1980 Metropolitan Appendix B: Source and Accuracy

Table of Contents

Albany-Schenectady-Troy 3

Allentown-Bethlehem-Easton 14

Birmingham 25

Grand Rapids 36

Indianapolis 47

Los Angeles-Long Beach 58

Louisville 70

Memphis 81

New York 92

Oklahoma City 104

Providence-Pawtucket-Warwick 115

Sacramento 126

Saginaw 137

Salt Lake City 148

St. Louis 159

Albany-Schenectady-Troy 1980

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1980 survey. App-42
Selection of the 1977 AHS-SMSA sample App-43
1977-1980 additions to the housing inventory App-44
Sample selection for the 1977 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
1980 housing inventory App-45
1977-1980 lost housing units . . . App-46
1977 estimation procedure App-46
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-47
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-170-80) are based on data collected from the 1980 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 the Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill., SMSA's, the data were collected during the 11-month period from April 1980 through February 1981 with one-eleventh of the sample housing units being visited each month. In the remaining SMSA's, the data were collected during the 12-month period from April 1980 through March 1981 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 1980 were interviewed previously in either 1976 or 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1980 group of SMSA's, the largest SMSA from 3 of the 4 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 1980 group are: Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill.

The remaining SMSA's in the 1980 group are: Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Birmingham, Ala.; Grand Rapids, Mich.; Indianapolis, Ind.; Louisville, Ky.-Ind.; Memphis, Tenn.-Ark.; Oklahoma City, Okla.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; and Salt Lake City, Utah.

In this SMSA, 5,242 housing units were eligible for interview. Of these sample housing units, 116 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 5,242 housing units eligible for interview, 494 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 1980 survey—The sample housing units designated to be interviewed in the 1980 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 1977 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 1977 survey. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1980 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 1977 survey. (This sample represented the housing units built in permit-issuing areas, since the 1977 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1977 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1977 survey.)
5. All sample housing units that were selected as part of the 1977 and 1980 Coverage Improvement Programs. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1977 AHS-SMSA sample—The sample for the SMSA's which, in 1970, 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 1970, the following three SMSA's were 100-percent permit-issuing: Los Angeles-Long Beach, Calif.; New York, N.Y.; and Sacramento, Calif. The remaining 12 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.

1977-1980 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1977 survey, was selected to represent housing units built in permit-issuing areas since the 1977 survey. Sampling procedures were identical to those used in selecting the 1970-1977 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1977, to identify any housing units missed in the 1977 survey or any housing units added since the 1977 survey.

Sample selection for the 1977 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 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with the Oklahoma City, Okla. SMSA receiving some updating as a part of the 1980 AHS. For each of the five SMSA's interviewed previously in 1977, the Coverage Improvement Program was conducted as a part of the 1977 AHS with the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; and Salt Lake City, Utah, SMSA's receiving some updating and refining as a part of the 1980 AHS. The following discussion applies to both the prior year (1976 or 1977) and 1980 coverage improvement procedures. For the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Oklahoma City, Okla.; and Salt Lake City, Utah, SMSA's estimates of housing units added by a specific procedure reflect units added in the prior year as well as any additions that resulted from the updating and refining in 1980.

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 1980 SMSA's. Two different procedures were used. For the first procedure, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. For the SMSA's previously interviewed in 1976, these new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units). For the SMSA's previously interviewed in 1977, sample units selected from one- and two-unit structures were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample, while units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage sample selection was a sample of permit offices, and the second stage was a sample of the 1969 permits within each of the selected permit offices. In the Grand Rapids, Mich., and New York, N.Y., SMSA's, 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. This procedure was employed for the 10 SMSA's previously interviewed in 1976.

For the Albany-Schenectady-Troy, N.Y.; Memphis, Tenn.-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's previously interviewed in 1977, the above procedures were performed for the first and second stages. For the third stage, structures of size three or more were divided into clusters of an expected size of two units and a sample of clusters was selected.

For the Los Angeles-Long Beach, Calif., SMSA, 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 units were then sampled at one-third the rate of regular AHS units. Since permits were not available for all sampled offices, the procedure was also used in parts of the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 903 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. In the Oklahoma City, Okla., SMSA, a sample of tracts was selected but not canvassed during the 1976 Coverage Improvement Program. This procedure was completed during the 1980 AHS and added 166 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 AHS sample housing units was selected from the 1970 sample universe at a rate of 1 in 24 for the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Memphis, Tenn-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's and at a rate of 1 in 22 for the remaining 10 SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing unit 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. The Albany-Schenectady-Troy, N.Y., SMSA, had these sampling procedures implemented in 1976, excluding identification of the housing units missed in the 1970 census. This procedure was completed as a part of the 1980 AHS. The Los Angeles-Long Beach, Calif., and Salt Lake City, Utah, SMSA's had this sampling procedure implemented for the first time in 1980. This procedure added an estimated 3,674 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 1,133 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 1980 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 1980 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1977 (i.e., 1977-1980 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1980 housing inventory—The AHS estimates of characteristics of the 1980 housing inventory were produced using a 2-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Grand Rapids, Mich.; Indianapolis, Ind.; Los Angeles-Long Beach, Calif.; Louisville, Ky.-Ind.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; St. Louis, Mo.-Ill.; and Salt Lake City, Utah, 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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., SMSA's. In the Birmingham, Ala.; Memphis, Tenn.-Ark.; and Oklahoma City, Okla., 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 New York, N.Y., SMSA, the sample selected correctly apportioned the new construction housing units between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 1980 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:

$$\frac{\text{Independent estimate of the October 1980 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 procedure, 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.

1977-1980 lost housing units—The 1977-1980 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 1977 housing inventory, as was described in the 1977 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1977-1980 lost housing units existed, by definition, in the 1977 housing inventory, there was a 1977 housing inventory weight associated with each 1977-1980 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1977-1980 lost housing units.

1977 estimation procedure—This report presents data on the housing characteristics of the 1977 housing inventory from the 1977 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 1977.

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 1977 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 1980 AHS-SMSA sample reinterview program were not available at the time this report was being prepared. However, a study was conducted for the 1977 AHS-SMSA sample. The results of which are presented in the Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1977."

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 1977 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 1980 survey, the number of missed housing units may be considerably less for 1980.

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

Rounding errors—For errors associated with processing, 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 1977 housing inventory can be found in the AHS Series H-170 reports for 1977.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1980 housing inventory as well as estimates of characteristics of the 1977-1980 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 11,780 for the total SMSA, 5,040 for the central city of the SMSA, and 10,550 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 1980 housing inventory as well as estimated percentages of the 1977-1980 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 1980 there were 176,200 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 2,000. 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
150,000	2,010
176,200	x
196,200	1,990

The entry for "x" is determined as follows by vertically interpolating between 2,010 and 1,990.

$$176{,}200-150{,}000 = 26{,}200$$
$$196{,}200-150{,}000 = 46{,}200$$

$$2{,}010 + \frac{26{,}200}{46{,}200}(1{,}990-2{,}010) = 2{,}000$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 174,200 to 178,200 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1980 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 173,000 to 179,400 housing units with 90 percent confidence; and that the average estimate lies within the interval from 172,200 to 180,200 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 176,200 owner-occupied housing units, 37,900, or 21.5 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 21.5 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	21.5	25 or 75
150,000.	0.6	a	0.9
176,200.		p	
200,000.	0.5	b	0.8

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

$$21.5-10.0 = 11.5$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{11.5}{15.0}(0.9-0.6) = 0.8$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.5 and 0.8.

$$21.5-10.0 = 11.5$$
$$25.0-10.0 = 15.0$$

$$0.5 + \frac{11.5}{15.0}(0.8-0.5) = 0.7$$

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

$$176{,}200-150{,}000 = 26{,}200$$
$$200{,}000-150{,}000 = 50{,}000$$

$$0.8 + \frac{26{,}200}{50{,}000}(0.7-0.8) = 0.7$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 20.8 to 22.2 percent; the 90-percent confidence interval is from 20.4 to 22.6 percent; and the 95-percent confidence interval is from 20.1 to 22.9 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; if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1976 and 1980 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1980 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1980 there were 87,200 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 49,300. Table I shows the standard error of 37,900 is approximately 1,380, and the standard error of 87,200 is approximately 1,920. Therefore, the standard error of the estimated difference of 49,300 is about 2,360:

$$2{,}360 = \sqrt{(1{,}380)^2 + (1{,}920)^2}$$

Consequently, the 68-percent confidence interval for the 49,300 difference is from 46,940 to 51,660 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 45,520 to 53,080 housing units, and the 95-percent confidence interval is from 44,580 to 54,020 housing units. Thus, we can conclude with 95 percent confidence that the number of 1980 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 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 176,200 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 176,200 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 three persons (for purposes of calculating the median, the category of three persons is considered to be from 2.5 to 3.5 persons) contains the 48.2 percent derived in step 2. About 81,000 housing units or 46.0 percent fall below this interval, and 34,700 housing units or 19.7 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(48.2-46.0)}{19.7} = 2.6$$

Similarly, the interval for owner-occupied housing units with three persons contains the 51.8 percent derived in step 2. About 81,000 housing units or 46.0 percent fall below this interval, and 34,700 housing units or 19.7 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(51.8-46.0)}{19.7} = 2.8$$

Thus, the 95-percent confidence interval ranges from 2.6 to 2.8 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1980 Housing Inventory and for Estimated Number of 1977-1980 Lost Units for the Albany-Schenectady-Troy, 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]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	60	50	60	50,000	1,580	1,160	1,540
100	80	70	80	75,000	1,840	990	1,710
200	110	100	110	99,200	1,990	–	1,760
500	170	160	180	100,000	2,000	–	1,760
700	210	190	210	150,000	2,010	–	1,500
1,000	250	230	250	196,200	1,990	–	–
2,500	390	360	400	200,000	1,970	–	–
5,000	540	510	560	250,000	1,520	–	–
10,000	760	700	780	295,400	–	–	–
25,000	1,170	1,010	1,180				

[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, 1.1 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 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 1980 Housing Inventory and for Estimated Percentages of 1977-1980 Lost Housing Units for the Albany-Schenectady-Troy, 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]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	38.8	38.8	38.8	38.8	38.8	39.8	25,000	0.3	0.5	1.1	1.5	2.2	2.5
200	24.0	24.0	24.0	24.0	24.4	28.1	50,000	0.13	0.4	0.8	1.1	1.5	1.8
500	11.2	11.2	11.2	11.2	15.4	17.8	75,000	0.08	0.3	0.6	0.9	1.3	1.5
700	8.3	8.3	8.3	9.0	13.0	15.0	100,000	0.06	0.3	0.5	0.8	1.1	1.3
1,000	6.0	6.0	6.0	7.5	10.9	12.6	150,000	0.04	0.2	0.4	0.6	0.9	1.0
2,500	2.5	2.5	3.5	4.8	6.9	8.0	200,000	0.03	0.2	0.4	0.5	0.8	0.9
5,000	1.3	1.3	2.5	3.4	4.9	5.6	250,000	0.03	0.2	0.3	0.5	0.7	0.8
10,000	0.6	0.8	1.7	2.4	3.4	4.0	300,000	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, 1.1 for the central city, and 1.2 for the balance (not in central city).

Allentown-Bethlehem-Easton 1980

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1980 survey. App-42
Selection of the 1976 AHS-SMSA sample App-43
1976-1980 additions to the housing inventory App-44
Sample selection for the 1976 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
1980 housing inventory App-45
1976-1980 lost housing units . . . App-46
1976 estimation procedure App-46
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-47
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-170-80) are based on data collected from the 1980 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 the Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill., SMSA's, the data were collected during the 11-month period from April 1980 through February 1981 with one-eleventh of the sample housing units being visited each month. In the remaining SMSA's, the data were collected during the 12-month period from April 1980 through March 1981 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 1980 were interviewed previously in either 1976 or 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1980 group of SMSA's, the largest SMSA from 3 of the 4 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 1980 group are: Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill.

The remaining SMSA's in the 1980 group are: Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Birmingham, Ala.; Grand Rapids, Mich.; Indianapolis, Ind.; Louisville, Ky.-Ind.; Memphis, Tenn.-Ark.; Oklahoma City, Okla.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; and Salt Lake City, Utah.

In this SMSA, 4,886 housing units were eligible for interview. Of these sample housing units, 102 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 4,886 housing units eligible for interview, 322 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 1980 survey—The sample housing units designated to be interviewed in the 1980 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 1980 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 and 1980 Coverage Improvement Programs. (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—The sample for the SMSA's which, in 1970, 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 1970, the following three SMSA's were 100-percent permit-issuing: Los Angeles-Long Beach, Calif.; New York, N.Y.; and Sacramento, Calif. The remaining 12 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.

1976-1980 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 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with the Oklahoma City, Okla. SMSA receiving some updating as a part of the 1980 AHS. For each of the five SMSA's interviewed previously in 1977, the Coverage Improvement Program was conducted as a part of the 1977 AHS with the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; and Salt Lake City, Utah, SMSA's receiving some updating and refining as a part of the 1980 AHS. The following discussion applies to both the prior year (1976 or 1977) and 1980 coverage improvement procedures. For the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Oklahoma City, Okla.; and Salt Lake City, Utah, SMSA's estimates of housing units added by a specific procedure reflect units added in the prior year as well as any additions that resulted from the updating and refining in 1980.

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 1980 SMSA's. Two different procedures were used. For the first procedure, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. For the SMSA's previously interviewed in 1976, these new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units). For the SMSA's previously interviewed in 1977, sample units selected from one- and two-unit structures were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample, while units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage sample selection was a sample of permit offices, and the second stage was a sample of the 1969 permits within each of the selected permit offices. In the Grand Rapids, Mich., and New York, N.Y., SMSA's, 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. This procedure was employed for the 10 SMSA's previously interviewed in 1976.

For the Albany-Schenectady-Troy, N.Y.; Memphis, Tenn.-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's previously interviewed in 1977, the above procedures were performed for the first and second stages. For the third stage, structures of size three or more were divided into clusters of an expected size of two units and a sample of clusters was selected.

For the Los Angeles-Long Beach, Calif., SMSA, 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 units were then sampled at one-third the rate of regular AHS units. Since permits were not available for all sampled offices, the procedure was also used in parts of the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 657 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. In the Oklahoma City, Okla., SMSA, a sample of tracts was selected but not canvassed during the 1976 Coverage Improvement Program. This procedure was completed during the 1980 AHS and added 406 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 AHS sample housing units was selected from the 1970 sample universe at a rate of 1 in 24 for the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Memphis, Tenn-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's and at a rate of 1 in 22 for the remaining 10 SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing unit 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. The Albany-Schenectady-Troy, N.Y., SMSA, had these sampling procedures implemented in 1976, excluding identification of the housing units missed in the 1970 census. This procedure was completed as a part of the 1980 AHS. The Los Angeles-Long Beach, Calif., and Salt Lake City, Utah, SMSA's had this sampling procedure implemented for the first time in 1980. This procedure added an estimated 5,063 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 626 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 1980 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 1980 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1976 (i.e., 1976-1980 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1980 housing inventory—The AHS estimates of characteristics of the 1980 housing inventory were produced using a 2-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Grand Rapids, Mich.; Indianapolis, Ind.; Los Angeles-Long Beach, Calif.; Louisville, Ky.-Ind.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; St. Louis, Mo.-Ill.; and Salt Lake City, Utah, 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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., SMSA's. In the Birmingham, Ala.; Memphis, Tenn.-Ark.; and Oklahoma City, Okla., 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 New York, N.Y., SMSA, the sample selected correctly apportioned the new construction housing units between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 1980 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:

$$\frac{\text{Independent estimate of the October 1980 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 procedure, 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.

1976-1980 lost housing units—The 1976-1980 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 1976-1980 lost housing units existed, by definition, in the 1976 housing inventory, there was a 1976 housing inventory weight associated with each 1976-1980 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1976-1980 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 1980 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 of which 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 1980 survey, the number of missed housing units may be considerably less for 1980.

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

Rounding errors—For errors associated with processing, 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 1976 housing inventory can be found in the AHS Series H-170 reports for 1976.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1980 housing inventory as well as estimates of characteristics of the 1976-1980 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 8,420 for the total SMSA, 4,380 for the central city of the SMSA, and 7,100 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.

Table II (page App-51) presents the standard errors of estimated percentages for the 1980 housing inventory as well as estimated percentages of the 1976-1980 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 1980 there were 152,400 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 1,570. 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
150,000	1,600
152,400	x
200,000	1,050

The entry for "x" is determined as follows by vertically interpolating between 1,600 and 1,050.

$$152{,}400-150{,}000 = 2{,}400$$
$$200{,}000-150{,}000 = 50{,}000$$

$$1{,}600 + \frac{2{,}400}{50{,}000}(1{,}050-1{,}600) = 1{,}570$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 150,830 to 153,970 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1980 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 149,890 to 154,910 housing units with 90 percent confidence; and that the average estimate lies within the interval from 149,260 to 155,540 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 152,400 owner-occupied housing units, 31,600, or 20.7 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 20.7 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	20.7	25 or 75
150,000.	0.6	a	0.8
152,400.		p	
200,000.	0.5	b	0.7

1. The entry for cell "a" is determined by horizontal interpolation between 0.6 and 0.8

$$20.7-10.0 = 10.7$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{10.7}{15.0}(0.8-0.6) = 0.7$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.5 and 0.7.

$$20.7-10.0 = 10.7$$
$$25.0-10.0 = 15.0$$

$$0.5 + \frac{10.7}{15.0}(0.7-0.5) = 0.6$$

3. The entry for "p" is then determined by vertical interpolation between 0.7 and 0.6.

$$152{,}400-150{,}000 = 2{,}400$$
$$200{,}000-150{,}000 = 50{,}000$$

$$0.7 + \frac{2{,}400}{50{,}000}(0.6-0.7) = 0.7$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 20.0 to 21.4 percent; the 90-percent confidence interval is from 19.6 to 21.8 percent; and the 95-percent confidence interval is from 19.3 to 22.1 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; if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1976 and 1980 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1980 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1980 there were 83,100 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 51,500. Table I shows the standard error of 31,600 is approximately 1,160, and the standard error of 83,100 is approximately 1,640. Therefore, the standard error of the estimated difference of 51,500 is about 2,010:

$$2{,}010 = \sqrt{(1{,}160)^2 + (1{,}640)^2}$$

Consequently, the 68-percent confidence interval for the 51,500 difference is from 49,490 to 53,510 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 48,280 to 54,720 housing units, and the 95-percent confidence interval is from 47,480 to 55,520 housing units. Thus, we can conclude with 95 percent confidence that the number of 1980 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. 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. 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 152,400 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 152,400 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 three persons (for purposes of calculating the median, the category of three persons is considered to be from 2.5 to 3.5 persons) contains the 48.2 percent derived in step 2. About 69,800 housing units or 45.8 percent fall below this interval, and 30,700 housing units or 20.1 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(48.2-45.8)}{20.1} = 2.6$$

Similary, the interval for owner-occupied housing units with three persons contains the 51.8 percent derived in step 2. About 69,800 housing units or 45.8 percent fall below this interval, and 30,700 housing units or 20.1 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(51.8-45.8)}{20.1} = 2.8$$

Thus, the 95-percent confidence interval ranges from 2.6 to 2.8 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1980 Housing Inventory and for Estimated Number of 1976-1980 Lost Units for the Allentown-Bethlehem-Easton, Pa.-N.J., 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]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	50	50	50	25,000	1,070	900	1,030
100	70	70	70	50,000	1,420	960	1,290
200	100	100	100	75,000	1,600	540	1,350
500	160	150	160	81,600	1,640	–	1,330
700	190	180	190	100,000	1,690	–	1,230
1,000	230	220	230	142,200	1,640	–	–
2,500	360	340	350	150,000	1,600	–	–
5,000	500	470	500	200,000	1,050	–	–
10,000	700	640	690	223,800	–	–	–

[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, 1.4 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 1980 Housing Inventory and for Estimated Percentages of 1976-1980 Lost Housing Units for the Allentown-Bethlehem-Easton, Pa.-N.J., SMSA, for the Central City and for the Balance (Not in the Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	34.0	34.0	34.0	34.0	34.0	35.9	25,000	0.2	0.5	1.0	1.4	2.0	2.3
200	20.5	20.5	20.5	20.5	22.0	25.4	50,000	0.10	0.3	0.7	1.0	1.4	1.6
500	9.4	9.4	9.4	9.6	13.9	16.1	75,000	0.07	0.3	0.6	0.8	1.1	1.3
700	6.9	6.9	6.9	8.1	11.8	13.6	100,000	0.05	0.2	0.5	0.7	1.0	1.1
1,000	4.9	4.9	5.0	6.8	9.8	11.4	150,000	0.03	0.2	0.4	0.6	0.8	0.9
2,500	2.0	2.0	3.1	4.3	6.2	7.2	200,000	0.03	0.2	0.4	0.5	0.7	0.8
5,000	1.0	1.0	2.2	3.0	4.4	5.1	250,000	0.02	0.14	0.3	0.4	0.6	0.7
10,000	0.5	0.7	1.6	2.2	3.1	3.6							

[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 1.2 for the total SMSA, 1.4 for the central city, and 1.1 for the balance (not in central city).

Birmingham 1980

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1980 survey. App-42
Selection of the 1976 AHS-SMSA sample App-43
1976-1980 additions to the housing inventory App-44
Sample selection for the 1976 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
1980 housing inventory App-45
1976-1980 lost housing units . . . App-46
1976 estimation procedure App-46
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-47
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-170-80) are based on data collected from the 1980 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 the Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill., SMSA's, the data were collected during the 11-month period from April 1980 through February 1981 with one-eleventh of the sample housing units being visited each month. In the remaining SMSA's, the data were collected during the 12-month period from April 1980 through March 1981 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 1980 were interviewed previously in either 1976 or 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1980 group of SMSA's, the largest SMSA from 3 of the 4 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 1980 group are: Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill.

The remaining SMSA's in the 1980 group are: Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Birmingham, Ala.; Grand Rapids, Mich.; Indianapolis, Ind.; Louisville, Ky.-Ind.; Memphis, Tenn.-Ark.; Oklahoma City, Okla.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; and Salt Lake City, Utah.

In this SMSA, 4,979 housing units were eligible for interview. Of these sample housing units, 122 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 4,979 housing units eligible for interview, 429 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 1980 survey—The sample housing units designated to be interviewed in the 1980 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 1980 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 and 1980 Coverage Improvement Programs. (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—The sample for the SMSA's which, in 1970, 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 1970, the following three SMSA's were 100-percent permit-issuing: Los Angeles-Long Beach, Calif.; New York, N.Y.; and Sacramento, Calif. The remaining 12 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.

1976-1980 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 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with the Oklahoma City, Okla. SMSA receiving some updating as a part of the 1980 AHS. For each of the five SMSA's interviewed previously in 1977, the Coverage Improvement Program was conducted as a part of the 1977 AHS with the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; and Salt Lake City, Utah, SMSA's receiving some updating and refining as a part of the 1980 AHS. The following discussion applies to both the prior year (1976 or 1977) and 1980 coverage improvement procedures. For the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Oklahoma City, Okla.; and Salt Lake City, Utah, SMSA's estimates of housing units added by a specific procedure reflect units added in the prior year as well as any additions that resulted from the updating and refining in 1980.

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 1980 SMSA's. Two different procedures were used. For the first procedure, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. For the SMSA's previously interviewed in 1976, these new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units). For the SMSA's previously interviewed in 1977, sample units selected from one- and two-unit structures were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample, while units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage sample selection was a sample of permit offices, and the second stage was a sample of the 1969 permits within each of the selected permit offices. In the Grand Rapids, Mich., and New York, N.Y., SMSA's, 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. This procedure was employed for the 10 SMSA's previously interviewed in 1976.

For the Albany-Schenectady-Troy, N.Y.; Memphis, Tenn.-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's previously interviewed in 1977, the above procedures were performed for the first and second stages. For the third stage, structures of size three or more were divided into clusters of an expected size of two units and a sample of clusters was selected.

For the Los Angeles-Long Beach, Calif., SMSA, 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 units were then sampled at one-third the rate of regular AHS units. Since permits were not available for all sampled offices, the procedure was also used in parts of the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added no 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. In the Oklahoma City, Okla., SMSA, a sample of tracts was selected but not canvassed during the 1976 Coverage Improvement Program. This procedure was completed during the 1980 AHS and added 737 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 AHS sample housing units was selected from the 1970 sample universe at a rate of 1 in 24 for the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Memphis, Tenn-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's and at a rate of 1 in 22 for the remaining 10 SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing unit 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. The Albany-Schenectady-Troy, N.Y., SMSA, had these sampling procedures implemented in 1976, excluding identification of the housing units missed in the 1970 census. This procedure was completed as a part of the 1980 AHS. The Los Angeles-Long Beach, Calif., and Salt Lake City, Utah, SMSA's had this sampling procedure implemented for the first time in 1980. This procedure added an estimated 8,344 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 655 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 1980 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 1980 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1976 (i.e., 1976-1980 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1980 housing inventory—The AHS estimates of characteristics of the 1980 housing inventory were produced using a 2-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Grand Rapids, Mich.; Indianapolis, Ind.; Los Angeles-Long Beach, Calif.; Louisville, Ky.-Ind.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; St. Louis, Mo.-Ill.; and Salt Lake City, Utah, 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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., SMSA's. In the Birmingham, Ala.; Memphis, Tenn.-Ark.; and Oklahoma City, Okla., 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 New York, N.Y., SMSA, the sample selected correctly apportioned the new construction housing units between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 1980 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:

$$\frac{\text{Independent estimate of the October 1980 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 procedure, 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.

1976-1980 lost housing units—The 1976-1980 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 1976-1980 lost housing units existed, by definition, in the 1976 housing inventory, there was a 1976 housing inventory weight associated with each 1976-1980 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1976-1980 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 1980 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 of which 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 1980 survey, the number of missed housing units may be considerably less for 1980.

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

Rounding errors—For errors associated with processing, 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 1976 housing inventory can be found in the AHS Series H-170 reports for 1976.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1980 housing inventory as well as estimates of characteristics of the 1976-1980 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 15,960 for the total SMSA, 7,060 for the central city of the SMSA, and 14,190 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-51) presents the standard errors of estimated percentages for the 1980 housing inventory as well as estimated percentages of the 1976-1980 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 1980 there were 201,800 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 2,260. 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
200,000	2,270
201,800	x
206,300	2,240

The entry for "x" is determined as follows by vertically interpolating between 2,270 and 2,240.

$$201{,}800 - 200{,}000 = 1{,}800$$
$$206{,}300 - 200{,}000 = 6{,}300$$

$$2{,}270 + \frac{1{,}800}{6{,}300}(2{,}240 - 2{,}270) = 2{,}260$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 199,540 to 204,060 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1980 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 198,180 to 205,420 housing units with 90 percent confidence; and that the average estimate lies within the interval from 197,280 to 206,320 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 201,800 owner-occupied housing units, 58,100, or 28.8 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 28.8 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		
	25 or 75	28.8	50
200,000	0.8	a	1.0
201,800		p	
250,000	0.7	b	0.9

1. The entry for cell "a" is determined by horizontal interpolation between 0.8 and 1.0.

$$28.8-25.0 = 3.8$$
$$50.0-25.0 = 25.0$$

$$0.8 + \frac{3.8}{25.0}(1.0-0.8) = 0.8$$

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

$$28.8-25.0 = 3.8$$
$$50.0-25.0 = 25.0$$

$$0.7 + \frac{3.8}{25.0}(0.9-0.7) = 0.7$$

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

$$201{,}800-200{,}000 = 1{,}800$$
$$250{,}000-200{,}000 = 50{,}000$$

$$0.8 + \frac{1{,}800}{50{,}000}(0.7-0.8) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 28.0 to 29.6 percent; the 90-percent confidence interval is from 27.5 to 30.1 percent; and the 95-percent confidence interval is from 27.2 to 30.4 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; if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1976 and 1980 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1980 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1980 there were 112,600 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 54,500. Table I shows the standard error of 58,100 is approximately 1,810, and the standard error of 112,600 is approximately 2,250. Therefore, the standard error of the estimated difference of 54,500 is about 2,890:

$$2{,}890 = \sqrt{(1{,}810)^2 + (2{,}250)^2}$$

Consequently, the 68-percent confidence interval for the 54,500 difference is from 51,610 to 57,390 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 49,880 to 59,120 housing units, and the 95-percent confidence interval is from 48,720 to 60,280 housing units. Thus, we can conclude with 95 percent confidence that the number of 1980 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 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 201,800 housing units.

1. Interpolating using table II shows that the standard error of 50 percent on a base of 201,800 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 three persons (for purposes of calculating the median, the category of three persons is considered to be from 2.5 to 3.5 persons) contains the 48.0 percent derived in step 2. About 95,200 housing units or 47.2 percent fall below this interval, and 41,500 housing units or 20.6 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\frac{(48.0-47.2)}{20.6} = 2.5$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.0 percent derived in step 2. About 95,200 housing units or 47.2 percent fall below this interval, and 41,500 housing units or 20.6 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\frac{(52.0-47.2)}{20.6} = 2.7$$

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 1980 Housing Inventory and for Estimated Number of 1976-1980 Lost Units for the Birmingham, Ala., 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]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	70	60	70	50,000	1,720	1,300	1,680
100	80	80	90	75,000	2,010	1,210	1,890
200	120	110	120	100,000	2,200	700	1,960
500	190	180	190	108,600	2,240	–	1,960
700	220	210	230	150,000	2,360	–	1,750
1,000	270	250	270	200,000	2,270	–	680
2,500	420	390	430	206,300	2,240	–	–
5,000	590	550	600	250,000	1,910	–	–
10,000	830	750	840	300,000	1,000	–	–
25,000	1,270	1,100	1,280	314,900	–	–	–

[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 1980 Housing Inventory and for Estimated Percentages of 1976-1980 Lost Housing Units for the Birmingham, Ala., 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]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	42.8	42.8	42.8	42.8	42.8	43.2	50,000	0.15	0.4	0.8	1.2	1.7	1.9
200	27.2	27.2	27.2	27.2	27.2	30.6	75,000	0.10	0.3	0.7	0.9	1.4	1.6
500	13.0	13.0	13.0	13.0	16.7	19.3	100,000	0.07	0.3	0.6	0.8	1.2	1.4
700	9.7	9.7	9.7	9.8	14.2	16.3	150,000	0.05	0.2	0.5	0.7	1.0	1.1
1,000	7.0	7.0	7.0	8.2	11.8	13.7	200,000	0.04	0.2	0.4	0.6	0.8	1.0
2,500	2.9	2.9	3.8	5.2	7.5	8.6	250,000	0.03	0.2	0.4	0.5	0.7	0.9
5,000	1.5	1.5	2.7	3.7	5.3	6.1	300,000	0.02	0.2	0.3	0.5	0.7	0.8
10,000	0.7	0.9	1.9	2.6	3.7	4.3	400,000	0.02	0.14	0.3	0.4	0.6	0.7
25,000	0.3	0.5	1.2	1.6	2.4	2.7							

[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.

Grand Rapids 1980

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey . . . App-42
Designation of sample housing units for the 1980 survey . . App-42
Selection of the 1976 AHS-SMSA sample . . . App-43
1976-1980 additions to the housing inventory . . . App-44
Sample selection for the 1976 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
1980 housing inventory . . App-45
1976-1980 lost housing units . . App-46
1976 estimation procedure . . App-46
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-47
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-170-80) are based on data collected from the 1980 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 the Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill., SMSA's, the data were collected during the 11-month period from April 1980 through February 1981 with one-eleventh of the sample housing units being visited each month. In the remaining SMSA's, the data were collected during the 12-month period from April 1980 through March 1981 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 1980 were interviewed previously in either 1976 or 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1980 group of SMSA's, the largest SMSA from 3 of the 4 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 1980 group are: Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill.

The remaining SMSA's in the 1980 group are: Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Birmingham, Ala.; Grand Rapids, Mich.; Indianapolis, Ind.; Louisville, Ky.-Ind.; Memphis, Tenn.-Ark.; Oklahoma City, Okla.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; and Salt Lake City, Utah.

In this SMSA, 5,153 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 5,153 housing units eligible for interview, 288 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 1980 survey—The sample housing units designated to be interviewed in the 1980 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 1980 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 and 1980 Coverage Improvement Programs. (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—The sample for the SMSA's which, in 1970, 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 1970, the following three SMSA's were 100-percent permit-issuing: Los Angeles-Long Beach, Calif.; New York, N.Y.; and Sacramento, Calif. The remaining 12 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.

1976-1980 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 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with the Oklahoma City, Okla. SMSA receiving some updating as a part of the 1980 AHS. For each of the five SMSA's interviewed previously in 1977, the Coverage Improvement Program was conducted as a part of the 1977 AHS with the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; and Salt Lake City, Utah, SMSA's receiving some updating and refining as a part of the 1980 AHS. The following discussion applies to both the prior year (1976 or 1977) and 1980 coverage improvement procedures. For the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Oklahoma City, Okla.; and Salt Lake City, Utah, SMSA's estimates of housing units added by a specific procedure reflect units added in the prior year as well as any additions that resulted from the updating and refining in 1980.

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 1980 SMSA's. Two different procedures were used. For the first procedure, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. For the SMSA's previously interviewed in 1976, these new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units). For the SMSA's previously interviewed in 1977, sample units selected from one- and two-unit structures were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample, while units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage sample selection was a sample of permit offices, and the second stage was a sample of the 1969 permits within each of the selected permit offices. In the Grand Rapids, Mich., and New York, N.Y., SMSA's, 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. This procedure was employed for the 10 SMSA's previously interviewed in 1976.

For the Albany-Schenectady-Troy, N.Y.; Memphis, Tenn.-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's previously interviewed in 1977, the above procedures were performed for the first and second stages. For the third stage, structures of size three or more were divided into clusters of an expected size of two units and a sample of clusters was selected.

For the Los Angeles-Long Beach, Calif., SMSA, 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 units were then sampled at one-third the rate of regular AHS units. Since permits were not available for all sampled offices, the procedure was also used in parts of the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 169 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. In the Oklahoma City, Okla., SMSA, a sample of tracts was selected but not canvassed during the 1976 Coverage Improvement Program. This procedure was completed during the 1980 AHS and added 1,081 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 AHS sample housing units was selected from the 1970 sample universe at a rate of 1 in 24 for the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Memphis, Tenn-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's and at a rate of 1 in 22 for the remaining 10 SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing unit 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. The Albany-Schenectady-Troy, N.Y., SMSA, had these sampling procedures implemented in 1976, excluding identification of the housing units missed in the 1970 census. This procedure was completed as a part of the 1980 AHS. The Los Angeles-Long Beach, Calif., and Salt Lake City, Utah, SMSA's had this sampling procedure implemented for the first time in 1980. This procedure added an estimated 1,947 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 187 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 1980 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 1980 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1976 (i.e., 1976-1980 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1980 housing inventory—The AHS estimates of characteristics of the 1980 housing inventory were produced using a 2-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Grand Rapids, Mich.; Indianapolis, Ind.; Los Angeles-Long Beach, Calif.; Louisville, Ky.-Ind.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; St. Louis, Mo.-Ill.; and Salt Lake City, Utah, 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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., SMSA's. In the Birmingham, Ala.; Memphis, Tenn.-Ark.; and Oklahoma City, Okla., 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 New York, N.Y., SMSA, the sample selected correctly apportioned the new construction housing units between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 1980 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:

$$\frac{\text{Independent estimate of the October 1980 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 procedure, 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.

1976-1980 lost housing units—The 1976-1980 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 1976-1980 lost housing units existed, by definition, in the 1976 housing inventory, there was a 1976 housing inventory weight associated with each 1976-1980 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1976-1980 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 1980 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 of which 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 1980 survey, the number of missed housing units may be considerably less for 1980.

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

Rounding errors—For errors associated with processing, 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 1976 housing inventory can be found in the AHS Series H-170 reports for 1976.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1980 housing inventory as well as estimates of characteristics of the 1976-1980 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 6,850 for the total SMSA, 3,490 for the central city of the SMSA, and 5,820 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 1980 housing inventory as well as estimated percentages of the 1976-1980 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 1980 there were 155,300 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 1,450. 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
151,700	1,490
155,300	x
200,000	950

The entry for "x" is determined as follows by vertically interpolation between 1,490 and 950.

$$155{,}300 - 151{,}700 = 3{,}600$$
$$200{,}000 - 151{,}700 = 48{,}300$$

$$1{,}490 + \frac{3{,}600}{48{,}300}(950 - 1{,}490) = 1{,}450$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 153,850 to 156,750 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1980 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 152,980 to 157,620 housing units with 90 percent confidence; and that the average estimate lies within the interval from 152,400 to 158,200 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 155,300 owner-occupied housing units, 35,600, or 22.9 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 22.9 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	22.9	25 or 75
150,000.	0.5	a	0.8
155,300.		p	
200,000.	0.5	b	0.7

1. The entry for cell "a" is determined by horizontal interpolation between 0.5 and 0.8.

$$22.9-10.0 = 12.9$$
$$25.0-10.0 = 15.0$$

$$0.5 + \frac{12.9}{15.0}(0.8-0.5) = 0.8$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.5 and 0.7.

$$22.9-10.0 = 12.9$$
$$25.0-10.0 = 15.0$$

$$0.5 + \frac{12.9}{15.0}(0.7-0.5) = 0.7$$

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

$$155{,}300-150{,}000 = 5{,}300$$
$$200{,}000-150{,}000 = 50{,}000$$

$$0.8 + \frac{5{,}300}{50{,}000}(0.7-0.8) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 22.1 to 23.7 percent; the 90-percent confidence interval is from 21.6 to 24.2 percent; and the 95-percent confidence interval is from 21.3 to 24.5 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; if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1976 and 1980 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1980 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1980 there were 81,800 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 units with three bedrooms is 46,200. Table I shows the standard error of 35,600 is approximately 1,160, and the standard error of 81,800 is approximately 1,540. Therefore, the standard error of the estimated difference of 46,200 is about 1,930:

$$1{,}930 = \sqrt{(1{,}160)^2 + (1{,}540)^2}$$

Consequently, the 68-percent confidence interval for the 46,200 difference is from 44,270 to 48,130 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 43,110 to 49,290 housing units, and the 95-percent confidence interval is from 42,340 to 50,060 housing units. thus, we can conclude with 95 percent confidence that the number of 1980 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 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 value 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 155,300 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 155,300 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 three persons (for purposes of calculating the median, the category of three persons is considered to be from 2.5 to 3.5 persons) contains the 48.2 percent derived in step 2. About 64,900 housing units or 41.8 percent fall below this interval, and 28,400 housing units or 18.3 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(48.2-41.8)}{18.3} = 2.8$$

Similarly, the interval for owner-occupied housing units with three persons contains the 51.8 percent derived in step 2. About 64,900 housing units or 41.8 percent fall below this interval, and 28,400 housing units or 18.3 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(51.8-41.8)}{18.3} = 3.0$$

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 1980 Housing Inventory and for Estimated Number of 1976-1980 Lost Units for the Grand Rapids, Mich., 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]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	50	40	50	25,000	1,020	830	1,000
100	70	70	70	50,000	1,340	780	1,270
200	100	90	100	70,000	1,490	–	1,340
500	150	150	150	75,000	1,520	–	1,350
700	180	170	180	100,000	1,600	–	1,270
1,000	220	210	220	150,000	1,500	–	220
2,500	340	320	340	151,700	1,490	–	–
5,000	480	450	480	200,000	950	–	–
10,000	670	600	670	221,700	–	–	–

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.3. 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 Percentages of Housing Units in the 1980 Housing Inventory and for Estimated Percentages of 1976-1980 Lost Housing Units for the Grand Rapids, Mich., 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]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	32.5	32.5	32.5	32.5	32.5	34.7	25,000	0.2	0.4	1.0	1.3	1.9	2.2
200	19.4	19.4	19.4	19.4	21.2	24.5	50,000	0.10	0.3	0.7	0.9	1.3	1.6
500	8.8	8.8	8.8	9.3	13.4	15.5	75,000	0.06	0.3	0.6	0.8	1.1	1.3
700	6.4	6.4	6.4	7.9	11.3	13.1	100,000	0.05	0.2	0.5	0.7	0.9	1.1
1,000	4.6	4.6	4.8	6.6	9.5	11.0	150,000	0.03	0.2	0.4	0.5	0.8	0.9
2,500	1.9	1.9	3.0	4.2	6.0	6.9	200,000	0.02	0.15	0.3	0.5	0.7	0.8
5,000	1.0	1.0	2.1	2.9	4.2	4.9	250,000	0.02	0.14	0.3	0.4	0.6	0.7
10,000	0.5	0.7	1.5	2.1	3.0	3.5							

[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 1.3.

Indianapolis 1980

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1980 survey. App-42
Selection of the 1976 AHS-SMSA sample App-43
1976-1980 additions to the housing inventory App-44
Sample selection for the 1976 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
1980 housing inventory App-45
1976-1980 lost housing units . . . App-46
1976 estimation procedure App-46
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-47
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-170-80) are based on data collected from the 1980 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 the Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill., SMSA's, the data were collected during the 11-month period from April 1980 through February 1981 with one-eleventh of the sample housing units being visited each month. In the remaining SMSA's, the data were collected during the 12-month period from April 1980 through March 1981 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 1980 were interviewed previously in either 1976 or 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1980 group of SMSA's, the largest SMSA from 3 of the 4 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 1980 group are: Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill.

The remaining SMSA's in the 1980 group are: Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Birmingham, Ala.; Grand Rapids, Mich.; Indianapolis, Ind.; Louisville, Ky.-Ind.; Memphis, Tenn.-Ark.; Oklahoma City, Okla.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; and Salt Lake City, Utah.

In this SMSA, 5,045 housing units were eligible for interview. Of these sample housing units, 119 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 5,045 housing units eligible for interview, 305 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 1980 survey—The sample housing units designated to be interviewed in the 1980 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 1980 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 and 1980 Coverage Improvement Programs. (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—The sample for the SMSA's which, in 1970, 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 1970, the following three SMSA's were 100-percent permit-issuing: Los Angeles-Long Beach, Calif.; New York, N.Y.; and Sacramento, Calif. The remaining 12 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.

1976-1980 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 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with the Oklahoma City, Okla. SMSA receiving some updating as a part of the 1980 AHS. For each of the five SMSA's interviewed previously in 1977, the Coverage Improvement Program was conducted as a part of the 1977 AHS with the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; and Salt Lake City, Utah, SMSA's receiving some updating and refining as a part of the 1980 AHS. The following discussion applies to both the prior year (1976 or 1977) and 1980 coverage improvement procedures. For the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Oklahoma City, Okla.; and Salt Lake City, Utah, SMSA's estimates of housing units added by a specific procedure reflect units added in the prior year as well as any additions that resulted from the updating and refining in 1980.

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 1980 SMSA's. Two different procedures were used. For the first procedure, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. For the SMSA's previously interviewed in 1976, these new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units). For the SMSA's previously interviewed in 1977, sample units selected from one- and two-unit structures were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample, while units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage sample selection was a sample of permit offices, and the second stage was a sample of the 1969 permits within each of the selected permit offices. In the Grand Rapids, Mich., and New York, N.Y., SMSA's, 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. This procedure was employed for the 10 SMSA's previously interviewed in 1976.

For the Albany-Schenectady-Troy, N.Y.; Memphis, Tenn.-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's previously interviewed in 1977, the above procedures were performed for the first and second stages. For the third stage, structures of size three or more were divided into clusters of an expected size of two units and a sample of clusters was selected.

For the Los Angeles-Long Beach, Calif., SMSA, 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 units were then sampled at one-third the rate of regular AHS units. Since permits were not available for all sampled offices, the procedure was also used in parts of the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 1,812 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. In the Oklahoma City, Okla., SMSA, a sample of tracts was selected but not canvassed during the 1976 Coverage Improvement Program. This procedure was completed during the 1980 AHS and added 792 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 AHS sample housing units was selected from the 1970 sample universe at a rate of 1 in 24 for the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Memphis, Tenn-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's and at a rate of 1 in 22 for the remaining 10 SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing unit 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. The Albany-Schenectady-Troy, N.Y., SMSA, had these sampling procedures implemented in 1976, excluding identification of the housing units missed in the 1970 census. This procedure was completed as a part of the 1980 AHS. The Los Angeles-Long Beach, Calif., and Salt Lake City, Utah, SMSA's had this sampling procedure implemented for the first time in 1980. This procedure added an estimated 5,255 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 1,789 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 1980 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 1980 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1976 (i.e., 1976-1980 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1980 housing inventory—The AHS estimates of characteristics of the 1980 housing inventory were produced using a 2-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Grand Rapids, Mich.; Indianapolis, Ind.; Los Angeles-Long Beach, Calif.; Louisville, Ky.-Ind.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; St. Louis, Mo.-Ill.; and Salt Lake City, Utah, 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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., SMSA's. In the Birmingham, Ala.; Memphis, Tenn.-Ark.; and Oklahoma City, Okla., 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 New York, N.Y., SMSA, the sample selected correctly apportioned the new construction housing units between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 1980 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:

$$\frac{\text{Independent estimate of the October 1980 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 procedure, 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.

1976-1980 lost housing units—The 1976-1980 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 1976-1980 lost housing units existed, by definition, in the 1976 housing inventory, there was a 1976 housing inventory weight associated with each 1976-1980 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1976-1980 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 1980 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 of which 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 1980 survey, the number of missed housing units may be considerably less for 1980.

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

Rounding errors—For errors associated with processing, 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 1976 housing inventory can be found in the AHS Series H-170 reports for 1976.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1980 housing inventory as well as estimates of characteristics of the 1976-1980 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 14,670 for the total SMSA, 9,920 for the central city of the SMSA, and 10,740 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 1980 housing inventory as well as estimated percentages of the 1976-1980 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 1980 there were 276,400 owner-occupied housing units in this SMSA. Interpolation using tabel I of this appendix shows that the standard error of an estimate of this size is approximately 3,300. 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	3,370
276,400	x
286,200	3,280

The entry for "x" is determined as follows by vertically interpolating between 3,370 and 3,280.

$$276{,}400-250{,}000 = 26{,}400$$
$$286{,}200-250{,}000 = 36{,}200$$

$$3{,}370+\frac{26{,}400}{36{,}200}(3{,}280-3{,}370) = 3{,}300$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 273,100 to 279,700 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1980 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 271,120 to 281,680 housing units with 90 percent confidence; and that the average estimate lies within the interval from 269,800 to 283,000 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 276,400 owner-occupied housing units, 83,600, or 30.2 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 30.2 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	30.2	50
250,000.	0.9	a	1.0
276,400.		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.

$$30.2-25.0 = 5.2$$
$$50.0-25.0 = 25.0$$

$$0.9 + \frac{5.2}{25.0}(1.0-0.9) = 0.9$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.8 and 0.9.

$$30.2-25.0 = 5.2$$
$$50.0-25.0 = 25.0$$

$$0.8 + \frac{5.2}{25.0}(0.9-0.8) = 0.8$$

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

$$276{,}400-250{,}000 = 26{,}400$$
$$300{,}000-250{,}000 = 50{,}000$$

$$0.9 + \frac{26{,}400}{50{,}000}(0.8-0.9) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 29.4 to 31.0 percent; the 90-percent confidence interval is from 28.9 to 31.5 percent; and the 95-percent confidence interval is from 28.6 to 31.8 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 differnce 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; if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1976 and 1980 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1980 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1980 there were 140,300 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 56,700. Table I shows the standard error of 83,600 is approximately 2,610, and the standard error of 140,300 is approximately 3,120. Therefore, the standard error of the estimated difference of 56,700 is about 4,070:

$$4{,}070 = \sqrt{(2{,}610)^2 + (3{,}120)^2}$$

Consequently, the 68-percent confidence interval for the 56,700 difference is from 52,630 to 60,770 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 50,190 to 63,210 housing units, and the 95-percent confidence interval is from 48,560 to 64,840 housing units. Thus, we can conclude with 95 percent confidence that the number of 1980 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 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 for 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 276,400 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 276,400 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 three persons (for purposes of calculating the median, the category of three persons is considered to be from 2.5 to 3.5 persons) contains the 48.2 percent derived in step 2. About 126,900 housing units or 45.9 percent fall below this interval, and 51,100 housing units or 18.5 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(48.2-45.9)}{18.5} = 2.6$$

Similarly, the interval for owner-occupied housing units with three persons contains the 51.8 percent derived in step 2. About 126,900 housing units or 45.9 percent fall below this interval, and 51,100 housing units or 18.5 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(51.8-45.9)}{18.5} = 2.8$$

Thus, the 95-percent confidence interval ranges from 2.6 to 2.8 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1980 Housing Inventory and for Estimated Number of 1976-1980 Lost Units for the Indianapolis, Ind., 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]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	100	100	100	75,000	2,510	2,360	2,040
100	100	100	100	100,000	2,810	2,560	2,020
200	140	140	140	150,000	3,190	2,680	1,300
500	220	220	220	169,200	3,280	2,640	–
700	270	270	260	200,000	3,360	2,460	–
1,000	320	320	310	250,000	3,370	1,789	–
2,500	500	500	500	286,200	3,280	–	–
5,000	710	700	700	300,000	3,210	–	–
10,000	990	990	970	400,000	2,220	–	–
25,000	1,540	1,520	1,460	455,400	–	–	–
50,000	2,120	2,040	1,870				

[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 and 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 1980 Housing Inventory and for Estimated Percentages of 1976-1980 Lost Housing Units for the Indianapolis, Ind., 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]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	50.2	50.2	50.2	50.2	50.2	50.2	50,000	0.2	0.4	1.0	1.3	1.9	2.2
200	33.5	33.5	33.5	33.5	33.5	35.5	75,000	0.13	0.4	0.8	1.1	1.6	1.8
500	16.8	16.8	16.8	16.8	19.5	22.5	100,000	0.10	0.3	0.7	1.0	1.4	1.6
700	12.6	12.6	12.6	12.6	16.4	19.0	150,000	0.07	0.3	0.6	0.8	1.1	1.3
1,000	9.2	9.2	9.2	9.5	13.8	15.9	200,000	0.05	0.2	0.5	0.7	1.0	1.1
2,500	3.9	3.9	4.4	6.0	8.7	10.0	250,000	0.04	0.2	0.4	0.6	0.9	1.0
5,000	2.0	2.0	3.1	4.3	6.2	7.1	300,000	0.03	0.2	0.4	0.6	0.8	0.9
10,000	1.0	1.0	2.2	3.0	4.4	5.0	400,000	0.03	0.2	0.3	0.5	0.7	0.8
25,000	0.4	0.6	1.4	1.9	2.8	3.2	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 and for the central city and 1.1 for the balance (not in central city).

Los Angeles-Long Beach 1980

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1980 survey. App-42
Selection of the 1977 AHS-SMSA sample App-43
1977-1980 additions to the housing inventory App-44
Sample selection for the 1977 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
1980 housing inventory App-45
1977-1980 lost housing units . . . App-46
1977 estimation procedure App-46
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-47
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-170-80) are based on data collected from the 1980 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 the Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill., SMSA's, the data were collected during the 11-month period from April 1980 through February 1981 with one-eleventh of the sample housing units being visited each month. In the remaining SMSA's, the data were collected during the 12-month period from April 1980 through March 1981 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 1980 were interviewed previously in either 1976 or 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1980 group of SMSA's, the largest SMSA from 3 of the 4 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 1980 group are: Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill.

The remaining SMSA's in the 1980 group are: Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Birmingham, Ala.; Grand Rapids, Mich.; Indianapolis, Ind.; Louisville, Ky.-Ind.; Memphis, Tenn.-Ark.; Oklahoma City, Okla.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; and Salt Lake City, Utah.

In this SMSA, 14,359 housing units were eligible for interview. Of these sample housing units, 905 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 14,359 housing units eligible for interview, 559 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 1980 survey—The sample housing units designated to be interviewed in the 1980 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 1977 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 1977 survey. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1980 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 1977 survey. (This sample represented the housing units built in permit-issuing areas, since the 1977 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1977 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1977 survey.)
5. All sample housing units that were selected as part of the 1977 and 1980 Coverage Improvement Programs. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1977 AHS-SMSA sample—The sample for the SMSA's which, in 1970, 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 1970, the following three SMSA's were 100-percent permit-issuing: Los Angeles-Long Beach, Calif.; New York, N.Y.; and Sacramento, Calif. The remaining 12 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.

1977-1980 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1977 survey, was selected to represent housing units built in permit-issuing areas since the 1977 survey. Sampling procedures were identical to those used in selecting the 1970-1977 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1977, to identify any housing units missed in the 1977 survey or any housing units added since the 1977 survey.

Sample selection for the 1977 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 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with the Oklahoma City, Okla. SMSA receiving some updating as a part of the 1980 AHS. For each of the five SMSA's interviewed previously in 1977, the Coverage Improvement Program was conducted as a part of the 1977 AHS with the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; and Salt Lake City, Utah, SMSA's receiving some updating and refining as a part of the 1980 AHS. The following discussion applies to both the prior years (1976 or 1977) and 1980 coverage improvement procedures. For the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Oklahoma City, Okla.; and Salt Lake City, Utah, SMSA's estimates of housing units added by a specific procedure reflect units added in the prior year as well as any additions that resulted from the updating and refining in 1980.

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 1980 SMSA's. Two different procedures were used. For the first procedure, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. For the SMSA's previously interviewed in 1976, these new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units). For the SMSA's previously interviewed in 1977, sample units selected from one- and two-unit structures were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample, while units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage sample selection was a sample of permit offices, and the second stage was a sample of the 1969 permits within each of the selected permit offices. In the Grand Rapids, Mich., and New York, N.Y., SMSA's, 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. This procedure was employed for the 10 SMSA's previously interviewed in 1976.

For the Albany-Schenectady-Troy, N.Y.; Memphis, Tenn.-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's previously interviewed in 1977, the above procedures were performed for the first and second stages. For the third stage, structures of size three or more were divided into clusters of an expected size of two units and a sample of clusters was selected.

For the Los Angeles-Long Beach, Calif., SMSA, 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 units were then sampled at one-third the rate of regular AHS units. Since permits were not available for all sampled offices, the procedure was also used in parts of the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 13,778 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. In the Oklahoma City, Okla., SMSA, a sample of tracts was selected but not canvassed during the 1976 Coverage Improvement Program. This procedure was completed during the 1980 AHS. This procedure added 12,746 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 AHS sample housing units was selected from the 1970 sample universe at a rate of 1 in 24 for the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Memphis, Tenn-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's and at a rate of 1 in 22 for the remaining 10 SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing unit 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. The Albany-Schenectady-Troy, N.Y., SMSA, had these sampling procedures implemented in 1976, excluding identification of the housing units missed in the 1970 census. This procedure was completed as a part of the 1980 AHS. The Los Angeles-Long Beach, Calif., and Salt Lake City, Utah, SMSA's had this sampling procedure implemented for the first time in 1980. This procedure added an estimated 75,822 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 1980 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 1980 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1977 (i.e., 1977-1980 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1980 housing inventory—The AHS estimates of characteristics of the 1980 housing inventory were produced using a 2-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Grand Rapids, Mich.; Indianapolis, Ind.; Los Angeles-Long Beach, Calif.; Louisville, Ky.-Ind.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; St. Louis, Mo.-Ill.; and Salt Lake City, Utah, 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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., SMSA's. In the Birmingham, Ala.; Memphis, Tenn.-Ark.; and Oklahoma City, Okla., 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 New York, N.Y., SMSA, the sample selected correctly apportioned the new construction housing units between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 1980 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:

$$\frac{\text{Independent estimate of the October 1980 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 procedure, 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.

1977-1980 lost housing units—The 1977-1980 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 1977 housing inventory, as was described in the 1977 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1977-1980 lost housing units existed, by definition, in the 1977 housing inventory, there was a 1977 housing inventory weight associated with each 1977-1980 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1977-1980 lost housing units.

1977 estimation procedure—This report presents data on the housing characteristics of the 1977 housing inventory from the 1977 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 1977.

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 1977 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 1980 AHS-SMSA sample reinterview program were not available at the time this report was being prepared. However, a study was conducted for the 1977 AHS-SMSA sample. The results of which are presented in the Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1977."

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 1977 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 1980 survey, the number of missed housing units may be considerably less for 1980.

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

Rounding errors—For errors associated with processing, 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 1977 housing inventory can be found in the AHS Series H-170 reports for 1977.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1980 housing inventory as well as estimates of characteristics of the 1977-1980 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 53,910 for the total SMSA, 31,260 for the central city of the SMSA, and 43,920 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-52) present the standard errors of estimated percentages for the 1980 housing inventory as well as estimated percentages of the 1977-1980 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, 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 1980 there were 1,304,300 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 12,820. 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
1,300,000	12,820
1,304,300	x
1,357,000	12,860

The entry for "x" is determined as follows by vertically interpolating between 12,820 and 12,860.

$$1{,}304{,}300 - 1{,}300{,}000 = 4{,}300$$
$$1{,}357{,}000 - 1{,}300{,}000 = 57{,}000$$

$$12{,}820 + \frac{4{,}300}{57{,}000}(12{,}860 - 12{,}820) = 12{,}820$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 1,291,480 to 1,317,120 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1980 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 1,283,790 to 1,324,810 housing units with 90 percent confidence; and that the average estimate lies within the interval from 1,278,660 to 1,329,940 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 1,304,300 owner-occupied housing units, 417,200 or 32.0 percent, had two bedrooms. Interpolation using table III of this appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 32.0 percent is approximately 0.6 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	25 or 75	32.0	50
1,300,000	0.6	a	0.7
1,304,300		p	
1,400,000	0.6	b	0.7

1. The entry for cell "a" is determined by horizontal interpolation between 0.6 and 0.7.

$$32.0-25.0 = 7.0$$
$$50.0-25.0 = 25.0$$

$$0.6 + \frac{7.0}{25.0}(0.7-0.6) = 0.6$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.6 and 0.7.

$$32.0-25.0 = 7.0$$
$$50.0-25.0 = 25.0$$

$$0.6 + \frac{7.0}{25.0}(0.7-0.6) = 0.6$$

3. The entry for "p" is then determined by vertical interpolation between 0.6 and 0.6. This interpolation is not necessary as "a and b" have the same value. The formula, however, would appear as follows.

$$1{,}304{,}300-1{,}300{,}000 = 4{,}300$$
$$1{,}400{,}000-1{,}300{,}000 = 100{,}000$$

$$0.6 + \frac{4{,}300}{100{,}000}(0.6-0.6) = 0.6$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 31.4 to 32.6 percent; the 90-percent confidence interval is from 31.0 to 33.0 percent; and the 95-percent confidence interval is from 30.8 to 33.2 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 1980 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1980 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1980 there were 614,100 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 196,900. Table I shows the standard error of 417,200 is approximately 9,070, and the standard error of 614,100 is approximately 10,550. Therefore, the standard error of the estimated difference of 196,900 is about 13,910.

$$13{,}910 = \sqrt{(9{,}070)^2 + (10{,}550)^2}$$

Consequently, the 68-percent confidence interval for the 196,900 difference is from 182,990 to 210,810 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 174,640 to 219,160 housing units, and the 95-percent confidence interval is from 169,080 to 224,720 housing units. Thus, we can conclude with 95 percent confidence that the number of 1980 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 2.7. The base of the distribution from which this median was determined is 1,304,300 housing units.

1. Interpolation using table III shows that the standard error of 50 percent on a base of 1,304,300 is approximately 0.7 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 48.6 and 51.4.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with three persons (for purposes of calculating the median, the category of three persons is considered to be from 2.5 to 3.5 persons) contains the 48.6 percent derived in step 2. About 612,000 housing units or 46.9 percent fall below this interval, and 245,300 housing units or 18.8 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(48.6-46.9)}{18.8} = 2.6$$

Similarly, the interval for owner-occupied housing units with three persons contains the 51.4 percent derived in step 2. About 612,000 housing units or 46.9 percent fall below this interval, and 245,300 housing units or 18.8 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(51.4-46.9)}{18.8} = 2.7$$

Thus, the 95-percent confidence interval ranges from 2.6 to 2.7 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1980 Housing Inventory and for Estimated Number of 1977-1980 Lost Units for the Los Angles-Long Beach, Calif., 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]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	230	200	260	500,000	9,770	7,920	9,330
100	230	200	260	600,000	10,470	8,160	9,700
200	230	200	260	700,000	11,060	8,210	9,890
500	340	320	360	800,000	11,540	8,080	9,900
700	400	370	430	900,000	11,950	7,760	9,740
1,000	480	450	510	1,000,000	12,270	7,230	9,400
2,500	760	700	810	1,100,000	12,520	6,430	8,840
5,000	1,070	990	1,140	1,200,000	12,700	5,250	8,050
10,000	1,520	1,400	1,610	1,300,000	12,820	3,290	6,920
25,000	2,390	2,210	2,530	1,357,000	12,860	–	6,060
50,000	3,370	3,090	3,540	1,400,000	12,870	–	5,250
75,000	4,110	3,750	4,300	1,500,000	12,860	–	1,940
100,000	4,720	4,290	4,930	1,514,600	12,860	–	–
150,000	5,730	5,150	5,930	1,750,000	12,560	–	–
200,000	6,560	5,820	6,720	2,000,000	11,840	–	–
250,000	7,260	6,370	7,360	2,500,000	8,640	–	–
300,000	7,880	6,810	7,910	2,871,600	–	–	–
400,000	8,920	7,490	8,750				

[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 1980 Housing Inventory and for Estimated Percentages of 1977-1980 Lost Housing Units for the Central City of the Los Angeles-Long Beach, Calif., SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
200	49.8	49.8	49.8	49.8	49.8	49.8	250,000	0.08	0.3	0.6	0.8	1.2	1.4
500	28.4	28.4	28.4	28.4	28.4	31.5	300,000	0.07	0.3	0.6	0.8	1.1	1.3
700	22.1	22.1	22.1	22.1	23.1	26.6	400,000	0.05	0.2	0.5	0.7	1.0	1.1
1,000	16.6	16.6	16.6	16.6	19.3	22.3	500,000	0.04	0.2	0.4	0.6	0.9	1.0
2,500	7.4	7.4	7.4	8.5	12.2	14.1	600,000	0.03	0.2	0.4	0.5	0.8	0.9
5,000	3.8	3.8	4.3	6.0	8.6	10.0	700,000	0.03	0.2	0.4	0.5	0.7	0.8
10,000	1.9	1.9	3.1	4.2	6.1	7.0	800,000	0.02	0.2	0.3	0.5	0.7	0.8
25,000	0.8	0.9	1.9	2.7	3.9	4.5	900,000	0.02	0.15	0.3	0.4	0.6	0.7
50,000	0.4	0.6	1.4	1.9	2.7	3.2	1,000,000	0.02	0.14	0.3	0.4	0.6	0.7
75,000	0.3	0.5	1.1	1.5	2.2	2.6	1,100,000	0.02	0.13	0.3	0.4	0.6	0.7
100,000	0.2	0.4	1.0	1.3	1.9	2.2	1,200,000	0.02	0.13	0.3	0.4	0.6	0.6
150,000	0.13	0.4	0.8	1.1	1.6	1.8	1,300,000	0.02	0.12	0.3	0.4	0.5	0.6
200,000	0.10	0.3	0.7	0.9	1.4	1.6	1,400,000	0.01	0.12	0.3	0.4	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.

TABLE III. Standard Errors for Estimated Percentages of Housing Units in the 1980 Housing Inventory and for Estimated Percentages of 1977-1980 Lost Housing Units for the Los Angeles-Long Beach, Calif., SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
500	34.2	34.2	34.2	34.2	34.2	36.0	600,000	0.04	0.2	0.5	0.6	0.9	1.0
700	27.1	27.1	27.1	27.1	27.1	30.5	700,000	0.04	0.2	0.4	0.6	0.8	1.0
1,000	20.6	20.6	20.6	20.6	22.1	25.5	800,000	0.03	0.2	0.4	0.5	0.8	0.9
2,500	9.4	9.4	9.4	9.7	14.0	16.1	900,000	0.03	0.2	0.4	0.5	0.7	0.8
5,000	4.9	4.9	5.0	6.8	9.9	11.4	1,000,000	0.03	0.2	0.4	0.5	0.7	0.8
10,000	2.5	2.5	3.5	4.8	7.0	8.1	1,100,000	0.02	0.15	0.3	0.5	0.7	0.8
25,000	1.0	1.0	2.2	3.1	4.4	5.1	1,200,000	0.02	0.15	0.3	0.4	0.6	0.7
50,000	0.5	0.7	1.6	2.2	3.1	3.6	1,300,000	0.02	0.14	0.3	0.4	0.6	0.7
75,000	0.3	0.6	1.3	1.8	2.5	2.9	1,400,000	0.02	0.14	0.3	0.4	0.6	0.7
100,000	0.3	0.5	1.1	1.5	2.2	2.5	1,500,000	0.02	0.13	0.3	0.4	0.6	0.7
150,000	0.2	0.4	0.9	1.2	1.8	2.1	1,750,000	0.01	0.12	0.3	0.4	0.5	0.6
200,000	0.13	0.4	0.8	1.1	1.6	1.8	2,000,000	0.01	0.11	0.2	0.3	0.5	0.6
250,000	0.10	0.3	0.7	1.0	1.4	1.6	2,250,000	0.01	0.11	0.2	0.3	0.5	0.5
300,000	0.09	0.3	0.6	0.9	1.3	1.5	2,500,000	0.01	0.10	0.2	0.3	0.4	0.5
400,000	0.06	0.3	0.6	0.8	1.1	1.3	3,000,000	0.01	0.09	0.2	0.3	0.4	0.5
500,000	0.05	0.2	0.5	0.7	1.0	1.1							

[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 1.3 for the central city.

Louisville 1980

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1980 survey. App-42
Selection of the 1976 AHS-SMSA sample App-43
1976-1980 additions to the housing inventory App-44
Sample selection for the 1976 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
1980 housing inventory App-45
1976-1980 lost housing units . . . App-46
1976 estimation procedure App-46
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-47
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-50

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 15 SMSA's in this report series (H-170-80) are based on data collected from the 1980 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 the Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill., SMSA's, the data were collected during the 11-month period from April 1980 through February 1981 with one-eleventh of the sample housing units being visited each month. In the remaining SMSA's, the data were collected during the 12-month period from April 1980 through March 1981 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 1980 were interviewed previously in either 1976 or 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1980 group of SMSA's, the largest SMSA from 3 of the 4 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 1980 group are: Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill.

The remaining SMSA's in the 1980 group are: Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Birmingham, Ala.; Grand Rapids, Mich.; Indianapolis, Ind.; Louisville, Ky.-Ind.; Memphis, Tenn.-Ark.; Oklahoma City, Okla.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; and Salt Lake City, Utah.

In this SMSA, 5,045 housing units were eligible for interview. Of these sample housing units, 204 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 5,045 housing units eligible for interview, 307 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 1980 survey—The sample housing units designated to be interviewed in the 1980 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 1980 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 and 1980 Coverage Improvement Programs. (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—The sample for the SMSA's which, in 1970, 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 1970, the following three SMSA's were 100-percent permit-issuing: Los Angeles-Long Beach, Calif.; New York, N.Y.; and Sacramento, Calif. The remaining 12 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.

1976-1980 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 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with the Oklahoma City, Okla. SMSA receiving some updating as a part of the 1980 AHS. For each of the five SMSA's interviewed previously in 1977, the Coverage Improvement Program was conducted as a part of the 1977 AHS with the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; and Salt Lake City, Utah, SMSA's receiving some updating and refining as a part of the 1980 AHS. The following discussion applies to both the prior year (1976 or 1977) and 1980 coverage improvement procedures. For the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Oklahoma City, Okla.; and Salt Lake City, Utah, SMSA's estimates of housing units added by a specific procedure reflect units added in the prior year as well as any additions that resulted from the updating and refining in 1980.

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 1980 SMSA's. Two different procedures were used. For the first procedure, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. For the SMSA's previously interviewed in 1976, these new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units). For the SMSA's previously interviewed in 1977, sample units selected from one- and two-unit structures were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample, while units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage sample selection was a sample of permit offices, and the second stage was a sample of the 1969 permits within each of the selected permit offices. In the Grand Rapids, Mich., and New York, N.Y., SMSA's, 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. This procedure was employed for the 10 SMSA's previously interviewed in 1976.

For the Albany-Schenectady-Troy, N.Y.; Memphis, Tenn.-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's previously interviewed in 1977, the above procedures were performed for the first and second stages. For the third stage, structures of size three or more were divided into clusters of an expected size of two units and a sample of clusters was selected.

For the Los Angeles-Long Beach, Calif., SMSA, 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 units were then sampled at one-third the rate of regular AHS units. Since permits were not available for all sampled offices, the procedure was also used in parts of the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 273 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. In the Oklahoma City, Okla., SMSA, a sample of tracts was selected but not canvassed during the 1976 Coverage Improvement Program. This procedure was completed during the 1980 AHS. This procedure added 615 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 AHS sample housing units was selected from the 1970 sample universe at a rate of 1 in 24 for the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Memphis, Tenn-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's and at a rate of 1 in 22 for the remaining 10 SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing unit 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. The Albany-Schenectady-Troy, N.Y., SMSA, had these sampling procedures implemented in 1976, excluding identification of the housing units missed in the 1970 census. This procedure was completed as a part of the 1980 AHS. The Los Angeles-Long Beach, Calif., and Salt Lake City, Utah, SMSA's had this sampling procedure implemented for the first time in 1980. This procedure added an estimated 2,778 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 3,241 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 1980 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 1980 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1976 (i.e., 1976-1980 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1980 housing inventory—The AHS estimates of characteristics of the 1980 housing inventory were produced using a 2-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Grand Rapids, Mich.; Indianapolis, Ind.; Los Angeles-Long Beach, Calif.; Louisville, Ky.-Ind.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; St. Louis, Mo.-Ill.; and Salt Lake City, Utah, 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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., SMSA's. In the Birmingham, Ala.; Memphis, Tenn.-Ark.; and Oklahoma City, Okla., 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 New York, N.Y., SMSA, the sample selected correctly apportioned the new construction housing units between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 1980 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:

$$\frac{\text{Independent estimate of the October 1980 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 procedure, 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.

1976-1980 lost housing units—The 1976-1980 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 1976-1980 lost housing units existed, by definition, in the 1976 housing inventory, there was a 1976 housing inventory weight associated with each 1976-1980 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1976-1980 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 1980 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 of which 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 1980 survey, the number of missed housing units may be considerably less for 1980.

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

Rounding errors—For errors associated with processing, 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 1980 housing inventory as well as estimates of characteristics of the 1976-1980 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 12,330 for the total SMSA, 6,810 for the central city of the SMSA, and 10,190 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-51) presents the standard errors of estimated percentages for the 1980 housing inventory as well as estimated percentages of the 1976-1980 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 1980 there were 209,300 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 2,360. 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
200,000	2,420
209,300	x
250,000	2,090

The entry for "x" is determined as follows by vertically interpolating between 2,420 and 2,090.

$$209{,}300 - 200{,}000 = 9{,}300$$
$$250{,}000 - 200{,}000 = 50{,}000$$

$$2{,}420 + \frac{9{,}300}{50{,}000}(2{,}090 - 2{,}420) = 2{,}360$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 206,940 to 211,660 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1980 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 205,520 to 213,080 housing units with 90 percent confidence; and that the average estimate lies within the interval from 204,580 to 214,020 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 209,300 owner-occupied housing units, 50,200, or 24.0 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 24.0 percent is approximately 0.9 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	24.0	25 or 75
200,000	0.6	a	0.9
209,300		p	
250,000	0.5	b	0.8

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

$$24.0-10.0=14.0$$
$$25.0-10.0=15.0$$

$$0.6+\frac{14.0}{15.0}(0.9-0.6)=0.9$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.5 and 0.8.

$$24.0-10.0=14.0$$
$$25.0-10.0=15.0$$

$$0.5+\frac{14.0}{15.0}(0.8-0.5)=0.8$$

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

$$209{,}300-200{,}000=9{,}300$$
$$250{,}000-200{,}000=50{,}000$$

$$0.9+\frac{9{,}300}{50{,}000}(0.8-0.9)=0.9$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 23.1 to 24.9 percent; the 90-percent confidence interval is from 22.6 to 25.4 percent; and the 95-percent confidence interval is from 22.2 to 25.8 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; if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1976 and 1980 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1980 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1980 there were 111,400 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 61,200. Table I shows the standard error of 50,200 is approximately 1,800, and the standard error of 111,400 is approximately 2,360. Therefore, the standard error of the estimated difference of 61,200 is about 2,970:

$$2{,}970=\sqrt{(1{,}800)^2+(2{,}360)^2}$$

Consequently, the 68-percent confidence interval for the 61,200 difference is from 58,230 to 64,170 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 56,450 to 65,950 housing units, and the 95-percent confidence interval is from 55,260 to 67,140 housing units. Thus, we can conclude with 95 percent confidence that the number of 1980 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 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 209,300 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 209,300 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 three persons (for purposes of calculating the median, the category of three persons is considered to be from 2.5 to 3.5 persons) contains the 48.0 percent derived in step 2. About 94,000 housing units or 44.9 percent fall below this interval, and 44,200 housing units or 21.1 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(48.0-44.9)}{21.1} = 2.6$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.0 percent derived in step 2. About 94,000 housing units or 44.9 percent fall below this interval, and 44,200 housing units or 21.1 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(52.0-44.9)}{21.1} = 2.8$$

Thus, the 95-percent confidence interval ranges from 2.6 to 2.8 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1980 Housing Inventory and for Estimated Number of 1976-1980 Lost Units for the Louisville, Ky.-Ind., 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]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	80	70	80	50,000	1,800	1,410	1,720
100	90	80	90	75,000	2,100	1,410	1,920
200	120	110	130	100,000	2,300	1,170	1,980
500	200	180	200	126,100	2,430	—	1,900
700	230	210	240	150,000	2,480	—	1,700
1,000	280	260	280	198,100	2,430	—	—
2,500	440	400	440	200,000	2,420	—	—
5,000	610	560	620	250,000	2,090	—	—
10,000	860	780	870	300,000	1,310	—	—
25,000	1,330	1,150	1,320	324,100	—	—	—

[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, 1.2 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 H. Standard Errors for Estimated Percentages of Housing Units in the 1980 Housing Inventory and for Estimated Percentages of 1976-1980 Lost Housing Units for the Louisville, Ky.-Ind., SMSA, for the Central City and for the Balance (Not in the Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	44.2	44.2	44.2	44.2	44.2	44.5	50,000	0.2	0.4	0.9	1.2	1.7	2.0
200	28.4	28.4	28.4	28.4	28.4	31.5	75,000	0.11	0.3	0.7	1.0	1.4	1.6
500	13.7	13.7	13.7	13.7	17.2	19.9	100,000	0.08	0.3	0.6	0.8	1.2	1.4
700	10.2	10.2	10.2	10.2	14.6	16.8	150,000	0.05	0.2	0.5	0.7	1.0	1.1
1,000	7.3	7.3	7.3	8.4	12.2	14.1	200,000	0.04	0.2	0.4	0.6	0.9	1.0
2,500	3.1	3.1	3.9	5.3	7.7	8.9	250,000	0.03	0.2	0.4	0.5	0.8	0.9
5,000	1.6	1.6	2.7	3.8	5.4	6.3	300,000	0.03	0.2	0.4	0.5	0.7	0.8
10,000	0.8	0.9	1.9	2.7	3.9	4.4	400,000	0.02	0.14	0.3	0.4	0.6	0.7
25,000	0.3	0.6	1.2	1.7	2.4	2.8							

[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 1.1 for the total SMSA, 1.2 for the central city, and 1.1 for the balance (not in central city).

Memphis 1980

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1980 survey. App-42
Selection of the 1977 AHS-SMSA sample App-43
1977-1980 additions to the housing inventory App-44
Sample selection for the 1977 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
1980 housing inventory App-45
1977-1980 lost housing units . . . App-46
1977 estimation procedure App-46
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-47
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-170-80) are based on data collected from the 1980 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 the Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill., SMSA's, the data were collected during the 11-month period from April 1980 through February 1981 with one-eleventh of the sample housing units being visited each month. In the remaining SMSA's, the data were collected during the 12-month period from April 1980 through March 1981 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 1980 were interviewed previously in either 1976 or 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1980 group of SMSA's, the largest SMSA from 3 of the 4 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 1980 group are: Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill.

The remaining SMSA's in the 1980 group are: Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Birmingham, Ala.; Grand Rapids, Mich.; Indianapolis, Ind.; Louisville, Ky.-Ind.; Memphis, Tenn.-Ark.; Oklahoma City, Okla.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; and Salt Lake City, Utah.

In this SMSA, 5,159 housing units were eligible for interview. Of these sample housing units, 157 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 5,159 housing units eligible for interview, 293 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 1980 survey—The sample housing units designated to be interviewed in the 1980 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 1977 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 1977 survey. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1980 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 1977 survey. (This sample represented the housing units built in permit-issuing areas, since the 1977 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1977 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1977 survey.)
5. All sample housing units that were selected as part of the 1977 and 1980 Coverage Improvement Programs. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1977 AHS-SMSA sample—The sample for the SMSA's which, in 1970, 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 1970, the following three SMSA's were 100-percent permit-issuing: Los Angeles-Long Beach, Calif.; New York, N.Y.; and Sacramento, Calif. The remaining 12 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.

1977-1980 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1977 survey, was selected to represent housing units built in permit-issuing areas since the 1977 survey. Sampling procedures were identical to those used in selecting the 1970-1977 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1977, to identify any housing units missed in the 1977 survey or any housing units added since the 1977 survey.

Sample selection for the 1977 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 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with the Oklahoma City, Okla. SMSA receiving some updating as a part of the 1980 AHS. For each of the five SMSA's interviewed previously in 1977, the Coverage Improvement Program was conducted as a part of the 1977 AHS with the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; and Salt Lake City, Utah, SMSA's receiving some updating and refining as a part of the 1980 AHS. The following discussion applies to both the prior year (1976 or 1977) and 1980 coverage improvement procedures. For the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Oklahoma City, Okla.; and Salt Lake City, Utah, SMSA's estimates of housing units added by a specific procedure reflect units added in the prior year as well as any additions that resulted from the updating and refining in 1980.

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 1980 SMSA's. Two different procedures were used. For the first procedure, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. For the SMSA's previously interviewed in 1976, these new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units). For the SMSA's previously interviewed in 1977, sample units selected from one- and two-unit structures were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample, while units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage sample selection was a sample of permit offices, and the second stage was a sample of the 1969 permits within each of the selected permit offices. In the Grand Rapids, Mich., and New York, N.Y., SMSA's, 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. This procedure was employed for the 10 SMSA's previously interviewed in 1976.

For the Albany-Schenectady-Troy, N.Y.; Memphis, Tenn.-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's previously interviewed in 1977, the above procedures were performed for the first and second stages. For the third stage, structures of size three or more were divided into clusters of an expected size of two units and a sample of clusters was selected.

For the Los Angeles-Long Beach, Calif., SMSA, 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 units were then sampled at one-third the rate of regular AHS units. Since permits were not available for all sampled offices, the procedure was also used in parts of the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 3,342 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. In the Oklahoma City, Okla., SMSA, a sample of tracts was selected but not canvassed during the 1976 Coverage Improvement Program. This procedure was completed during the 1980 AHS and added 958 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 AHS sample housing units was selected from the 1970 sample universe at a rate of 1 in 24 for the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Memphis, Tenn-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's and at a rate of 1 in 22 for the remaining 10 SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing unit 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. The Albany-Schenectady-Troy, N.Y., SMSA, had these sampling procedures implemented in 1976, excluding identification of the housing units missed in the 1970 census. This procedure was completed as a part of the 1980 AHS. The Los Angeles-Long Beach, Calif., and Salt Lake City, Utah, SMSA's had this sampling procedure implemented for the first time in 1980. This procedure added an estimated 8,310 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 993 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 1980 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 1980 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1977 (i.e., 1977-1980 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1980 housing inventory—The AHS estimates of characteristics of the 1980 housing inventory were produced using a 2-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York,N.Y.; and Oklahoma City, Okla., 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 Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Grand Rapids, Mich.; Indianapolis, Ind.; Los Angeles-Long Beach, Calif.; Louisville, Ky.-Ind.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; St. Louis, Mo.-Ill.; and Salt Lake City, Utah, 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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., SMSA's. In the Birmingham, Ala.; Memphis, Tenn.-Ark.; and Oklahoma City, Okla., 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 New York, N.Y., SMSA, the sample selected correctly apportioned the new construction housing units between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 1980 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:

$$\frac{\text{Independent estimate of the October 1980 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 procedure, 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.

1977-1980 lost housing units—The 1977-1980 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 1977 housing inventory, as was described in the 1977 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1977-1980 lost housing units existed, by definition, in the 1977 housing inventory, there was a 1977 housing inventory weight associated with each 1977-1980 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1977-1980 lost housing units.

1977 estimation procedure—This report presents data on the housing characteristics of the 1977 housing inventory from the 1977 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 1977.

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 1977 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 1980 AHS-SMSA sample reinterview program were not available at the time this report was being prepared. However, a study was conducted for the 1977 AHS-SMSA sample. The results of which are presented in the Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1977."

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 1977 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 1980 survey, the number of missed housing units may be considerably less for 1980.

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

Rounding errors—For errors associated with processing, 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 1977 housing inventory can be found in the AHS Series H-170 reports for 1977.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1980 housing inventory as well as estimates of characteristics of the 1977-1980 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 13,600 for the total SMSA, 8,820 for the central city of the SMSA, and 10,260 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 1980 housing inventory as well as estimated percentages of the 1977-1980 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 1980 there were 166,800 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 2,250. 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
150,000	2,290
166,800	x
200,000	2,180

The entry for "x" is determined as follows by vertically interpolating between 2,290 and 2,180.

$$166{,}800-150{,}000 = 16{,}800$$
$$200{,}000-150{,}000 = 50{,}000$$

$$2{,}290+\frac{16{,}800}{50{,}000}(2{,}180-2{,}290) = 2{,}250$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 164,550 to 169,050 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1980 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 163,200 to 170,400 housing units with 90 percent confidence; and that the average estimate lies within the interval from 162,300 to 171,300 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 166,800 owner-occupied housing units, 40,500, or 24.3 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 24.3 percent is approximately 0.9 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	24.3	25 or 75
150,000.	0.7	a	1.0
166,800.		p	
200,000.	0.6	b	0.8

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

$$24.3-10.0 = 14.3$$
$$25.0-10.0 = 15.0$$

$$0.7 + \frac{14.3}{15.0}(1.0-0.7) = 1.0$$

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

$$24.3-10.0 = 14.3$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{14.3}{15.0}(0.8-0.6) = 0.8$$

3. The entry for "p" is then determined by vertical interpolation between 1.0 and 0.8.

$$166{,}800-150{,}000 = 16{,}800$$
$$200{,}000-150{,}000 = 50{,}000$$

$$1.0 + \frac{16{,}800}{50{,}000}(0.8-1.0) = 0.9$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 23.4 to 25.2 percent; the 90-percent confidence interval is from 22.9 to 25.7 percent; and the 95-percent confidence interval is from 22.5 to 26.1 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; if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1977 and 1980 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1977 and 1980 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1980 there were 92,100 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 51,600. Table I shows the standard error of 40,500 is approximately 1,520, and the standard error of 92,100 is approximately 2,100. Therefore, the standard error of the estimated difference of 51,600 is about 2,590:

$$2{,}590 = \sqrt{(1{,}520)^2 + (2{,}100)^2}$$

Consequently, the 68-percent confidence interval for the 51,600 difference is from 49,010 to 54,190 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 47,460 to 55,740 housing units, and the 95-percent confidence interval is from 46,420 to 56,780 housing units. Thus, we can conclude with 95 percent confidence that the number of 1980 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 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 for 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 is determined is 166,800 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 166,800 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 three persons (for purposes of calculating the median, the category of three persons is considered to be from 2.5 to 3.5 persons) contains the 47.8 percent derived in step 2. About 76,100 housing units or 45.6 percent fall below this interval, and 33,700 housing units or 20.2 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(47.8-45.6)}{20.2} = 2.6$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.2 percent derived in step 2. About 76,100 housing units or 45.6 percent fall below this interval, and 33,700 housing units or 20.2 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(52.2-45.6)}{20.2} = 2.8$$

Thus, the 95-percent confidence interval ranges from 2.6 to 2.8 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1980 Housing Inventory and for Estimated Number of 1977-1980 Lost Units for the Memphis, Tenn.-Ark., 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]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	70	60	80	50,000	1,690	1,580	1,250
100	80	80	90	75,000	1,970	1,790	810
200	120	110	120	84,700	2,050	1,840	–
500	180	180	200	100,000	2,150	1,890	–
700	220	210	230	150,000	2,290	1,770	–
1,000	260	250	280	200,000	2,180	1,130	–
2,500	410	400	430	222,000	2,050	–	–
5,000	580	560	600	250,000	1,780	–	–
10,000	810	790	820	300,000	670	–	–
25,000	1,250	1,200	1,160	306,700	–	–	–

[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 1.0 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 derive these estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1980 Housing Inventory and for Estimated Percentages of 1977-1980 Lost Housing Units for the Memphis, Tenn.-Ark., 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]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	43.4	43.4	43.4	43.4	43.4	43.8	50,000	0.15	0.4	0.9	1.2	1.7	2.0
200	27.7	27.7	27.7	27.7	27.7	30.9	75,000	0.10	0.3	0.7	1.0	1.4	1.6
500	13.3	13.3	13.3	13.3	16.9	19.6	100,000	0.08	0.3	0.6	0.8	1.2	1.4
700	9.9	9.9	9.9	9.9	14.3	16.5	150,000	0.05	0.2	0.5	0.7	1.0	1.1
1,000	7.1	7.1	7.1	8.3	12.0	13.8	200,000	0.04	0.2	0.4	0.6	0.8	1.0
2,500	3.0	3.0	3.8	5.3	7.6	8.8	250,000	0.03	0.2	0.4	0.5	0.8	0.9
5,000	1.5	1.5	2.7	3.7	5.4	6.2	300,000	0.03	0.2	0.3	0.5	0.7	0.8
10,000	0.8	0.9	1.9	2.6	3.8	4.4	400,000	0.02	0.14	0.3	0.4	0.6	0.7
25,000	0.3	0.6	1.2	1.7	2.4	2.8							

[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 1.1 for the total SMSA, 1.3 for the central city, and 1.0 for the balance (not in central city).

New York 1980

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1980 survey. App-42
Selection of the 1976 AHS-SMSA sample App-43
1976-1980 additions to the housing inventory App-44
Sample selection for the 1976 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
1980 housing inventory App-45
1976-1980 lost housing units . . . App-46
1976 estimation procedure App-46
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-47
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-170-80) are based on data collected from the 1980 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 the Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill., SMSA's, the data were collected during the 11-month period from April 1980 through February 1981 with one-eleventh of the sample housing units being visited each month. In the remaining SMSA's, the data were collected during the 12-month period from April 1980 through March 1981 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 1980 were interviewed previously in either 1976 or 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1980 group of SMSA's, the largest SMSA from 3 of the 4 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 1980 group are: Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill.

The remaining SMSA's in the 1980 group are: Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Birmingham, Ala.; Grand Rapids, Mich.; Indianapolis, Ind.; Louisville, Ky.-Ind.; Memphis, Tenn.-Ark.; Oklahoma City, Okla.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; and Salt Lake City, Utah.

In this SMSA, 13,727 housing units were eligible for interview. Of these sample housing units, 965 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 13,727 housing units eligible for interview, 1,096 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 1980 survey—The sample housing units designated to be interviewed in the 1980 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 1980 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 and 1980 Coverage Improvement Programs. (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—The sample for the SMSA's which, in 1970, 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 1970, the following three SMSA's were 100-percent permit-issuing: Los Angeles-Long Beach, Calif.; New York, N.Y.; and Sacramento, Calif. The remaining 12 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 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.

1976-1980 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 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with the Oklahoma City, Okla. SMSA receiving some updating as a part of the 1980 AHS. For each of the five SMSA's interviewed previously in 1977, the Coverage Improvement Program was conducted as a part of the 1977 AHS with the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; and Salt Lake City, Utah, SMSA's receiving some updating and refining as a part of the 1980 AHS. The following discussion applies to both the prior years (1976 or 1977) and 1980 coverage improvement procedures. For the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Oklahoma City, Okla.; and Salt Lake City, Utah, SMSA's estimates of housing units added by a specific procedure reflect units added in the prior year as well as any additions that resulted from the updating and refining in 1980.

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 1980 SMSA's. Two different procedures were used. For the first procedure, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. For the SMSA's previously interviewed in 1976, these new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units). For the SMSA's previously interviewed in 1977, sample units selected from one- and two-unit structures were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample, while units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage sample selection was a sample of permit offices, and the second stage was a sample of the 1969 permits within each of the selected permit offices. In the Grand Rapids, Mich., and New York, N.Y., SMSA's, 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. This procedure was employed for the 10 SMSA's previously interviewed in 1976.

For the Albany-Schenectady-Troy, N.Y.; Memphis, Tenn.-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's previously interviewed in 1977, the above procedures were performed for the first and second stages. For the third stage, structures of size three or more were divided into clusters of an expected size of two units and a sample of clusters was selected.

For the Los Angeles-Long Beach, Calif.; SMSA, 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 units were then sampled at one-third the rate of regular AHS units. Since permits were not available for all sampled offices, the procedure was also used in parts of the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 11,162 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. In the Oklahoma City, Okla., SMSA, a sample of tracts was selected but not canvassed during the 1976 Coverage Improvement Program. This procedure was completed during the 1980 AHS. 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 AHS sample housing units was selected from the 1970 sample universe at a rate of 1 in 24 for the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Memphis, Tenn-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's and at a rate of 1 in 22 for the remaining 10 SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing unit 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. The Albany-Schenectady-Troy, N.Y., SMSA, had these sampling procedures implemented in 1976, excluding identification of the housing units missed in the 1970 census. This procedure was completed as a part of the 1980 AHS. The Los Angeles-Long Beach, Calif., and Salt Lake City, Utah, SMSA's had this sampling procedure implemented for the first time in 1980. This procedure added an estimated 76,968 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 14,000 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 1980 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 1980 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1976 (i.e., 1976-1980 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1980 housing inventory—The AHS estimates of characteristics of the 1980 housing inventory were produced using a 2-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J., Grand Rapids, Mich.; Indianapolis, Ind.; Los Angeles-Long Beach, Calif.; Louisville, Ky.-Ind.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; St. Louis, Mo.-Ill.; and Salt Lake City, Utah, 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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., SMSA's. In the Birmingham, Ala.; Memphis, Tenn.-Ark.; and Oklahoma City, Okla., 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 New York, N.Y., SMSA, the sample selected correctly apportioned the new construction housing units between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 1980 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:

$$\frac{\text{Independent estimate of the October 1980 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 procedure, 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.

1976-1980 lost housing units—The 1976-1980 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 1976-1980 lost housing units existed, by definition, in the 1976 housing inventory, there was a 1976 housing inventory weight associated with each 1976-1980 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1976-1980 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 1980 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 of which 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 1980 survey, the number of missed housing units may be considerably less for 1980.

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

Rounding errors—For errors associated with processing, 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 1976 housing inventory can be found in the AHS Series H-170 reports for 1976.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1980 housing inventory as well as estimates of characteristics of the 1976-1980 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 92,610 for the total SMSA, 77,360 for the central city of the SMSA, and 50,840 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-52) present the standard errors of estimated percentages for the 1980 housing inventory as well as estimated percentages of the 1976-1980 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, 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 1980 there were 1,650,800 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 20,170. 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
1,500,000	19,910
1,650,800	x
2,000,000	20,770

The entry for "x" is determined as follows by vertically interpolating between 19,910 and 20,770.

$$1{,}650{,}800-1{,}500{,}000 = 150{,}800$$
$$2{,}000{,}000-1{,}500{,}000 = 500{,}000$$

$$19{,}910 + \frac{150{,}800}{500{,}000}(20{,}770-19{,}910) = 20{,}170$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 1,630,630 to 1,670,970 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1980 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 1,618,530 to 1,683,070 housing units with 90 percent confidence; and that the average estimate lies within the interval from 1,610,460 to 1,691,140 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 1,650,800 owner-occupied housing units, 372,000, or 22.5 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 22.5 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	22.5	25 or 75
1,500,000	0.5	a	0.8
1,650,800		p	
1,750,000	0.5	b	0.7

1. The entry for cell "a" is determined by horizontal interpolation between 0.5 and 0.8.

$$22.5-10.0 = 12.5$$
$$25.0-10.0 = 15.0$$

$$0.5 + \frac{12.5}{15.0}(0.8-0.5) = 0.8$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.5 and 0.7.

$$22.5-10.0 = 12.5$$
$$25.0-10.0 = 15.0$$

$$0.5 + \frac{12.5}{15.0}(0.7-0.5) = 0.7$$

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

$$1{,}650{,}800-1{,}500{,}000 = 150{,}800$$
$$1{,}750{,}000-1{,}500{,}000 = 250{,}000$$

$$0.8 + \frac{150{,}800}{250{,}000}(0.7-0.8) = 0.7$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 21.8 to 23.2 percent; the 90-percent confidence interval is from 21.4 to 23.6 percent; and the 95-percent confidence interval is from 21.1 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 1980 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1980 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1980 there were 749,100 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 377,100. Table I shows the standard error of 372,000 is approximately 11,770, and the standard error of 749,100 is approximately 15,880. Therefore, the standard error of the estimated difference of 377,100 is about 19,770:

$$19{,}770 = \sqrt{(11{,}770)^2 + (15{,}880)^2}$$

Consequently, the 68-percent confidence interval for the 377,100 difference is from 357,330 to 396,870 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 345,470 to 408,730 housing units, and the 95-percent confidence interval is from 337,560 to 416,640 housing units. Thus, we can conclude with 95 percent confidence that the number of 1980 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 3.1. The base of the distribution from which this median was determined is 1,650,800 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 1,650,800 is approximately 0.8 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 48.4 and 51.6.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with three persons (for purposes of calculating the median, the category of three persons is considered to be from 2.5 to 3.5 persons) contains the 48.4 percent derived in step 2. About 636,800 housing units or 38.6 percent fall below this interval, and 321,000 housing units or 19.4 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(48.4-38.6)}{19.4} = 3.0$$

Similarly, the interval for owner-occupied housing units with three persons contains the 51.6 percent derived in step 2. About 636,800 housing units or 38.6 percent fall below this interval, and 321,000 housing units or 19.4 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(51.6-38.6)}{19.4} = 3.2$$

Thus, the 95-percent confidence interval ranges from 3.0 to 3.2 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1980 Housing Inventory and for Estimated Number of 1976-1980 Lost Units for the New York, 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]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	410	500	180	400,000	12,190	13,160	7,070
100	410	500	180	500,000	13,440	14,420	7,430
200	410	500	190	600,000	14,530	15,480	7,600
500	450	500	300	700,000	15,470	16,360	7,570
700	540	590	360	800,000	16,310	17,090	7,350
1,000	640	710	430	900,000	17,690	17,700	6,920
2,500	1,010	1,120	670	1,000,000	18,260	18,200	6,230
5,000	1,430	1,580	950	1,270,800	19,080	19,030	–
10,000	2,020	2,230	1,340	1,500,000	19,910	19,210	–
25,000	3,190	3,520	2,110	2,000,000	20,770	17,950	–
50,000	4,500	4,960	2,960	2,500,000	20,440	13,790	–
75,000	5,500	6,050	3,590	2,947,500	19,080	–	–
100,000	6,330	6,960	4,100	3,000,000	18,850	–	–
150,000	7,700	8,450	4,910	3,500,000	15,630	–	–
200,000	8,840	9,670	5,540	4,000,000	9,210	–	–
250,000	9,820	10,710	6,050	4,218,300	–	–	–
300,000	10,690	11,620	6,460				

[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 1980 Housing Inventory and for Estimated Percentages of 1976-1980 Lost Housing Units for the New York, N.Y., SMSA and Central City of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
700	41.7	41.7	41.7	41.7	41.7	42.3	800,000	0.06	0.2	0.5	0.8	1.1	1.3
1,000	33.4	33.4	33.4	33.4	33.4	35.4	900,000	0.06	0.2	0.5	0.7	1.0	1.2
2,500	16.7	16.7	16.7	16.7	19.4	22.4	1,000,000	0.05	0.2	0.5	0.7	1.0	1.1
5,000	9.1	9.1	9.1	9.5	13.7	15.8	1,100,000	0.05	0.2	0.5	0.6	0.9	1.1
10,000	4.8	4.8	4.9	6.7	9.7	11.2	1,200,000	0.04	0.2	0.4	0.6	0.9	1.0
25,000	2.0	2.0	3.1	4.2	6.1	7.1	1,300,000	0.04	0.2	0.4	0.6	0.9	1.0
50,000	1.0	1.0	2.2	3.0	4.3	5.0	1,400,000	0.04	0.2	0.4	0.6	0.8	0.9
75,000	0.7	0.8	1.8	2.5	3.5	4.1	1,500,000	0.03	0.2	0.4	0.5	0.8	0.9
100,000	0.5	0.7	1.5	2.1	3.1	3.5	1,750,000	0.03	0.2	0.4	0.5	0.7	0.8
150,000	0.3	0.6	1.3	1.7	2.5	2.9	2,000,000	0.03	0.2	0.3	0.5	0.7	0.8
200,000	0.2	0.5	1.1	1.5	2.2	2.5	2,250,000	0.02	0.15	0.3	0.4	0.6	0.7
250,000	0.2	0.4	1.0	1.3	1.9	2.2	2,500,000	0.02	0.14	0.3	0.4	0.6	0.7
300,000	0.2	0.4	0.9	1.2	1.8	2.0	3,000,000	0.02	0.13	0.3	0.4	0.6	0.6
400,000	0.13	0.4	0.8	1.1	1.5	1.8	3,500,000	0.01	0.12	0.3	0.4	0.5	0.6
500,000	0.10	0.3	0.7	0.9	1.4	1.6	4,000,000	0.01	0.11	0.2	0.3	0.5	0.6
600,000	0.08	0.3	0.6	0.9	1.3	1.4	4,250,000	0.01	0.10	0.2	0.3	0.5	0.5
700,000	0.07	0.3	0.6	0.8	1.2	1.3							

[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 1.3 for the central city.

TABLE III. Standard Errors for Estimated Percentages of Housing Units in the 1980 Housing Inventory and for Estimated Percentages of 1976-1980 Lost Housing Units for the Balance (Outside Central City) of the New York, N.Y., SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
200	47.7	47.7	47.7	47.7	47.7	47.7	250,000	0.07	0.3	0.6	0.8	1.2	1.4
500	26.7	26.7	26.7	26.7	26.7	30.2	300,000	0.06	0.2	0.5	0.7	1.1	1.2
700	20.7	20.7	20.7	20.7	22.1	25.5	400,000	0.05	0.2	0.5	0.6	0.9	1.1
1,000	15.4	15.4	15.4	15.4	18.5	21.3	500,000	0.04	0.2	0.4	0.6	0.8	1.0
2,500	6.8	6.8	6.8	8.1	11.7	13.5	600,000	0.03	0.2	0.4	0.5	0.8	0.9
5,000	3.5	3.5	4.2	5.7	8.3	9.5	700,000	0.03	0.2	0.4	0.5	0.7	0.8
10,000	1.8	1.8	2.9	4.1	5.8	6.8	800,000	0.02	0.15	0.3	0.5	0.7	0.8
25,000	0.7	0.8	1.9	2.6	3.7	4.3	900,000	0.02	0.14	0.3	0.4	0.6	0.7
50,000	0.4	0.6	1.3	1.8	2.6	3.0	1,000,000	0.02	0.13	0.3	0.4	0.6	0.7
75,000	0.2	0.5	1.1	1.5	2.1	2.5	1,100,000	0.02	0.13	0.3	0.4	0.6	0.6
100,000	0.2	0.4	0.9	1.3	1.8	2.1	1,200,000	0.02	0.12	0.3	0.4	0.5	0.6
150,000	0.12	0.3	0.8	1.0	1.5	1.7	1,300,000	0.01	0.12	0.3	0.4	0.5	0.6
200,000	0.09	0.3	0.7	0.9	1.3	1.5							

[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.

Oklahoma City 1980

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1980 survey. App-42
Selection of the 1976 AHS-SMSA sample App-43
1976-1980 additions to the housing inventory App-44
Sample selection for the 1976 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
1980 housing inventory App-45
1976-1980 lost housing units . . . App-46
1976 estimation procedure App-46
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-47
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-170-80) are based on data collected from the 1980 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 the Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill., SMSA's, the data were collected during the 11-month period from April 1980 through February 1981 with one-eleventh of the sample housing units being visited each month. In the remaining SMSA's, the data were collected during the 12-month period from April 1980 through March 1981 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 1980 were interviewed previously in either 1976 or 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1980 group of SMSA's, the largest SMSA from 3 of the 4 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 1980 group are: Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill.

The remaining SMSA's in the 1980 group are: Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Birmingham, Ala.; Grand Rapids, Mich.; Indianapolis, Ind.; Louisville, Ky.-Ind.; Memphis, Tenn.-Ark.; Oklahoma City, Okla.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; and Salt Lake City, Utah.

In this SMSA, 5,252 housing units were eligible for interview. Of these sample housing units, 177 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 5,252 housing units eligible for interview, 446 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 1980 survey—The sample housing units designated to be interviewed in the 1980 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 1980 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 and 1980 Coverage Improvement Programs. (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—The sample for the SMSA's which, in 1970, 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 1970, the following three SMSA's were 100-percent permit-issuing: Los Angeles-Long Beach, Calif.; New York, N.Y.; and Sacramento, Calif. The remaining 12 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.

1976-1980 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 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with the Oklahoma City, Okla. SMSA receiving some updating as a part of the 1980 AHS. For each of the five SMSA's interviewed previously in 1977, the Coverage Improvement Program was conducted as a part of the 1977 AHS with the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; and Salt Lake City, Utah, SMSA's receiving some updating and refining as a part of the 1980 AHS. The following discussion applies to both the prior year (1976 or 1977) and 1980 coverage improvement procedures: For the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Oklahoma City, Okla.; and Salt Lake City, Utah, SMSA's estimates of housing units added by a specific procedure reflect units added in the prior year as well as any additions that resulted from the updating and refining in 1980.

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 1980 SMSA's. Two different procedures were used. For the first procedure, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. For the SMSA's previously interviewed in 1976, these new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units). For the SMSA's previously interviewed in 1977, sample units selected from one- and two-unit structures were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample, while units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage sample selection was a sample of permit offices, and the second stage was a sample of the 1969 permits within each of the selected permit offices. In the Grand Rapids, Mich., and New York, N.Y., SMSA's, 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. This procedure was employed for the 10 SMSA's previously interviewed in 1976.

For the Albany-Schenectady-Troy, N.Y.; Memphis, Tenn.-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's previously interviewed in 1977, the above procedures were performed for the first and second stages. For the third stage, structures of size three or more were divided into clusters of an expected size of two units and a sample of clusters was selected.

For the Los Angeles-Long Beach, Calif., SMSA, 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 units were then sampled at one-third the rate of regular AHS units. Since permits were not available for all sampled offices, the procedure was also used in parts of the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 2,419 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. In the Oklahoma City, Okla., SMSA, a sample of tracts was selected but not canvassed during the 1976 Coverage Improvement Program. This procedure was completed during the 1980 AHS and added 1,389 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 AHS sample housing units was selected from the 1970 sample universe at a rate of 1 in 24 for the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Memphis, Tenn-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's and at a rate of 1 in 22 for the remaining 10 SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing unit 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. The Albany-Schenectady-Troy, N.Y., SMSA, had these sampling procedures implemented in 1976, excluding identification of the housing units missed in the 1970 census. This procedure was completed as a part of the 1980 AHS. The Los Angeles-Long Beach, Calif., and Salt Lake City, Utah, SMSA's had this sampling procedure implemented for the first time in 1980. This procedure added an estimated 2,345 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 572 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 1980 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 1980 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1976 (i.e., 1976-1980 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1980 housing inventory—The AHS estimates of characteristics of the 1980 housing inventory were produced using a 2-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Grand Rapids, Mich.; Indianapolis, Ind.; Los Angeles-Long Beach, Calif.; Louisville, Ky.-Ind.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; St. Louis, Mo.-Ill.; and Salt Lake City, Utah, 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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., SMSA's. In the Birmingham, Ala.; Memphis, Tenn.-Ark.; and Oklahoma City, Okla., 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 New York, N.Y., SMSA, the sample selected correctly apportioned the new construction housing units between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 1980 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:

$$\frac{\text{Independent estimate of the October 1980 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 procedure, 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.

1976-1980 lost housing units—The 1976-1980 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 1976-1980 lost housing units existed, by definition, in the 1976 housing inventory, there was a 1976 housing inventory weight associated with each 1976-1980 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1976-1980 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 1980 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 of which 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 1980 survey, the number of missed housing units may be considerably less for 1980.

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

Rounding errors—For errors associated with processing, 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 1976 housing inventory can be found in the AHS Series H-170 reports for 1976.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1980 housing inventory as well as estimates of characteristics of the 1977-1980 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 14,540 for the total SMSA, 9,640 for the central city of the SMSA, and 10,790 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 1980 housing inventory as well as estimated percentages of the 1976-1980 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 1980 there were 189,600 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 2,280. 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
179,100	2,310
189,600	x
200,000	2,250

The entry for "x" is determined as follows by vertically interpolating between 2,310 and 2,250.

$$189{,}600 - 179{,}100 = 10{,}500$$
$$200{,}000 - 179{,}100 = 20{,}900$$

$$2{,}310 + \frac{10{,}500}{20{,}900}(2{,}250 - 2{,}310) = 2{,}280$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 187,320 to 191,880 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1980 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 185,950 to 193,250 housing units with 90 percent confidence; and that the average estimate lies within the interval from 185,040 to 194,160 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 189,600 owner-occupied housing units, 53,800, or 28.4 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 28.4 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	28.4	50
150,000.	0.9	a	1.1
189,600.		p	
200,000.	0.8	b	0.9

1. The entry for cell "a" is determined by horizontal interpolation between 0.9 and 1.1.

$$28.4-25.0 = 3.4$$
$$50.0-25.0 = 25.0$$

$$0.9 + \frac{3.4}{25.0}(1.1-0.9) = 0.9$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.8 and 0.9.

$$28.4-25.0 = 3.4$$
$$50.0-25.0 = 25.0$$

$$0.8 + \frac{3.4}{25.0}(0.9-0.8) = 0.8$$

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

$$189{,}600-150{,}000 = 39{,}600$$
$$200{,}000-150{,}000 = 50{,}000$$

$$0.9 + \frac{39{,}600}{50{,}000}(0.8-0.9) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 27.6 to 29.2 percent; the 90-percent confidence interval is from 27.1 to 29.7 percent; and the 95-percent confidence interval is from 26.8 to 30.0 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; if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1976 and 1980 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1980 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1980 there were 107,500 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 53,700. Table I shows the standard error of 53,800 is approximately 1,740, and the standard error of 107,500 is approximately 2,200. Therefore, the standard error of the estimated difference of 53,700 is about 2,800:

$$2{,}800 = \sqrt{(1{,}740)^2 + (2{,}200)^2}$$

Consequently, the 68-percent confidence interval for the 53,700 difference is from 50,900 to 56,500 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 49,220 to 58,180 housing units, and the 95-percent confidence interval is from 48,100 to 59,300 housing units. Thus, we can conclude with 95 percent confidence that the number of 1980 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 corrsponding 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 2.4. The base of the distribution from which this median was determined is 189,600 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 189,600 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 tables 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 31,800 housing units or 16.8 percent fall below this interval, and 66,400 housing units or 35.0 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$1.5 + (2.5-1.5)\,\frac{(48.2-16.8)}{35.0} = 2.4$$

Similarly, the interval for owner-occupied housing units with two persons contains the 51.8 percent derived in step 2. About 31,800 housing units or 16.8 percent fall below this interval, and 66,400 housing units or 35.0 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$1.5 + (2.5-1.5)\,\frac{(51.8-16.8)}{35.0} = 2.5$$

Thus, the 95-percent confidence interval ranges from 2.4 to 2.5 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1980 Housing Inventory and for Estimated Number of 1976-1980 Lost Units for the Oklahoma City, Okla., 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]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	70	70	70	50,000	1,700	1,550	1,510
100	80	80	80	75,000	1,980	1,700	1,560
200	120	120	120	100,000	2,170	1,710	1,400
500	190	180	190	137,200	2,310	1,460	—
700	220	220	220	150,000	2,330	1,270	—
1,000	260	260	270	179,100	2,310	—	—
2,500	410	400	420	200,000	2,250	—	—
5,000	580	570	590	250,000	1,900	—	—
10,000	820	790	820	300,000	1,030	—	—
25,000	1,260	1,200	1,210	316,300	—	—	—

[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 derive these estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1980 Housing Inventory and for Estimated Percentages of 1976-1980 Lost Housing Units for the Oklahoma City, Okla., 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]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	41.8	41.8	41.8	41.8	41.8	42.4	50,000	0.14	0.4	0.8	1.1	1.6	1.9
200	26.4	26.4	26.4	26.4	26.4	30.0	75,000	0.10	0.3	0.7	0.9	1.3	1.5
500	12.6	12.6	12.6	12.6	16.4	19.0	100,000	0.07	0.3	0.6	0.8	1.2	1.3
700	9.3	9.3	9.3	9.6	13.9	16.0	150,000	0.05	0.2	0.5	0.7	0.9	1.1
1,000	6.7	6.7	6.7	8.0	11.6	13.4	200,000	0.04	0.2	0.4	0.6	0.8	0.9
2,500	2.8	2.8	3.7	5.1	7.3	8.5	250,000	0.03	0.2	0.4	0.5	0.7	0.8
5,000	1.4	1.4	2.6	3.6	5.2	6.0	300,000	0.02	0.2	0.3	0.5	0.7	0.8
10,000	0.7	0.8	1.8	2.5	3.7	4.2	400,000	0.02	0.13	0.3	0.4	0.6	0.7
25,000	0.3	0.5	1.2	1.6	2.3	2.7							

[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 1.0 for the total SMSA, 1.1 for the central city, and 1.0 for the balance (not in central city).

Providence-Pawtucket-Warwick 1980

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1980 survey. App-42
Selection of the 1976 AHS-SMSA sample App-43
1976-1980 additions to the housing inventory App-44
Sample selection for the 1976 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
1980 housing inventory App-45
1976-1980 lost housing units . . . App-46
1976 estimation procedure App-46
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-47
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-170-80) are based on data collected from the 1980 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 the Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill., SMSA's, the data were collected during the 11-month period from April 1980 through February 1981 with one-eleventh of the sample housing units being visited each month. In the remaining SMSA's, the data were collected during the 12-month period from April 1980 through March 1981 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 1980 were interviewed previously in either 1976 or 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1980 group of SMSA's, the largest SMSA from 3 of the 4 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 1980 group are: Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill.

The remaining SMSA's in the 1980 group are: Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Birmingham, Ala.; Grand Rapids, Mich.; Indianapolis, Ind.; Louisville, Ky.-Ind.; Memphis, Tenn.-Ark.; Oklahoma City, Okla.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; and Salt Lake City, Utah.

In this SMSA, 5,067 housing units were eligible for interview. Of these sample housing units, 142 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 5,067 housing units eligible for interview, 341 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 1980 survey—The sample housing units designated to be interviewed in the 1980 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 1980 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 and 1980 Coverage Improvement Programs. (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—The sample for the SMSA's which, in 1970, 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 1970, the following three SMSA's were 100-percent permit-issuing: Los Angeles-Long Beach, Calif.; New York, N.Y.; and Sacramento, Calif. The remaining 12 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.

1976-1980 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 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with the Oklahoma City, Okla. SMSA receiving some updating as a part of the 1980 AHS. For each of the five SMSA's interviewed previously in 1977, the Coverage Improvement Program was conducted as a part of the 1977 AHS with the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; and Salt Lake City, Utah, SMSA's receiving some updating and refining as a part of the 1980 AHS. The following discussion applies to both the prior year (1976 or 1977) and 1980 coverage improvement procedures. For the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Oklahoma City, Okla.; and Salt Lake City, Utah, SMSA's estimates of housing units added by a specific procedure reflect units added in the prior year as well as any additions that resulted from the updating and refining in 1980.

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 1980 SMSA's. Two different procedures were used. For the first procedure, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. For the SMSA's previously interviewed in 1976, these new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units). For the SMSA's previously interviewed in 1977, sample units selected from one- and two-unit structures were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample, while units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage sample selection was a sample of permit offices, and the second stage was a sample of the 1969 permits within each of the selected permit offices. In the Grand Rapids, Mich., and New York, N.Y., SMSA's, 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. This procedure was employed for the 10 SMSA's previously interviewed in 1976.

For the Albany-Schenectady-Troy, N.Y.; Memphis, Tenn.-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's previously interviewed in 1977, the above procedures were performed for the first and second stages. For the third stage, structures of size three or more were divided into clusters of an expected size of two units and a sample of clusters was selected.

For the Los Angeles-Long Beach, Calif., SMSA, 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 units were then sampled at one-third the rate of regular AHS units. Since permits were not available for all sampled offices, the procedure was also used in parts of the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 552 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. In the Oklahoma City, Okla., SMSA, a sample of tracts was selected but not canvassed during the 1976 Coverage Improvement Program. This procedure was completed during the 1980 AHS and added 620 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 AHS sample housing units was selected from the 1970 sample universe at a rate of 1 in 24 for the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Memphis, Tenn-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's and at a rate of 1 in 22 for the remaining 10 SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing unit 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. The Albany-Schenectady-Troy, N.Y., SMSA, had these sampling procedures implemented in 1976, excluding identification of the housing units missed in the 1970 census. This procedure was completed as a part of the 1980 AHS. The Los Angeles-Long Beach, Calif., and Salt Lake City, Utah, SMSA's had this sampling procedure implemented for the first time in 1980. This procedure added an estimated 10,363 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 3,457 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 1980 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 1980 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1976 (i.e., 1976-1980 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1980 housing inventory—The AHS estimates of characteristics of the 1980 housing inventory were produced using a 2-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York,N.Y.; and Oklahoma City, Okla., 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 Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Grand Rapids, Mich.; Indianapolis, Ind.; Los Angeles-Long Beach, Calif.; Louisville, Ky.-Ind.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; St. Louis, Mo.-Ill.; and Salt Lake City, Utah, 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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., SMSA's. In the Birmingham, Ala.; Memphis, Tenn.-Ark.; and Oklahoma City, Okla., 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 New York, N.Y., SMSA, the sample selected correctly apportioned the new construction housing units between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 1980 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:

$$\frac{\text{Independent estimate of the October 1980 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 procedure, 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.

1976-1980 lost housing units—The 1976-1980 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 1976-1980 lost housing units existed, by definition, in the 1976 housing inventory, there was a 1976 housing inventory weight associated with each 1976-1980 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1976-1980 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 1980 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 of which 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 1980 survey, the number of missed housing units may be considerably less for 1980.

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

Rounding errors—For errors associated with processing, 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 1976 housing inventory can be found in the AHS Series H-170 reports for 1976.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1980 housing inventory as well as estimates of characteristics of the 1976-1980 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 11,230 for the total SMSA, 6,210 for the central city of the SMSA, and 9,290 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 1980 housing inventory as well as estimated percentages of the 1976-1980 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 1980 there were 198,400 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 2,550. 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
150,000	2,540
198,400	x
200,000	2,550

The entry for "x" is determined as follows by vertically interpolating between 2,540 and 2,550.

$$198{,}400-150{,}000 = 48{,}400$$
$$200{,}000-150{,}000 = 50{,}000$$

$$2{,}540 + \frac{48{,}400}{50{,}000}(2{,}550-2{,}540) = 2{,}550$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 195,850 to 200,950 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1980 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 194,320 to 202,480 housing units with 90 percent confidence; and that the average estimate lies within the interval from 193,300 to 203,500 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 198,400 owner-occupied housing units, 57,100, or 28.8 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 28.8 percent is approximately 0.9 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	28.8	50
150,000.	1.0	a	1.2
198,400.		p	
200,000.	0.9	b	1.0

1. The entry for cell "a" is determined by horizontal interpolation between 1.0 and 1.2.

$$28.8-25.0 = 3.8$$
$$50.0-25.0 = 25.0$$

$$1.0 + \frac{3.8}{25.0}(1.2-1.0) = 1.0$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.9 and 1.0.

$$28.8-25.0 = 3.8$$
$$50.0-25.0 = 25.0$$

$$0.9 + \frac{3.8}{25.0}(1.0-0.9) = 0.9$$

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

$$198{,}400-150{,}000 = 48{,}400$$
$$200{,}000-150{,}000 = 50{,}000$$

$$1.0 + \frac{48{,}400}{50{,}000}(0.9-1.0) = 0.9$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 27.9 to 29.7 percent; the 90-percent confidence interval is from 27.4 to 30.2 percent; and the 95-percent confidence interval is from 27.0 to 30.6 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; if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1976 and 1980 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1980 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1980 there were 95,500 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 38,400. Table I shows the standard error of 57,100 is approximately 1,880, and the standard error of 95,500 is approximately 2,270. Therefore, the standard error of the estimated difference of 38,400 is about 2,950:

$$2{,}950 = \sqrt{(1{,}880)^2 + (2{,}270)^2}$$

Consequently, the 68-percent confidence interval for the 38,400 difference is from 35,450 to 41,350 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 33,680 to 43,120 housing units, and the 95-percent confidence interval is from 32,500 to 44,300 housing units. Thus, we can conclude with 95 percent confidence that the number of 1980 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 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 198,400 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 198,400 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 three persons (for purposes of calculating the median, the category of three persons is considered to be from 2.5 to 3.5 persons) contains the 48.0 percent derived in step 2. About 83,000 housing units or 41.8 percent fall below this interval, and 35,900 housing units or 18.1 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(48.0-41.8)}{18.1} = 2.8$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.0 percent derived in step 2. About 83,000 housing units or 41.8 percent fall below this interval, and 35,900 housing units or 18.1 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(52.0-41.8)}{18.1} = 3.1$$

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 1980 Housing Inventory and for Estimated Number of 1976-1980 Lost Units for the Providence-Pawtucket-Warwick, R.I.-Mass., 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]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	70	70	80	50,000	1,790	1,480	1,780
100	90	80	90	75,000	2,100	1,500	2,020
200	120	120	130	100,000	2,310	1,280	2,130
500	190	190	200	130,100	2,470	–	2,110
700	230	220	240	150,000	2,540	–	2,020
1,000	270	270	280	200,000	2,550	–	1,340
2,500	430	420	450	224,900	2,470	–	–
5,000	610	590	630	250,000	2,340	–	–
10,000	850	810	880	300,000	1,860	–	–
25,000	1,310	1,200	1,350	355,000	–	–	–

[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, 1.1 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 1980 Housing Inventory and for Estimated Percentages of 1976-1980 Lost Units for the Providence-Pawtucket-Warwick, R.I.-Mass., 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]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	44.9	44.9	44.9	44.9	44.9	45.1	50,000	0.2	0.4	0.9	1.2	1.7	2.0
200	28.9	28.9	28.9	28.9	28.9	31.9	75,000	0.11	0.3	0.7	1.0	1.4	1.6
500	14.0	14.0	14.0	14.0	17.5	20.2	100,000	0.08	0.3	0.6	0.9	1.2	1.4
700	10.4	10.4	10.4	10.4	14.8	17.1	150,000	0.05	0.2	0.5	0.7	1.0	1.2
1,000	7.5	7.5	7.5	8.6	12.4	14.3	200,000	0.04	0.2	0.4	0.6	0.9	1.0
2,500	3.2	3.2	3.9	5.4	7.8	9.0	250,000	0.03	0.2	0.4	0.5	0.8	0.9
5,000	1.6	1.6	2.8	3.8	5.5	6.4	300,000	0.03	0.2	0.4	0.5	0.7	0.8
10,000	0.8	0.9	2.0	2.7	3.9	4.5	400,000	0.02	0.14	0.3	0.4	0.6	0.7
25,000	0.3	0.6	1.2	1.7	2.5	2.9							

[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, 1.1 for the central city, and 1.2 for the balance (not in central city).

*U.S. GOVERNMENT PRINTING OFFICE: 1984-0-421-997/5303

Sacramento 1980

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1980 survey. App-42
Selection of the 1976 AHS-SMSA sample App-43
1976-1980 additions to the housing inventory App-44
Sample selection for the 1976 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
1980 housing inventory App-45
1976-1980 lost housing units . . . App-46
1976 estimation procedure App-46
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-47
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-170-80) are based on data collected from the 1980 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 the Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill., SMSA's, the data were collected during the 11-month period from April 1980 through February 1981 with one-eleventh of the sample housing units being visited each month. In the remaining SMSA's, the data were collected during the 12-month period from April 1980 through March 1981 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 1980 were interviewed previously in either 1976 or 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1980 group of SMSA's, the largest SMSA from 3 of the 4 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 1980 group are: Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill.

The remaining SMSA's in the 1980 group are: Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Birmingham, Ala.; Grand Rapids, Mich.; Indianapolis, Ind.; Louisville, Ky.-Ind.; Memphis, Tenn.-Ark.; Oklahoma City, Okla.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; and Salt Lake City, Utah.

In this SMSA, 5,868 housing units were eligible for interview. Of these sample housing units, 228 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 5,868 housing units eligible for interview, 244 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 1980 survey—The sample housing units designated to be interviewed in the 1980 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 1980 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 and 1980 Coverage Improvement Programs. (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—The sample for the SMSA's which, in 1970, 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 1970, the following three SMSA's were 100-percent permit-issuing: Los Angeles-Long Beach, Calif.; New York, N.Y.; and Sacramento, Calif. The remaining 12 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.

1976-1980 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 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with the Oklahoma City, Okla. SMSA receiving some updating as a part of the 1980 AHS. For each of the five SMSA's interviewed previously in 1977, the Coverage Improvement Program was conducted as a part of the 1977 AHS with the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; and Salt Lake City, Utah, SMSA's receiving some updating and refining as a part of the 1980 AHS. The following discussion applies to both the prior year (1976 or 1977) and 1980 coverage improvement procedures. For the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Oklahoma City, Okla.; and Salt Lake City, Utah, SMSA's estimates of housing units added by a specific procedure reflect units added in the prior year as well as any additions that resulted from the updating and refining in 1980.

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 1980 SMSA's. Two different procedures were used. For the first procedure, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. For the SMSA's previously interviewed in 1976, these new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units). For the SMSA's previously interviewed in 1977, sample units selected from one- and two-unit structures were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample, while units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage sample selection was a sample of permit offices, and the second stage was a sample of the 1969 permits within each of the selected permit offices. In the Grand Rapids, Mich., and New York, N.Y., SMSA's, 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. This procedure was employed for the 10 SMSA's previously interviewed in 1976.

For the Albany-Schenectady-Troy, N.Y.; Memphis, Tenn.-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's previously interviewed in 1977, the above procedures were performed for the first and second stages. For the third stage, structures of size three or more were divided into clusters of an expected size of two units and a sample of clusters was selected.

For the Los Angeles-Long Beach, Calif., SMSA, 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 units were then sampled at one-third the rate of regular AHS units. Since permits were not available for all sampled offices, the procedure was also used in parts of the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 6,719 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. In the Oklahoma City, Okla., SMSA, a sample of tracts was selected but not canvassed during the 1976 Coverage Improvement Program. This procedure was completed during the 1980 AHS and added 2,317 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 AHS sample housing units was selected from the 1970 sample universe at a rate of 1 in 24 for the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Memphis, Tenn-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's and at a rate of 1 in 22 for the remaining 10 SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing unit 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. The Albany-Schenectady-Troy, N.Y., SMSA, had these sampling procedures implemented in 1976, excluding identification of the housing units missed in the 1970 census. This procedure was completed as a part of the 1980 AHS. The Los Angeles-Long Beach, Calif., and Salt Lake City, Utah, SMSA's had this sampling procedure implemented for the first time in 1980. This procedure added an estimated 5,041 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 211 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 1980 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 1980 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1976 (i.e., 1976-1980 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1980 housing inventory—The AHS estimates of characteristics of the 1980 housing inventory were produced using a 2-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Grand Rapids, Mich.; Indianapolis, Ind.; Los Angeles-Long Beach, Calif.; Louisville, Ky.-Ind.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; St. Louis, Mo.-Ill.; and Salt Lake City, Utah, 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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., SMSA's. In the Birmingham, Ala.; Memphis, Tenn.-Ark.; and Oklahoma City, Okla., 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 New York, N.Y., SMSA, the sample selected correctly apportioned the new construction housing units between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 1980 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:

$$\frac{\text{Independent estimate of the October 1980 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 procedure, 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.

1976-1980 lost housing units—The 1976-1980 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 1976-1980 lost housing units existed, by definition, in the 1976 housing inventory, there was a 1976 housing inventory weight associated with each 1976-1980 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1976-1980 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 1980 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 of which 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 1980 survey, the number of missed housing units may be considerably less for 1980.

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

Rounding errors—For errors associated with processing, 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 1976 housing inventory can be found in the AHS Series H-170 reports for 1976.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1980 housing inventory as well as estimates of characteristics of the 1976-1980 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 15,230 for the total SMSA, 8,520 for the central city of the SMSA, and 12,530 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 1980 housing inventory as well as estimated percentages of the 1976-1980 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 1980 there were 235,100 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 3,060. 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
200,000	3,080
235,100	x
250,000	3,050

The entry for "x" is determined as follows by vertically interpolating between 3,080 and 3,050.

$$235{,}100-200{,}000 = 35{,}100$$
$$250{,}000-200{,}000 = 50{,}000$$

$$3{,}080 + \frac{35{,}100}{50{,}000}(3{,}050-3{,}080) = 3{,}060$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 232,040 to 238,160 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1980 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 230,200 to 240,000 housing units with 90 percent confidence; and that the average estimate lies within the interval from 228,980 to 241,220 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 235,100 owner-occupied housing units, 54,700, or 23.3 percent, had two bedrooms. Interpolating using table II of this appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 23.3 percent is approximately 0.9 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	23.3	25 or 75
200,000.	0.7	a	1.0
235,100.		p	
250,000.	0.6	b	0.9

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

$$23.3-10.0 = 13.3$$
$$25.0-10.0 = 15.0$$

$$0.7 + \frac{13.3}{15.0}(1.0-0.7) = 1.0$$

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

$$23.3-10.0 = 13.3$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{13.3}{15.0}(0.9-0.6) = 0.9$$

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

$$235,100-200,000 = 35,100$$
$$250,000-200,000 = 50,000$$

$$1.0 + \frac{35,100}{50,000}(0.9-1.0) = 0.9$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 22.4 to 24.2 percent; the 90-percent confidence interval is from 21.9 to 24.7 percent; and the 95-percent confidence interval is from 21.5 to 25.1 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; if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1976 and 1980 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1980 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1980 there were 126,100 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 71,400. Table I shows the standard error of 54,700 is approximately 2,050, and the standard error of 126,100 is approximately 2,810. Therefore, the standard error of the estimated difference of 71,400 is about 3,480:

$$3,480 = \sqrt{(2,050)^2 + (2,810)^2}$$

Consequently, the 68-percent confidence interval for the 71,400 difference is from 67,920 to 74,880 housing units. Therefore, a conclusion that the average estimate derived from all possible samples, of this difference, lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 65,830 to 76,970 housing units, and the 95-percent confidence interval is from 64,440 to 78,360 housing units. Thus, we can conclude with 95 percent confidence that the number of 1980 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 substract 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 necessary to know 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 which interval of the distribution the upper percentage limit falls. These two distribution intervals could be different.

For about 68 out 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 2.4. The base of the distribution from which this median was determined is 235,100 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 235,100 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 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 48.0 percent derived in step 2. About 35,900 housing units or 15.3 percent fall below this interval, and 85,500 housing units or 36.4 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$1.5 + (2.5-1.5)\,\frac{(48.0-15.3)}{36.4} = 2.4$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.0 percent derived in step 2. About 121,400 housing units or 51.6 percent fall below this interval, and 42,600 housing units or 18.1 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(52.0-51.6)}{18.1} = 2.5$$

Thus, the 95-percent confidence interval ranges from 2.4 to 2.5 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1980 Housing Inventory and for Estimated Number of 1976-1980 Lost Units for the Sacramento, Calif., 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]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	90	100	90	75,000	2,350	1,720	2,190
100	90	100	90	100,000	2,610	1,400	2,390
200	130	140	130	124,600	2,800	–	2,500
500	210	220	210	150,000	2,950	–	2,540
700	250	260	240	200,000	3,080	–	2,420
1,000	300	310	290	250,000	3,050	–	1,950
2,500	470	490	460	300,000	2,830	–	590
5,000	660	690	650	304,200	2,800	–	–
10,000	930	950	910	400,000	1,550	–	–
25,000	1,450	1,400	1,400	428,800	–	–	–
50,000	1,980	1,720	1,890				

[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, 1.1 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 Percentage of Housing Units in the 1980 Housing Inventory and for Estimated Percentages of 1976-1980 Lost Housing Units for the Sacramento, Calif., 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]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	49.6	49.6	49.6	49.6	49.6	49.6	50,000	0.2	0.4	1.0	1.3	1.9	2.2
200	33.0	33.0	33.0	33.0	33.0	35.1	75,000	0.13	0.4	0.8	1.1	1.6	1.8
500	16.5	16.5	16.5	16.5	19.2	22.2	100,000	0.10	0.3	0.7	0.9	1.4	1.6
700	12.3	12.3	12.3	12.3	16.3	18.8	150,000	0.07	0.3	0.6	0.8	1.1	1.3
1,000	9.0	9.0	9.0	9.4	13.6	15.7	200,000	0.05	0.2	0.5	0.7	1.0	1.1
2,500	3.8	3.8	4.3	6.0	8.6	9.9	250,000	0.04	0.2	0.4	0.6	0.9	1.0
5,000	1.9	1.9	3.1	4.2	6.1	7.0	300,000	0.03	0.2	0.4	0.5	0.8	0.9
10,000	1.0	1.0	2.2	3.0	4.3	5.0	400,000	0.02	0.2	0.3	0.5	0.7	0.8
25,000	0.4	0.6	1.4	1.9	2.7	3.1	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, 1.1 for the central city, and 1.2 for the balance (not in central city).

Saginaw 1980

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1980 survey. App-42
Selection of the 1977 AHS-SMSA sample App-43
1977-1980 additions to the housing inventory App-44
Sample selection for the 1977 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
1980 housing inventory App-45
1977-1980 lost housing units . . . App-46
1977 estimation procedure App-46
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-47
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-170-80) are based on data collected from the 1980 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 the Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill., SMSA's, the data were collected during the 11-month period from April 1980 through February 1981 with one-eleventh of the sample housing units being visited each month. In the remaining SMSA's, the data were collected during the 12-month period from April 1980 through March 1981 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 1980 were interviewed previously in either 1976 or 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1980 group of SMSA's, the largest SMSA from 3 of the 4 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 1980 group are: Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill.

The remaining SMSA's in the 1980 group are: Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Birmingham, Ala.; Grand Rapids, Mich.; Indianapolis, Ind.; Louisville, Ky.-Ind.; Memphis, Tenn.-Ark.; Oklahoma City, Okla.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; and Salt Lake City, Utah.

In this SMSA, 5,224 housing units were eligible for interview. Of these sample housing units, 246 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 5,224 housing units eligible for interview, 257 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 1980 survey—The sample housing units designated to be interviewed in the 1980 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 1977 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 1977 survey. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1980 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 1977 survey. (This sample represented the housing units built in permit-issuing areas, since the 1977 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1977 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1977 survey.)
5. All sample housing units that were selected as part of the 1977 and 1980 Coverage Improvement Programs. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1977 AHS-SMSA sample—The sample for the SMSA's which, in 1970, 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 1970, the following three SMSA's were 100-percent permit-issuing: Los Angeles-Long Beach, Calif.; New York, N.Y.; and Sacramento, Calif. The remaining 12 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.

1977-1980 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1977 survey, was selected to represent housing units built in permit-issuing areas since the 1977 survey. Sampling procedures were identical to those used in selecting the 1970-1977 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1977, to identify any housing units missed in the 1977 survey or any housing units added since the 1977 survey.

Sample selection for the 1977 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 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with the Oklahoma City, Okla. SMSA receiving some updating as a part of the 1980 AHS. For each of the five SMSA's interviewed previously in 1977, the Coverage Improvement Program was conducted as a part of the 1977 AHS with the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; and Salt Lake City, Utah, SMSA's receiving some updating and refining as a part of the 1980 AHS. The following discussion applies to both the prior year (1976 or 1977) and 1980 coverage improvement procedures. For the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Oklahoma City, Okla.; and Salt Lake City, Utah, SMSA's estimates of housing units added by a specific procedure reflect units added in the prior year as well as any additions that resulted from the updating and refining in 1980.

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 1980 SMSA's. Two different procedures were used. For the first procedure, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. For the SMSA's previously interviewed in 1976, these new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units). For the SMSA's previously interviewed in 1977, sample units selected from one- and two-unit structures were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample, while units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage sample selection was a sample of permit offices, and the second stage was a sample of the 1969 permits within each of the selected permit offices. In the Grand Rapids, Mich., and New York, N.Y., SMSA's, 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. This procedure was employed for the 10 SMSA's previously interviewed in 1976.

For the Albany-Schenectady-Troy, N.Y.; Memphis, Tenn.-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's previously interviewed in 1977, the above procedures were performed for the first and second stages. For the third stage, structures of size three or more were divided into clusters of an expected size of two units and a sample of clusters was selected.

For the Los Angeles-Long Beach, Calif., SMSA, 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 units were then sampled at one-third the rate of regular AHS units. Since permits were not available for all sampled offices, the procedure was also used in parts of the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 218 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. In the Oklahoma City, Okla., SMSA, a sample of tracts was selected but not canvassed during the 1976 Coverage Improvement Program. This procedure was completed during the 1980 AHS and added 749 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 AHS sample housing units was selected from the 1970 sample universe at a rate of 1 in 24 for the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Memphis, Tenn-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's and at a rate of 1 in 22 for the remaining 10 SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing unit 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. The Albany-Schenectady-Troy, N.Y., SMSA, had these sampling procedures implemented in 1976, excluding identification of the housing units missed in the 1970 census. This procedure was completed as a part of the 1980 AHS. The Los Angeles-Long Beach, Calif., and Salt Lake City, Utah, SMSA's had this sampling procedure implemented for the first time in 1980. This procedure added an estimated 2,039 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 42 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 1980 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 1980 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1977 (i.e., 1977-1980 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1980 housing inventory—The AHS estimates of characteristics of the 1980 housing inventory were produced using a 2-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York,N.Y.; and Oklahoma City, Okla., 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 Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Grand Rapids, Mich.; Indianapolis, Ind.; Los Angeles-Long Beach, Calif.; Louisville, Ky.-Ind.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; St. Louis, Mo.-Ill.; and Salt Lake City, Utah, 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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., SMSA's. In the Birmingham, Ala.; Memphis, Tenn.-Ark.; and Oklahoma City, Okla., 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 New York, N.Y., SMSA, the sample selected correctly apportioned the new construction housing units between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 1980 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:

$$\frac{\text{Independent estimate of the October 1980 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 procedure, 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.

1977-1980 lost housing units—The 1977-1980 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 1977 housing inventory, as was described in the 1977 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1977-1980 lost housing units existed, by definition, in the 1977 housing inventory, there was a 1977 housing inventory weight associated with each 1977-1980 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1977-1980 lost housing units.

1977 estimation procedure—This report presents data on the housing characteristics of the 1977 housing inventory from the 1977 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 1977.

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 1977 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 1980 AHS-SMSA sample reinterview program were not available at the time this report was being prepared. However, a study was conducted for the 1977 AHS-SMSA sample. The results of which are presented in the Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1977."

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 1977 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 1980 survey, the number of missed housing units may be considerably less for 1980.

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

Rounding errors—For errors associated with processing, 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 1977 housing inventory can be found in the AHS Series H-170 reports for 1977.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1980 housing inventory as well as estimates of characteristics of the 1977-1980 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 3,680 for the total SMSA, 1,320 for the central city of the SMSA, and 3,310 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 1980 housing inventory as well as estimated percentages of the 1977-1980 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 1980 there were 57,400 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 460. 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
51,500	530
57,400	x
75,000	270

The entry for "x" is determined as follows by vertically interpolating between 530 and 270.

$$57{,}400-51{,}500 = 5{,}900$$
$$75{,}000-51{,}500 = 23{,}500$$

$$530+\frac{5{,}900}{23{,}500}(270-530) = 460$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 51,940 to 57,860 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1980 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 56,660 to 58,140 housing units with 90 percent confidence; and that the average estimate lies within the interval from 56,480 to 58,320 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 57,400 owner-occupied housing units, 12,900, or 22.5 percent, had two bedrooms. Interpolation using table II of this appendix (i.e., interpolation on both the base and present) shows that the standard error of the 22.5 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	22.5	25 or 75
50,000	0.5	a	0.8
57,400		p	
75,000	0.4	b	0.6

1. The entry for cell "a" is determined by horizontal interpolation between 0.5 and 0.8.

$$22.5-10.0 = 12.5$$
$$25.0-10.0 = 15.0$$

$$0.5 + \frac{12.5}{15.0}(0.8-0.5) = 0.8$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.4 and 0.6.

$$22.5-10.0 = 12.5$$
$$25.0-10.0 = 15.0$$

$$0.4 + \frac{12.5}{15.0}(0.6-0.4) = 0.6$$

3. The entry for "p" is then determined by vertical interpolation between 0.8 and 0.6.

$$57{,}400-50{,}000 = 7{,}400$$
$$75{,}000-50{,}000 = 25{,}000$$

$$0.8 + \frac{7{,}400}{25{,}000}(0.6-0.8) = 0.7$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 21.8 to 23.2 percent; the 90-percent confidence interval is from 21.4 to 23.6 percent; and the 95-percent confidence interval is from 21.1 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 characterisitcs in the same SMSA. If there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error; if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1977 and 1980 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1977 and 1980 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1980 there were 32,800 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 units with three bedrooms is 19,900. Table I shows the standard error of 12,900 is approximately 400, and the standard error of 32,800 is approximately 530. Therefore, the standard error of the estimated difference of 19,900 is about 660:

$$660 = \sqrt{(400)^2 + (530)^2}$$

Consequently, the 68-percent confidence interval for the 19,900 difference is from 19,240 to 20,560 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 18,840 to 20,960 housing units, and the 95-percent confidence interval is from 18,580 to 21,220 housing units. Thus, we can conclude with 95 percent confidence that the number of 1980 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 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 3.1. The base of the distribution from which this median was determined is 57,400 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 57,400 is approximately 0.8 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 48.4 and 51.6.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with three persons (for purposes of calculating the median, the category of three persons is considered to be from 2.5 to 3.5 persons) contains the 48.4 percent derived in step 2. About 22,300 housing units or 38.9 percent fall below this interval, and 10,400 housing units or 18.1 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(48.4-38.9)}{18.1} = 3.0$$

Similarly, the interval for owner-occupied housing units with three persons contains the 51.6 percent derived in step 2. About 22,300 housing units or 38.9 percent fall below this interval, and 10,400 housing units or 18.1 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(51.6-38.9)}{18.1} = 3.2$$

Thus, the 95-percent confidence interval ranges from 3.0 to 3.2 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1980 Housing Inventory and for Estimated Number of 1977-1980 Lost Units for the Saginaw, Mich., 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]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	20	10	20	10,000	370	300	360
100	40	40	40	25,000	510	210	450
200	60	50	60	28,700	530	–	450
500	90	80	90	50,000	540	–	150
700	100	100	110	51,500	530	–	–
1,000	120	110	130	75,000	270	–	–
2,500	190	180	200	80,200	–	–	–
5,000	270	240	270				

[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 1980 Housing Inventory and for Estimated Percentages of 1977-1980 Lost Units for the Saginaw, Mich., SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	13.8	13.8	13.8	13.8	17.3	20.0	5,000	0.3	0.6	1.2	1.7	2.4	2.8
200	7.4	7.4	7.4	8.5	12.2	14.1	10,000	0.2	0.4	0.9	1.2	1.7	2.0
500	3.1	3.1	3.9	5.4	7.7	8.9	25,000	0.06	0.3	0.6	0.8	1.1	1.3
700	2.2	2.2	3.3	4.5	6.5	7.6	50,000	0.03	0.2	0.4	0.5	0.8	0.9
1,000	1.6	1.6	2.8	3.8	5.5	6.3	75,000	0.02	0.15	0.3	0.4	0.6	0.7
2,500	0.6	0.8	1.7	2.4	3.5	4.0	100,000	0.02	0.13	0.3	0.4	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 total SMSA, 1.3 for the central city, and 1.2 for the balance (not in central city).

Salt Lake City 1980

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1980 survey. App-42
Selection of the 1977 AHS-SMSA sample App-43
1977-1980 additions to the housing inventory App-44
Sample selection for the 1977 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
1980 housing inventory App-45
1977-1980 lost housing units . . . App-46
1977 estimation procedure App-46
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-47
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-170-80) are based on data collected from the 1980 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 the Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill., SMSA's, the data were collected during the 11-month period from April 1980 through February 1981 with one-eleventh of the sample housing units being visited each month. In the remaining SMSA's, the data were collected during the 12-month period from April 1980 through March 1981 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 1980 were interviewed previously in either 1976 or 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1980 group of SMSA's, the largest SMSA from 3 of the 4 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 1980 group are: Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill.

The remaining SMSA's in the 1980 group are: Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Birmingham, Ala.; Grand Rapids, Mich.; Indianapolis, Ind.; Louisville, Ky.-Ind.; Memphis, Tenn.-Ark.; Oklahoma City, Okla.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; and Salt Lake City, Utah.

In this SMSA, 6,386 housing units were eligible for interview. Of these sample housing units, 265 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 6,386 housing units eligible for interview, 319 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 1980 survey—The sample housing units designated to be interviewed in the 1980 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 1977 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 1977 survey. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1980 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 1977 survey. (This sample represented the housing units built in permit-issuing areas, since the 1977 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1977 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1977 survey.)
5. All sample housing units that were selected as part of the 1977 and 1980 Coverage Improvement Programs. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1977 AHS-SMSA sample—The sample for the SMSA's which, in 1970, 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 1970, the following three SMSA's were 100-percent permit-issuing: Los Angeles-Long Beach, Calif.; New York, N.Y.; and Sacramento, Calif. The remaining 12 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.

1977-1980 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1977 survey, was selected to represent housing units built in permit-issuing areas since the 1977 survey. Sampling procedures were identical to those used in selecting the 1970-1977 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1977, to identify any housing units missed in the 1977 survey or any housing units added since the 1977 survey.

Sample selection for the 1977 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 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with the Oklahoma City, Okla. SMSA receiving some updating as a part of the 1980 AHS. For each of the five SMSA's interviewed previously in 1977, the Coverage Improvement Program was conducted as a part of the 1977 AHS with the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; and Salt Lake City, Utah, SMSA's receiving some updating and refining as a part of the 1980 AHS. The following discussion applies to both the prior year (1976 or 1977) and 1980 coverage improvement procedures. For the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Oklahoma City, Okla.; and Salt Lake City, Utah, SMSA's estimates of housing units added by a specific procedure reflect units added in the prior year as well as any additions that resulted from the updating and refining in 1980.

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 1980 SMSA's. Two different procedures were used. For the first procedure, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. For the SMSA's previously interviewed in 1976, these new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units). For the SMSA's previously interviewed in 1977, sample units selected from one- and two-unit structures were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample, while units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage sample selection was a sample of permit offices, and the second stage was a sample of the 1969 permits within each of the selected permit offices. In the Grand Rapids, Mich., and New York, N.Y., SMSA's, 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. This procedure was employed for the 10 SMSA's previously interviewed in 1976.

For the Albany-Schenectady-Troy, N.Y.; Memphis, Tenn.-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's previously interviewed in 1977, the above procedures were performed for the first and second stages. For the third stage, structures of size three or more were divided into clusters of an expected size of two units and a sample of clusters was selected.

For the Los Angeles-Long Beach, Calif., SMSA, 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 units were then sampled at one-third the rate of regular AHS units. Since permits were not available for all sampled offices, the procedure was also used in parts of the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 747 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. In the Oklahoma City, Okla., SMSA, a sample of tracts was selected but not canvassed during the 1976 Coverage Improvement Program. This procedure was completed during the 1980 AHS and added 1,774 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 AHS sample housing units was selected from the 1970 sample universe at a rate of 1 in 24 for the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Memphis, Tenn-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's and at a rate of 1 in 22 for the remaining 10 SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing unit 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. The Albany-Schenectady-Troy, N.Y., SMSA, had these sampling procedures implemented in 1976, excluding identification of the housing units missed in the 1970 census. This procedure was completed as a part of the 1980 AHS. The Los Angeles-Long Beach, Calif., and Salt Lake City, Utah, SMSA's had this sampling procedure implemented for the first time in 1980. This procedure added an estimated 11,743 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 1980 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 1980 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1977 (i.e., 1977-1980 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1980 housing inventory—The AHS estimates of characteristics of the 1980 housing inventory were produced using a 2-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Grand Rapids, Mich.; Indianapolis, Ind.; Los Angeles-Long Beach, Calif.; Louisville, Ky.-Ind.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; St. Louis, Mo.-Ill.; and Salt Lake City, Utah, 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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., SMSA's. In the Birmingham, Ala.; Memphis, Tenn.-Ark.; and Oklahoma City, Okla., 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 New York, N.Y., SMSA, the sample selected correctly apportioned the new construction housing units between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 1980 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:

$$\frac{\text{Independent estimate of the October 1980 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 procedure, 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.

1977-1980 lost housing units—The 1977-1980 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 1977 housing inventory, as was described in the 1977 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1977-1980 lost housing units existed, by definition, in the 1977 housing inventory, there was a 1977 housing inventory weight associated with each 1977-1980 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1977-1980 lost housing units.

1977 estimation procedure—This report presents data on the housing characteristics of the 1977 housing inventory from the 1977 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 1977.

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 1977 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 1980 AHS-SMSA sample reinterview program were not available at the time this report was being prepared. However, a study was conducted for the 1977 AHS-SMSA sample. The results of which are presented in the Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1977."

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 1977 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 1980 survey, the number of missed housing units may be considerably less for 1980.

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

Rounding errors—For errors associated with processing, 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 1977 housing inventory can be found in the AHS Series H-170 reports for 1977.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1980 housing inventory as well as estimates of characteristics of the 1977-1980 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 9,560 for the total SMSA, 3,890 for the central city of the SMSA, and 8,620 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 1980 housing inventory as well as estimated percentages of the 1977-1980 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 1980 there were 167,800 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 1,690. 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
150,000	1,770
167,800	x
187,600	1,610

The entry for "x" is determined as follows by vertically interpolating between 1,770 and 1,610.

$$167{,}800-150{,}000 = 17{,}800$$
$$187{,}600-150{,}000 = 37{,}600$$

$$1{,}770+\frac{17{,}800}{37{,}600}(1{,}610-1{,}770) = 1{,}690$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 166,110 to 169,490 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1980 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 165,100 to 170,500 housing units with 90 percent confidence; and that the average estimate lies within the interval from 164,420 to 171,180 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 167,800 owner-occupied housing units, 39,900, or 23.8 percent, had two bedrooms. Interpolation using table II of this appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 23.8 percent is approximately 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	23.8	25 or 75
150,000.	0.6	a	0.8
167,800.		p	
200,000.	0.5	b	0.7

1. The entry for cell "a" is determined by horizontal interpolation between 0.6 and 0.8.

$$23.8-10.0 = 13.8$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{13.8}{15.0}(0.8-0.6) = 0.8$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.5 and 0.7.

$$23.8-10.0 = 13.8$$
$$25.0-10.0 = 15.0$$

$$0.5 + \frac{13.8}{15.0}(0.7-0.5) = 0.7$$

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

$$167{,}800-150{,}000 = 17{,}800$$
$$200{,}000-150{,}000 = 50{,}000$$

$$0.8 + \frac{17{,}800}{50{,}000}(0.7-0.8) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 23.0 to 24.6 percent; the 90-percent confidence interval is from 22.5 to 25.1 percent; and the 95-percent confidence interval is from 22.2 to 25.4 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; if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1977 and 1980 AHS-SMSA samples a positive correlation should be expected when making compairsons between 1977 and 1980 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1980 there were 69,400 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 29,500. Table I shows the standard error of 39,900 is approximately 1,260, and the standard error of 69,400 is approximately 1,580. Therefore, the standard error of the estimated difference of 29,500 is about 2,020:

$$2{,}020 = \sqrt{(1{,}260)^2 + (1{,}580)^2}$$

Consequently, the 68-percent confidence interval for the 29,500 difference is from 27,480 to 31,520 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 26,270 to 32,730 housing units, and the 95-percent confidence interval is from 25,460 to 33,540 housing units. Thus, we can conclude with 95 percent confidence that the number of 1980 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 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 of owner-occupied housing units is 3.2. The base of the distribution from which this median was determined is 167,800 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 167,800 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 three persons (for purposes of calculating the median, the category of three persons is considered to be from 2.5 to 3.5 persons) contains the 48.2 percent derived in step 2. About 63,300 housing units or 37.7 percent fall below this interval, and 27,500 housing units or 16.4 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(48.2-37.7)}{16.4} = 3.1$$

Similarly, the interval for owner-occupied housing units with three persons contains the 51.8 percent derived in step 2. About 63,300 housing units or 37.7 percent fall below this interval, and 27,500 housing units or 16.4 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(51.8-37.7)}{16.4} = 3.4$$

Thus, the 95-percent confidence interval ranges from 3.1 to 3.4 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1980 Housing Inventory and for Estimated Number of 1977-1980 Lost Units for the Salt Lake City, Utah, 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]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	50	40	50	50,000	1,410	820	1,390
100	70	60	70	73,200	1,610	–	1,530
200	100	90	100	75,000	1,620	–	1,540
500	160	140	160	100,000	1,740	–	1,570
700	190	170	190	150,000	1,770	–	1,260
1,000	220	200	230	187,600	1,610	–	–
2,500	350	320	360	200,000	1,510	–	–
5,000	490	440	510	250,000	710	–	–
10,000	690	600	710	260,800	–	–	–
25,000	1,050	830	1,070				

[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 1.0 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 derive these estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1980 Housing Inventory and for Estimated Percentages of 1977-1980 Lost Units for the Salt Lake City, Utah, SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	34.5	34.5	34.5	34.5	34.5	36.3	25,000	0.2	0.5	1.0	1.4	2.0	2.3
200	20.9	20.9	20.9	20.9	22.2	25.7	50,000	0.11	0.3	0.7	1.0	1.4	1.6
500	9.5	9.5	9.5	9.7	14.1	16.2	75,000	0.07	0.3	0.6	0.8	1.1	1.3
700	7.0	7.0	7.0	8.2	11.9	13.7	100,000	0.05	0.2	0.5	0.7	1.0	1.1
1,000	5.0	5.0	5.0	6.9	9.9	11.5	150,000	0.04	0.2	0.4	0.6	0.8	0.9
2,500	2.1	2.1	3.2	4.4	6.3	7.3	200,000	0.03	0.2	0.4	0.5	0.7	0.8
5,000	1.0	1.0	2.2	3.1	4.4	5.1	250,000	0.02	0.14	0.3	0.4	0.6	0.7
10,000	0.5	0.7	1.6	2.2	3.1	3.6	300,000	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 for the total SMSA, 1.3 for the central city, and 1.0 for the balance (not in central city).

St. Louis 1980

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1980 survey. App-42
Selection of the 1976 AHS-SMSA sample App-43
1976-1980 additions to the housing inventory App-44
Sample selection for the 1976 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
1980 housing inventory App-45
1976-1980 lost housing units . . . App-46
1976 estimation procedure App-46
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-47
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-170-80) are based on data collected from the 1980 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 the Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill., SMSA's, the data were collected during the 11-month period from April 1980 through February 1981 with one-eleventh of the sample housing units being visited each month. In the remaining SMSA's, the data were collected during the 12-month period from April 1980 through March 1981 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 1980 were interviewed previously in either 1976 or 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

For the 1980 group of SMSA's, the largest SMSA from 3 of the 4 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 1980 group are: Los Angeles-Long Beach, Calif.; New York, N.Y.; and St. Louis, Mo.-Ill.

The remaining SMSA's in the 1980 group are: Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Birmingham, Ala.; Grand Rapids, Mich.; Indianapolis, Ind.; Louisville, Ky.-Ind.; Memphis, Tenn.-Ark.; Oklahoma City, Okla.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; and Salt Lake City, Utah.

In this SMSA, 12,818 housing units were eligible for interview. Of these sample housing units, 419 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 12,818 housing units eligible for interview, 1,320 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 1980 survey—The sample housing units designated to be interviewed in the 1980 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 1980 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 and 1980 Coverage Improvement Programs. (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—The sample for the SMSA's which, in 1970, 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 1970, the following three SMSA's were 100-percent permit-issuing: Los Angeles-Long Beach, Calif.; New York, N.Y.; and Sacramento, Calif. The remaining 12 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.

1976-1980 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 10 SMSA's interviewed for the first time in 1976, the Coverage Improvement Program was conducted as a part of the 1976 AHS with the Oklahoma City, Okla. SMSA receiving some updating as a part of the 1980 AHS. For each of the five SMSA's interviewed previously in 1977, the Coverage Improvement Program was conducted as a part of the 1977 AHS with the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; and Salt Lake City, Utah, SMSA's receiving some updating and refining as a part of the 1980 AHS. The following discussion applies to both the prior year (1976 or 1977) and 1980 coverage improvement procedures. For the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Oklahoma City, Okla.; and Salt Lake City, Utah, SMSA's estimates of housing units added by a specific procedure reflect units added in the prior year as well as any additions that resulted from the updating and refining in 1980.

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 1980 SMSA's. Two different procedures were used. For the first procedure, the sampling was carried out in two stages for one- and two-unit structures and in three stages for three-or-more-unit structures. For the SMSA's previously interviewed in 1976, these new construction housing units were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample (regular AHS units). For the SMSA's previously interviewed in 1977, sample units selected from one- and two-unit structures were sampled at one-fourth the rate of units originally selected for the AHS-SMSA sample, while units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage sample selection was a sample of permit offices, and the second stage was a sample of the 1969 permits within each of the selected permit offices. In the Grand Rapids, Mich., and New York, N.Y., SMSA's, 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. This procedure was employed for the 10 SMSA's previously interviewed in 1976.

For the Albany-Schenectady-Troy, N.Y.; Memphis, Tenn.-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's previously interviewed in 1977, the above procedures were performed for the first and second stages. For the third stage, structures of size three or more were divided into clusters of an expected size of two units and a sample of clusters was selected.

For the Los Angeles-Long Beach, Calif., SMSA, 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 units were then sampled at one-third the rate of regular AHS units. Since permits were not available for all sampled offices, the procedure was also used in parts of the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 2,286 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. In the Oklahoma City, Okla., SMSA, a sample of tracts was selected but not canvassed during the 1976 Coverage Improvement Program. This procedure was completed during the 1980 AHS and added 3,270 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 AHS sample housing units was selected from the 1970 sample universe at a rate of 1 in 24 for the Albany-Schenectady-Troy, N.Y.; Los Angeles-Long Beach, Calif.; Memphis, Tenn-Ark.; Saginaw, Mich.; and Salt Lake City, Utah, SMSA's and at a rate of 1 in 22 for the remaining 10 SMSA's. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample housing unit 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. The Albany-Schenectady-Troy, N.Y., SMSA, had these sampling procedures implemented in 1976, excluding identification of the housing units missed in the 1970 census. This procedure was completed as a part of the 1980 AHS. The Los Angeles-Long Beach, Calif., and Salt Lake City, Utah, SMSA's had this sampling procedure implemented for the first time in 1980. This procedure added an estimated 11,564 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 1,409 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 1980 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 1980 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1976 (i.e., 1976-1980 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1980 housing inventory—The AHS estimates of characteristics of the 1980 housing inventory were produced using a 2-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 Albany-Schenectady-Troy, N.Y.; Allentown-Bethlehem-Easton, Pa.-N.J.; Grand Rapids, Mich.; Indianapolis, Ind.; Los Angeles-Long Beach, Calif.; Louisville, Ky.-Ind.; Providence-Pawtucket-Warwick, R.I.-Mass.; Sacramento, Calif.; Saginaw, Mich.; St. Louis, Mo.-Ill.; and Salt Lake City, Utah, 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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., SMSA's. In the Birmingham, Ala.; Memphis, Tenn.-Ark.; and Oklahoma City, Okla., 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 New York, N.Y., SMSA, the sample selected correctly apportioned the new construction housing units between the central city and the balance.

The final ratio estimation procedure was employed as a second-stage ratio estimation procedure for the Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 1980 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:

$$\frac{\text{Independent estimate of the October 1980 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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 Birmingham, Ala.; Memphis, Tenn.-Ark.; New York, N.Y.; and Oklahoma City, Okla., 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 procedure, 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.

1976-1980 lost housing units—The 1976-1980 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 1976-1980 lost housing units existed, by definition, in the 1976 housing inventory, there was a 1976 housing inventory weight associated with each 1976-1980 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1976-1980 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 1980 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 of which 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 1980 survey, the number of missed housing units may be considerably less for 1980.

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

Rounding errors—For errors associated with processing, 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 1976 housing inventory can be found in the AHS Series H-170 reports for 1976.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1980 housing inventory as well as estimates of characteristics of the 1976-1980 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 23,110 for the total SMSA, 7,080 for the central city of the SMSA, and 21,680 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, III, and IV (page App-51) present the standard errors of estimated percentages for the 1980 housing inventory as well as estimated percentages of the 1976-1980 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, III, and IV.

Included in tables I through IV 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 through IV 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 1980 there were 573,400 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 4,010. 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
500,000	4,190
573,400	x
600,000	3,950

The entry for "x" is determined as follows by vertically interpolating between 4,190 and 3,950.

$$573{,}400 - 500{,}000 = 73{,}400$$
$$600{,}000 - 500{,}000 = 100{,}000$$

$$4{,}190 + \frac{73{,}400}{100{,}000}(3{,}950 - 4{,}190) = 4{,}010$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 569,390 to 577,410 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1980 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 566,980 to 579,820 housing units with 90 percent confidence; and that the average estimate lies within the interval from 565,380 to 581,420 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 573,400 owner-occupied housing units, 171,700, or 29.9 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 29.9 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	29.9	50
500,000.	0.6	a	0.6
573,400.		p	
600,000.	0.5	b	0.6

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

$$29.9-25.0 = 4.9$$
$$50.0-25.0 = 25.0$$

$$0.6 + \frac{4.9}{25.0}(0.6-0.6) = 0.6$$

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

$$29.9-25.0 = 4.9$$
$$50.0-25.0 = 25.0$$

$$0.5 + \frac{4.9}{25.0}(0.6-0.5) = 0.5$$

3. The entry for "p" is then determined by vertical interpolation between 0.6 and 0.5.

$$573{,}400-500{,}000 = 73{,}400$$
$$600{,}000-500{,}000 = 100{,}000$$

$$0.6 + \frac{73{,}400}{100{,}000}(0.5-0.6) = 0.5$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 29.4 to 30.4 percent; the 90-percent confidence interval is from 29.1 to 30.7 percent; and the 95-percent confidence interval is from 28.9 to 30.9 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; if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1976 and 1980 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1976 and 1980 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1980 there were 263,900 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 92,200. Table I shows the standard error of 171,700 is approximately 3,330, and the standard error of 263,900 is approximately 3,860. Therefore, the standard error of the estimated difference of 92,200 is about 5,100:

$$5{,}100 = \sqrt{(3{,}330)^2 + (3{,}860)^2}$$

Consequently, the 68-percent confidence interval for the 92,200 difference is from 87,100 to 97,300 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 84,040 to 100,360 housing units, and the 95-percent confidence interval is from 82,000 to 102,400 housing units. Thus, we can conclude with 95 percent confidence that the number of 1980 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, III, and IV 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. 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. 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 573,400 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 573,400 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 three persons (for purposes of calculating the median, the category of three persons is considered to be from 2.5 to 3.5 persons) contains the 48.8 percent derived in step 2. About 261,500 housing units or 45.6 percent fall below this interval, and 109,500 housing units or 19.1 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(48.8-45.6)}{19.1} = 2.7$$

Similarly, the interval for owner-occupied housing units with three persons contains the 51.2 percent derived in step 2. About 261,500 housing units or 45.6 percent fall below this interval, and 109,500 housing units or 19.1 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(51.2-45.6)}{19.1} = 2.8$$

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

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1980 Housing Inventory and for Estimated Number of 1976-1980 Lost Units for the St. Louis, Mo.-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. For estimates pertaining to lost units, the standard errors in the table should be multiplied by a factor of 1.0 for the total SMSA, 0.9 for the central city and 0.9 for the balance (not in central city) estimates)

Size of estimate	Standard error[1]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	80	40	110	150,000	3,170	1,230	3,620
100	90	60	110	200,000	3,540	110	3,980
200	130	90	150	200,300	3,540	–	3,980
500	200	140	240	250,000	3,810	–	4,210
700	240	170	280	300,000	4,000	–	4,340
1,000	280	200	330	400,000	4,210	–	4,310
2,500	450	320	530	500,000	4,190	–	3,880
5,000	630	440	750	600,000	3,950	–	2,880
10,000	890	620	1,050	684,400	3,540	–	–
25,000	1,400	940	1,640	700,000	3,440	–	–
50,000	1,950	1,230	2,280	800,000	2,490	–	–
75,000	2,360	1,380	2,730	884,700	–	–	–
100,000	2,680	1,420	3,090				

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.4 for the total SMSA, 1.0 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 1980 Housing Inventory and for Estimated Percentages of 1976-1980 Lost Housing Units for the St. Louis, Mo.-Ill., SMSA

(68 chances our of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	44.7	44.7	44.7	44.7	44.7	44.9	100,000	0.08	0.3	0.6	0.9	1.2	1.4
200	28.8	28.8	28.8	28.8	28.8	31.8	150,000	0.05	0.2	0.5	0.7	1.0	1.2
500	13.9	13.9	13.9	13.9	17.4	20.1	200,000	0.04	0.2	0.4	0.6	0.9	1.0
700	10.3	10.3	10.3	10.3	14.7	17.0	250,000	0.03	0.2	0.4	0.5	0.8	0.9
1,000	7.5	7.5	7.5	8.5	12.3	14.2	300,000	0.03	0.2	0.4	0.5	0.7	0.8
2,500	3.1	3.1	3.9	5.4	7.8	9.0	400,000	0.02	0.14	0.3	0.4	0.6	0.7
5,000	1.6	1.6	2.8	3.8	5.5	6.4	500,000	0.02	0.13	0.3	0.4	0.6	0.6
10,000	0.8	0.9	2.0	2.7	3.9	4.5	600,000	0.01	0.12	0.3	0.3	0.5	0.6
25,000	0.3	0.6	1.2	1.7	2.5	2.8	700,000	0.01	0.11	0.2	0.3	0.5	0.5
50,000	0.2	0.4	0.9	1.2	1.7	2.0	800,000	0.01	0.10	0.2	0.3	0.4	0.5
75,000	0.11	0.3	0.7	1.0	1.4	1.6	900,000	0.01	0.09	0.2	0.3	0.4	0.5

[1] Standard errors are presented to the nearest one-tenth of the 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 1.4.

TABLE III. Standard Errors for Estimated Percentages of Housing Units in the 1980 Housing Inventory and for Estimated Percentages of 1976-1980 Lost Housing Units for the Central City of the St. Louis, Mo.-Ill., SMSA

(68 chances out of 100. For estimates pertaining to lost units, the standard errors in the table should be multiplied by a factor of 0.9)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	28.8	28.8	28.8	28.8	28.8	31.8	25,000	0.2	0.4	0.9	1.2	1.7	2.0
200	16.8	16.8	16.8	16.8	19.5	22.5	50,000	0.08	0.3	0.6	0.9	1.2	1.4
500	7.5	7.5	7.5	8.5	12.3	14.2	75,000	0.05	0.2	0.5	0.7	1.0	1.2
700	5.5	5.5	5.5	7.2	10.4	12.0	100,000	0.04	0.2	0.4	0.6	0.9	1.0
1,000	3.9	3.9	4.4	6.0	8.7	10.1	150,000	0.03	0.2	0.4	0.5	0.7	0.8
2,500	1.6	1.6	2.8	3.8	5.5	6.4	200,000	0.02	0.14	0.3	0.4	0.6	0.7
5,000	0.8	0.9	2.0	2.7	3.9	4.5	250,000	0.02	0.13	0.3	0.4	0.6	0.6
10,000	0.4	0.6	1.4	1.9	2.8	3.2							

[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.

TABLE IV. Standard Errors for Estimated Percentages of Housing Units in the 1980 Housing Inventory and for Estimated Percentages of 1976-1980 Lost Housing Units for the Balance (Not in Central City) of the St. Louis, Mo.-Ill., SMSA

(68 chances out of 100. For estimates pertaining to lost units, the standard errors in the table should be multiplied by a factor of 0.9)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	52.8	52.8	52.8	52.8	52.8	52.9	75,000	0.15	0.4	0.8	1.2	1.7	1.9
200	35.9	35.9	35.9	35.9	35.9	37.4	100,000	0.11	0.3	0.7	1.0	1.4	1.7
500	18.3	18.3	18.3	18.3	20.5	23.7	150,000	0.07	0.3	0.6	0.8	1.2	1.4
700	13.8	13.8	13.8	13.8	17.3	20.0	200,000	0.06	0.2	0.5	0.7	1.0	1.2
1,000	10.1	10.1	10.1	10.1	14.5	16.7	250,000	0.04	0.2	0.5	0.6	0.9	1.1
2,500	4.3	4.3	4.6	6.3	9.2	10.6	300,000	0.04	0.2	0.4	0.6	0.8	1.0
5,000	2.2	2.2	3.3	4.5	6.5	7.5	400,000	0.03	0.2	0.4	0.5	0.7	0.8
10,000	1.1	1.1	2.3	3.2	4.6	5.3	500,000	0.02	0.15	0.3	0.4	0.6	0.7
25,000	0.4	0.7	1.5	2.0	2.9	3.3	600,000	0.02	0.14	0.3	0.4	0.6	0.7
50,000	0.2	0.5	1.0	1.4	2.0	2.4	700,000	0.02	0.13	0.3	0.4	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 1.2.