented outside the central cities for these SMSA's. Based on a cost-benefit analysis, this procedure was updated as part of the 1979 AHS only for the Denver, Colo. and Las Vegas, Nev., SMSA's. This procedure added 687 housing units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed housing units were sampled by one of two procedures. The first procedure was designed to represent units from the following types of missed structures (structures that had no chance of selection for the AHS):

1. Structures missed in the 1970 census.
2. Structures that were completely nonresidential in the 1970 census but now contain residential housing units.
3. Mobile homes and trailers that had been placed outside parks since the 1970 census and have a utility hookup, or were on the site during the present survey but not occupied on April 1, 1970, or had no utility hookup but were occupied by persons with no usual residence elsewhere.
4. Housing units that had been moved onto their present site since the 1970 census.

Initially, a subsample of regular AHS units was selected at a rate of 1 in 24 for the Chicago, Ill.; Hartford, Conn.; Miami, Fla.; Milwaukee, Wis.; and Portland, Oreg.-Wash.; SMSA's, and at a rate of 1 in 22.2152 for the 10 remaining SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing units were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and housing units within

these structures were interviewed. In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. Although the Baltimore, Md.; Buffalo, N.Y.; Cleveland, Ohio; Houston, Tex.; Omaha, Nebr.-Iowa; and Raleigh, N.C.; SMSA's, were interviewed for the first time in 1976 and had these sampling procedures implemented at that time, some 1976 coverage improvement assignments were not sent out to be interviewed due to time limitations. The sampling was completed as part of the 1979 AHS and this procedure added an estimated 17,186 housing units to the coverage of the housing inventory of this SMSA.

The second procedure was designed to represent missed housing units from structures represented in the AHS. These missed housing units were:

1. Housing units missed in the 1970 census.
2. Nonresidential space converted to residential use since the 1970 census in structures that contained some residential housing units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structures selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 1,149 housing units to the coverage of the housing inventory for this SMSA.

1970 Census of Population and Housing—The estimates pertaining to the 1970 housing inventory (i.e., the housing inventory that existed at the time of the 1970 census) are based on either 20-, 15-, or 5-percent sample data collected in April 1970 for the Decennial Census of Population and Housing. A detailed description of the sample design employed for the 1970 census can be obtained in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

ESTIMATION

The 1979 AHS sample produced two types of estimates for each SMSA: Estimates pertaining to characteristics of the housing inventory at the time of the interview (i.e., the 1979 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1976 (i.e., 1976-1979 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1979 housing inventory—The AHS estimates of characteristics of the 1979 housing inventory were produced using a 2-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and a 3-stage ratio estimation procedure for the remaining 11 SMSA's. Prior to 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 the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of one or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes and trailers from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample housing units from both the nonpermit universe and the coverage improvement universe (if units were not included in any of the previous cells).

The following first-stage 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 permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of housing units enumerated in areas under the jurisdiction of permit-issuing offices. The denominators of the ratios were obtained from weighted estimates of all the AHS sample housing units within the corresponding ratio estimation categories using the existing weight (i.e., the basic weight times the noninterview factor). The computed ratio estimation factor was then applied to the existing weight for each sample housing unit within the corresponding ratio estimation category.

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 permit-issuing universe. Prior to the AHS sample selection within each SMSA, 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 sample. Since the number of housing units deleted from the AHS universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the Houston, Tex., SMSA, a second ratio estimation procedure was employed for new construction housing units outside the Houston central city. Due to rapid growth in the

nonpermit-issuing universe the sampling methods did not adequately represent the new construction outside the Houston central city. Because of this deficiency, separate factors were computed for the following cells:

1. Conventional housing units (non-mobile home or trailer) outside the central city, which were built between the 1970 census and the 1976 survey.
2. Mobile homes and trailers outside the central city with a model year between the 1970 census and 1976 survey.
3. Conventional housing units outside the central city built between the 1976 and 1979 surveys.
4. Mobile homes and trailers outside the central city with a model year between the 1976 and 1979 surveys.

In a specific cell, the ratio estimation factor was equal to the following:

$$\frac{\text{Weighted count of housing units in the cell} + \text{Estimate of the undercoverage in the cell}}{\text{Weighted count of housing units in the cell}}$$

The computed second-stage ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell. For the Baltimore, Md.; Buffalo, N.Y.; Chicago, Ill.; Cleveland, Ohio; Denver, Colo.; Hartford, Conn.; Las Vegas, Nev.; Miami, Fla.; Milwaukee, Wis.; and Seattle-Everett, Wash.; SMSA's, a second-stage ratio estimation procedure was employed to adjust the central city/balance distribution of the weighted sample estimate of new construction housing units built since the last survey in permit-issuing areas to an independently derived estimate of this distribution.

This ratio estimation factor was calculated separately for the central city and balance of each SMSA and was applied to all new construction housing units from permit-issuing areas within the corresponding sector (central city or balance of the SMSA). This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}{\text{(Sample estimate of the proportion of new construction housing units from permit-issuing areas built since the last survey in the corresponding sector of the SMSA)}}$$

The independent estimates of new construction were based upon the number of authorized building permits which were determined from the Survey of Construction (SOC). The sample estimates were obtained from the weighted estimate of the AHS-SMSA sample housing units after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

This second-stage ratio estimation procedure was not implemented in the Honolulu, Hawaii; Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's. In the Houston, Tex.; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C., SMSA's, the central city permit offices, sources of SOC information on permits, had expanded their coverage of permits to include areas greater than those covered by the 1970 census central city definitions. Since the AHS-SMSA survey uses the 1970 definitions, use of this ratio estimation procedure would not correctly adjust the central city/balance distribution of new construction housing units. In the Honolulu, Hawaii, SMSA, this procedure could not be used because the permit office covers the entire SMSA and does not differentiate between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Neb.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and as a third-stage ratio estimation procedure for the remaining 11 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the October 1979 housing inventory of each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sectors. This ratio estimation factor equaled the following:

$$\frac{\text{(Independent estimate of the October 1979 housing inventory for the corresponding sector of the SMSA)}}{\text{(AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA)}}$$

The numerator of this ratio was derived using 1970 and 1980 census counts. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight after the first-stage ratio estimation procedure for the Honolulu, Hawaii; Omaha, Nebr.-Iowa; Portland, Oreg.-Wash.; and Raleigh, N.C.; SMSA's, and the existing weight after the second-stage ratio estimation procedure for the 11 remaining SMSA's.

The computed ratio estimation factors for the central city and balance of the SMSA's were then applied to the existing weight for all corresponding sample units and the resulting product was used as the final weight for tabulation purposes.

The effect of the total housing unit ratio estimation procedure, as well as the overall 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 SMSA as a whole, it can be expected that the sample estimates will be improved when the sample housing population, or different portions of it, are brought into agreement with known good estimates of the SMSA housing population.

1976-1979 lost housing units—The 1976-1979 lost housing unit (housing unit removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1976 housing inventory, as was described in the 1976 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1976-1979 lost housing units existed, by definition, in the 1976 housing inventory, there was a 1976 housing inventory weight associated with each 1976-1979 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1976-1979 lost housing units.

1976 estimation procedure—This report presents data on the housing characteristics of the 1976 housing inventory from the 1976 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1976.

Ratio estimation procedure of the 1970 Census of Population and Housing—This report presents data on the housing characteristics of the 1970 housing inventory from the 1970 Census of Population and Housing. The statistics based on 1970 census sample data employed a ratio estimation procedure which was applied separately for each of the three census samples. A detailed description of this ratio estimation procedure can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

RELIABILITY OF THE ESTIMATES

There are two types of possible errors associated with estimates based on data from sample surveys—sampling and nonsampling errors. The following is a description of the sampling and nonsampling errors associated with the AHS-SMSA sample and of the nonsampling errors associated with the 1970 census estimates. A description of the sampling errors associated with the sample estimates from the 1970 census can be found in the 1970 Census of Housing report, Volume I, *Housing Characteristics for States, Cities, and Counties,* Part 1.

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; and other errors of collection, response, processing, 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 1970 Census of Population and Housing and the 1976 AHS-SMSA sample.

1970 census—A number of studies were conducted to measure two types of general errors associated with 1970 census estimates—"coverage" and "content" errors. The "coverage" errors determined how completely housing units were counted in the census and the extent to which occupancy status was erroneously reported. The "content" errors measured the accuracy of the data collected for enumerated housing units. These errors were measured by reinterviews, record checks, and other surveys.

The detailed results of these studies, as well as the methodology employed, are presented in the 1970 Census of Population and Housing Evaluation and Research Program Reports, Series PHC(E)-5, *The Coverage of Housing in the 1970 Census,* and PHC(E)-10, *Accuracy of Data for Selected Housing Characteristics as Measured by Reinterviews.*

AHS-SMSA—Results from the 1979 AHS-SMSA sample reinterview program were not available at the time this report was being prepared. However, a study was conducted for the 1976 AHS-SMSA sample. The results are presented in the Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1976."

Coverage errors—In errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional (non-mobile home or trailer) new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units. Due to 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, in which case, 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, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit.

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 estimated that the 1976 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 for the 1979 survey, the number of missed housing units may be considerably less for 1979.

The final ratio estimation procedure corrects for these deficiencies as far as the count of total housing is concerned; i.e., it adjusts the estimate of the total housing inventory to the best available estimate. However, biases of subtotals would still remain.

Rounding errors—For errors associated with processing, the rounding of estimates introduces another source of error in the data, the severity of which depends on 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-SMSA sample—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 each other. 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 due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by 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, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. 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.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two 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 (page App-49) are approximations to the standard errors of various estimates shown in this report for this SMSA. 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. Standard errors applicable to estimates of characteristics of the 1976 housing inventory can be found in the AHS Series H-170 reports for 1976.

Table I (page App-49) presents the standard errors applicable to estimates of characteristics of the 1979 housing inventory as well as estimates of characteristics of the 1976-1979 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 9,300 for the total SMSA, 5,380 for the central city of the SMSA, and 7,570 for the balance of the SMSA.

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.

Tables II and III (page App-50) present the standard errors of estimated percentages for the 1979 housing inventory as well as estimated percentages of the 1976-1979 lost housing (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in tables II and III.

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

For ratios, 100 (x/y), where x is not a subclass of y, tables II and III underestimate 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:

$$(100)\left(\frac{x}{y}\right)\sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}$$

where: x = the numerator of the ratio
y = the denominator of the ratio
σ_x = the standard error of the numerator
σ_y = the standard error of the denominator

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1979 there were 393,500 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 2,710. 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
300,000	2,770
393,500	x
396,000	2,710

The entry for "x" is determined as follows by vertically interpolating between 2,770 and 2,710.

$$393{,}500-300{,}000 = 93{,}500$$
$$396{,}000-300{,}000 = 96{,}000$$

$$2{,}770 + \frac{93{,}500}{96{,}000}(2{,}710-2{,}770) = 2{,}710$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 390,790 to 396,210 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1979 owner-occupied housing units lies within a range computed in this way, would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 389,160 to 397,840 housing units with 90 percent confidence; and that the average estimate lies within the interval from 388,080 to 398,920 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 393,500 owner-occupied housing units, 96,800, or 24.6 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 24.6 percent is approximately 0.5 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	24.6	25 or 75
300,000.	0.4	a	0.6
393,500.		p	
400,000.	0.3	b	0.5

1. The entry for cell "a" is determined by horizontal interpolation between 0.4 and 0.6.

$$24.6-10.0 = 14.6$$
$$25.0-10.0 = 15.0$$

$$0.4 + \frac{14.6}{15.0}(0.6-0.4) = 0.6$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.3 and 0.5.

$$24.6-10.0 = 14.6$$
$$25.0-10.0 = 15.0$$

$$0.3 + \frac{14.6}{15.0}(0.5-0.3) = 0.5$$

3. The entry for "p" is then determined by vertical interpolation between 0.6 and 0.5.

$$393{,}500-300{,}000 = 93{,}500$$
$$400{,}000-300{,}000 = 100{,}000$$

$$0.6 + \frac{93{,}500}{100{,}000}(0.5-0.6) = 0.5$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 24.1 to 25.1 percent; the 90-percent confidence interval is from 23.8 to 25.4 percent; and the 95-percent confidence interval is from 23.6 to 25.6 percent.

Differences—The standard errors shown are not directly applicable to the differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. 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. Due to the overlap of the 1976 and 1979 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1979 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1979 there were 174,000 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 77,200. Table I shows the standard error of 174,000 is approximately 2,450, and the standard error of 96,800 is approximately 1,980. Therefore, the standard error of the estimated difference of 77,200 is about 3,150.

$$3{,}150 = \sqrt{(2{,}450)^2 + (1{,}980)^2}$$

Consequently, the 68-percent confidence interval for the 77,200 difference is from 74,050 to 80,350 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 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 72,160 to 82,240 housing units, and the 95-percent confidence interval is from 70,900 to 83,500 housing units. Thus, we can conclude with 95-percent confidence that the number of 1979 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied housing units with two bedrooms, since the 95-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 deter-

mine an interval about the estimated median so that there is a stated degree of confidence 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 tables II and III, 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. These two distribution intervals could be different.

For about 68 out of 100 possible samples, the average median from all possible samples would lie between these two values.

A two-standard-error confidence interval may be determined by finding the values corresponding to 50 percent plus and minus twice the standard error determined in step 1. For about 95 out of 100 possible samples, the average median from all possible samples would lie between these two values.

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.5. The base of the distribution from which this median was determined is 393,500 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 393,500 is approximately 0.6 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.8 and 51.2.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with 2 persons (for purposes of calculating the median, the category of 2 persons is considered to be from 1.5 to 2.5 persons) contains the 48.8 percent derived in step 2. About 58,900 housing units or 15.0 percent fall below this interval, and 134,700 housing units or 34.2 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$1.5 + (2.5-1.5)\,\frac{(48.8-15.0)}{34.2} = 2.5$$

Similarly, the interval for owner-occupied housing units with 3 persons contains the 51.2 percent derived in step 2. About 193,600 housing units or 49.2 percent fall below this interval, and 70,000 housing units or 17.8 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(51.2-49.2)}{17.8} = 2.6$$

Thus, the 95-percent confidence interval ranges from 2.5 to 2.6 persons. Although it appears that this confidence interval has the sample estimate as the lower limit, it actually is a reflection of the rounding error associated with this median (see the paragraph on rounding error in the nonsampling error section of this appendix).

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1979 Housing Inventory and for Estimated Number of 1976-1979 Lost Units for the Seattle-Everett, Wash., SMSA, for the Central Cities of the SMSA and for the Balance (Not in Central Cities) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]		
	SMSA	In central cities	Not in central cities
0	50	40	50
100	70	60	70
200	100	90	100
500	150	140	160
700	180	170	190
1,000	220	200	230
2,500	340	320	360
5,000	490	450	510
10,000	680	620	720
25,000	1,070	960	1,120
50,000	1,480	1,280	1,520
75,000	1,780	1,470	1,800
100,000	2,010	1,570	1,990
150,000	2,340	1,580	2,230
200,000	2,570	1,310	2,300
250,000	2,710	380	2,210
253,600	2,710	—	2,200
300,000	2,770	—	1,970
396,000	2,710	—	—
400,000	2,700	—	—
500,000	2,340	—	—
600,000	1,480	—	—
649,600	—	—	—

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2 for the total SMSA, 1.3 for the central cities, and 1.1 for the balance (not in central cities) estimates. Standard errors of estimates pertaining to total housing units for the central cities, balance, and total SMSA are assumed to be equal to zero since these estimates were derived from census data which are not subject to sampling error. However, these estimates are subject to the nonsampling errors associated with the 1970 and 1980 censuses and with the interpolation procedures used to derive these estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1979 Housing Inventory and for Estimated Percentages of 1976-1979 Lost Housing Units for the Seattle-Everett, Wash., SMSA, and for the Balance (Not in Central Cities) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
200	21.0	21.0	21.0	21.0	22.3	25.8
500	9.6	9.6	9.6	9.8	14.1	16.3
700	7.1	7.1	7.1	8.3	11.9	13.8
1,000	5.1	5.1	5.1	6.9	10.0	11.5
2,500	2.1	2.1	3.2	4.4	6.3	7.3
5,000	1.1	1.1	2.2	3.1	4.5	5.2
10,000	0.5	0.7	1.6	2.2	3.2	3.6
25,000	0.2	0.5	1.0	1.4	2.0	2.3
50,000	0.11	0.3	0.7	1.0	1.4	1.6
75,000	0.07	0.3	0.6	0.8	1.2	1.3
100,000	0.05	0.2	0.5	0.7	1.0	1.2
150,000	0.04	0.2	0.4	0.6	0.8	0.9
200,000	0.03	0.2	0.4	0.5	0.7	0.8
250,000	0.02	0.15	0.3	0.4	0.6	0.7
300,000	0.02	0.13	0.3	0.4	0.6	0.7
400,000	0.01	0.11	0.3	0.3	0.5	0.6
500,000	0.01	0.10	0.2	0.3	0.4	0.5
600,000	0.01	0.09	0.2	0.3	0.4	0.5
700,000	0.01	0.09	0.2	0.3	0.4	0.4

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than 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 1.2 for the total SMSA, and 1.1 for the balance (not in central cities).

TABLE III. Standard Errors for Estimated Percentages of Housing Units in the 1979 Housing Inventory and for Estimated Percentages of 1976-1979 Lost Housing Units for the Central Cities of the Seattle-Everett, Wash., SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
200	16.9	16.9	16.9	16.9	19.5	22.5
500	7.5	7.5	7.5	8.6	12.3	14.3
700	5.5	5.5	5.5	7.2	10.4	12.1
1,000	3.9	3.9	4.4	6.0	8.7	10.1
2,500	1.6	1.6	2.8	3.8	5.5	6.4
5,000	0.8	0.9	2.0	2.7	3.9	4.5
10,000	0.4	0.6	1.4	1.9	2.8	3.2
25,000	0.2	0.4	0.9	1.2	1.7	2.0
50,000	0.08	0.3	0.6	0.9	1.2	1.4
75,000	0.05	0.2	0.5	0.7	1.0	1.2
100,000	0.04	0.2	0.4	0.6	0.9	1.0
150,000	0.03	0.2	0.4	0.5	0.7	0.8
200,000	0.02	0.14	0.3	0.4	0.6	0.7
250,000	0.02	0.13	0.3	0.4	0.6	0.6
300,000	0.01	0.12	0.3	0.3	0.5	0.6

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than 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 1.3 for the central cities of the SMSA.