1981 Metropolitan Appendix B: Source and Accuracy

Table of Contents

Anaheim-Santa Ana-Garden Grove 3

Boston 14

Dallas 26

Detroit 37

Ft. Worth 49

Madison 60

Minneapolis-St. Paul 71

Newark 82

Orlando 93

Phoenix 104

Pittsburgh 115

Spokane 126

Tacoma 137

Washington 149

Wichita 161

Anaheim-Santa Ana-Garden Grove 1981

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN . . . App-42
Annual Housing Survey . . . App-42
Designation of sample housing units for the 1981 survey . . . App-42
Selection of the original AHS-SMSA sample . . . App-43
1977-1981 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-45
1970 Census of Population and Housing . . . App-45
ESTIMATION . . . App-45
1981 housing inventory . . . App-45
1977-1981 lost housing units . . . App-47
1977 estimation procedure . . . App-47
Ratio estimation procedure of the 1970 Census of Population and Housing . . . App-47
RELIABILITY OF THE ESTIMATES . . . App-47
Nonsampling errors . . . App-47
1970 census . . . App-47
AHS-SMSA . . . App-47
Coverage errors . . . App-47
Rounding errors . . . App-48
Sampling errors for the AHS-SMSA sample . . . App-48
Illustration of the use of the standard error tables . . . App-49
Differences . . . App-49
Illustration of the computation of the standard error of a difference . . . App-50
Medians . . . App-50
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-170-81) are based on data collected from the 1981 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.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1981 were interviewed previously in 1974 (1975 for Madison, Wis.) and 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

The three largest SMSA's (Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va.) in the 1981 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1974 and 1977 surveys evenly divided between the central city and the balance of the respective SMSA. The 12 remaining SMSA's (Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.) in the 1981 group were represented by a sample of 5,000 designated housing units in the 1974 (1975 for Madison, Wis.) and 1977 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1974 survey, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panel 3 was dropped from all SMSA's for the 1977 survey. Due to additional budget limitations for the 1981 survey, it also became necessary to drop panels 1, 2, 4, 5, 9, 11, and 12 from the Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va., SMSA's and panels 1 and 2 from the other 12 SMSA's. The designated sample size was reduced by two-thirds for the three largest SMSA's and by one-fourth for the other 12 SMSA's.

The interviewing was done for the three largest SMSA's during the period June 1981 through November 1981 with one-sixth of the sample housing units being visited each month. The interviewing was done for the remaining 12 SMSA's during the period April 1981 through December 1981 with one-ninth of the designated housing units visited each month.

In this SMSA, 4,798 housing units were eligible for interview. Of these sample housing units, 186 interviews were not obtained because, for occupied sample units, the occupants refused to be interviewed, were not at home after repeated visits, or were unavailable for some other reason; or, for vacant units, no informed respondent could be found after repeated visits. In addition to the 4,798 housing units eligible for interview, 136 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 1981 survey—The sample housing units designated to be interviewed in the 1981

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 and remained in sample after the 1981 panel reduction.
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 and remained in sample after the 1981 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1981 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 and remained in sample after the 1981 panel reduction. (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 and remained in sample after the 1981 panel reduction. (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 1981 Coverage Improvement Programs. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the original 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 five SMSA's were 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 10 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 for the 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[2]	1 2 3 4 5[2]
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.

1977-1981 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 from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes 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. Houses that have been moved onto their present site since the 1970 census.
6. Mobile homes placed outside parks since the 1970 census or vacant at the time of the 1970 census.

Each of the 15 SMSA's was previously interviewed in 1974 (1975 for Madison, Wis.) and 1977. The Coverage Improvement Program was conducted as part of the 1977 AHS with the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., SMSA's receiving some updating and refining as part of the 1981 AHS. The following discussion applies to both the prior year 1977 and the 1981 coverage improvement procedures. For the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., 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 1981.

Coverage improvement for deficiency 1—A sample of new construction units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 15 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. 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 (regular AHS units), while sample units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage 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 Fort Worth, Tex., SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second-stage sampling. For the third stage, structures of size three or more units were divided into clusters of an expected size of two units and a sample of clusters was selected. This procedure was employed in the following SMSA's: Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Orlando, Fla.; Spokane, Wash.; and Wichita, Kans.

In the remaining 9 SMSA's, 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, this procedure was also used in parts of the Anaheim-Santa Ana-Garden Grove, Calif., SMSA to supplement the sample described above. These procedures added an estimated 21,216 new construction 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 placed in a park that was 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 units represented the same number of units that the regular AHS sample units represented. This procedure added an estimated 14,678 units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiences 3-6—The remaining missed 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 units converted to residential use.
3. Mobile homes 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 a usual residence elsewhere.
4. Houses that had been moved onto their present site since the 1970 census.

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample 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 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 second procedure was designed to represent missed units from structures represented in the AHS. These missed units were:

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential 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 structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply both procedures during the 1977 Coverage Improvement Program to all SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in the Anaheim-Santa Ana-Garden Grove, Calif.; and Spokane, Wash., SMSA's. These SMSA's did not receive application of the second procedure. The Dallas, Tex.; Minneapolis-St. Paul, Minn.; and Pittsburgh, Pa., SMSA's did not receive application of either procedure during the 1977 Coverage Improvement Program. The Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and the Spokane, Wash., SMSA's had the first procedure completed as part of the 1981 Coverage Improvement Program. The remaining 10 1981 SMSA's received the full application of both procedures as part of the 1977 Coverage Improvement Program.

The first procedure added an estimated 10,243 housing units to the coverage of the housing inventory of 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 1981 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 1981 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1977 (i.e., 1977-1981 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1981 housing inventory—The AHS estimates of characteristics of the 1981 housing inventory were produced using a one-stage ratio estimation procedure for the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; and Madison, Wis., SMSA's, a two-stage ratio estimation procedure for the Detroit, Mich.; Orlando, Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, and a three-stage ratio estimation procedure for the remaining seven SMSA's.

Prior to the implementation of the ratio estimation procedures, 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 nonpermit 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 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.

The next ratio estimation procedure was applied in the Boston, Mass.; Detroit, Mich.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's. This procedure was utilized 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 ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

The next ratio estimation procedure was applied in 10 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1981 housing inventory in each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1981 housing unit inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

The independent estimates of total housing units that were used as the numerator of this ratio are described below. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight.

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for each of the 15 1981 SMSA's. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1981 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units after the permit new construction ratio estimation (i.e., the existing sample estimate) were compared to the corresponding independent estimates for the central city, balance, and total SMSA for each of the 15 SMSA's and the estimate which showed the most likely level of net growth since the 1980 census in both the central city and balance as well as the total SMSA were used in this ratio estimation. As a result of this analysis, these independent estimates were used in the Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's.

The remaining five SMSA's used a combination of these independent estimates and sample estimates. For the Boston, Mass.; and Minneapolis-St. Paul, Minn., SMSA's, the sample estimate was used for the total SMSA. For the sectors, the sample estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the independent estimates.

For the Orlando Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, the independent estimate of units was used for the total SMSA. For the sectors, the independent estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the sample estimates. Due to the central city

boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimate did not reflect the 1970 central city/balance SMSA definition. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definition were used.

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 in these 10 SMSA's, and the resulting product was used as the final weight for tabulation purposes. For the other five SMSA's, the sample estimates were used as the final weight for tabulation purposes. These SMSA's included Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Detroit, Mich.; Fort Worth, Tex.; and Madison, Wis., SMSA's.

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

1977-1981 lost housing units—The 1977-1981 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1977 housing inventory for 14 of the SMSA's and the corresponding two-stage ratio estimation procedure used for the Pittsburgh, Pa., SMSA as was described in the 1977 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1977-1981 lost housing units existed, by definition, in the 1977 housing inventory, there was a 1977 housing inventory weight associated with each 1977-1981 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1977-1981 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 two-stage ratio estimation process for the Pittsburgh, Pa., SMSA and a one-stage ratio estimation process for the remaining SMSA's. 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; 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 1981 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, and the results 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) 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, the 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 presented.

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 1981 survey, the number of missed housing units may be considerably less for 1981.

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 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 1981 housing inventory as well as estimates of characteristics of the 1977-1981 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 28,570 for the total SMSA, 14,390 for the central city of the SMSA, and 24,600 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 1981 housing inventory as well as estimated percentages of the 1977-1981 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, tables II through IV 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 1981 there were 425,000 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 7,680. The following interpolation procedure was used.

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

Size of estimate	Standard error
400,000	7,520
425,000	x
500,000	8,170

The entry of "x" is determined as follows by vertically interpolating between 7,520 and 8,170.

$$425{,}000-400{,}000 = 25{,}000$$
$$500{,}000-400{,}000 = 100{,}000$$

$$7{,}520 + \frac{25{,}000}{100{,}000}(8{,}170-7{,}520) = 7{,}680$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 417,320 to 432,680 housing units. Therefore, a conclusion that the average estimate of 1981 owner-occupied housing units (derived from all possible samples) 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 412,710 to 437,290 housing units with 90 percent confidence; and that the average estimate lies within the interval from 409,640 to 440,360 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 425,000 owner-occupied housing units, 95,500, 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 0.9 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	22.5	25 or 75
400,000	0.6	a	0.9
425,000		p	
500,000	0.6	b	0.8

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

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

$$0.6 + \frac{12.5}{15.0}(0.9-0.6) = 0.9$$

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

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

$$0.6 + \frac{12.5}{15.0}(0.8-0.6) = 0.8$$

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

$$425{,}000-400{,}000 = 25{,}000$$
$$500{,}000-400{,}000 = 100{,}000$$

$$0.9 + \frac{25{,}000}{100{,}000}(0.8-0.9) = 0.9$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 21.6 to 23.4 percent; the 90-percent confidence interval is from 21.1 to 23.9 percent; and the 95-percent confidence interval is from 20.7 to 24.3 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 1977 and 1981 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1977 and 1981 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1981 there were 175,800 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 80,300. Table I shows the standard error of 95,500 is approximately 3,970, and the standard error of 175,800 is approximately 5,280. Therefore, the standard error of the estimated difference of 80,300 is about 6,610.

$$6{,}610 = \sqrt{(3{,}970)^2 + (5{,}280)^2}$$

Consequently, the 68-percent confidence interval for the 80,300 difference is from 73,690 to 86,910 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 69,720 to 90,880 housing units, and the 95-percent confidence interval is from 67,080 to 93,520 housing units. Thus, we can conclude with 95 percent confidence that the number of 1981 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied 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 to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.7. The base of the distribution from which this median was determined is 425,000 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 425,000 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 199,800 housing units or 47.0 percent fall below this interval, and 74,100 housing units or 17.4 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(47.8-47.0)}{17.4} = 2.5$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.2 percent derived in step 2. About 199,800 housing units or 47.0 percent fall below this interval, and 74,100 housing units or 17.4 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(52.2-47.0)}{17.4} = 2.8$$

Thus, the 95-percent confidence interval ranges from 2.5 to 2.8 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1981 Housing Inventory and for Estimated Number of 1977-1981 Lost Units for the Anaheim-Santa Ana-Garden Grove, Calif., SMSA, for the Central Cities of the SMSA and for the Balance (Not in Central Cities) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]			Size of estimate	Standard error[1]		
	SMSA	In central cities	Not in central cities		SMSA	In central cities	Not in central cities
0	170	160	180	75,000	3,540	3,490	3,520
100	170	160	180	100,000	4,060	4,060	4,010
200	190	180	190	150,000	4,920	5,030	4,790
500	290	280	300	200,000	5,610	5,870	5,380
700	350	330	350	250,000	6,190	—	5,840
1,000	420	400	420	300,000	6,690	—	6,210
2,500	660	630	670	400,000	7,520	—	6,700
5,000	930	890	940	500,000	8,170	—	6,930
10,000	1,310	1,250	1,330	600,000	8,670	—	6,910
25,000	2,070	1,990	2,080	700,000	9,070	—	—
50,000	2,910	2,830	2,910	800,000	9,360	—	—

[1]For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.3 for the total SMSA, 1.2 for the central cities, and 1.2 for the balance (not in central cities) estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Anaheim-Santa Ana-Garden Grove, Calif., SMSA, for the Central Cities and for the Balance (Not in Central Cities) 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	63.9	63.9	63.9	63.9	63.9	66.6	100,000	0.2	0.4	0.9	1.3	1.8	2.1
200	47.0	47.0	47.0	47.0	47.0	47.1	150,000	0.12	0.3	0.7	1.0	1.5	1.7
500	26.2	26.2	26.2	26.2	26.2	29.8	200,000	0.09	0.3	0.6	0.9	1.3	1.5
700	20.2	20.2	20.2	20.2	21.8	25.2	250,000	0.07	0.3	0.6	0.8	1.2	1.3
1,000	15.1	15.1	15.1	15.1	18.2	21.0	300,000	0.06	0.2	0.5	0.7	1.1	1.2
2,500	6.6	6.6	6.6	8.0	11.5	13.3	400,000	0.04	0.2	0.5	0.6	0.9	1.1
5,000	3.4	3.4	4.1	5.6	8.2	9.4	500,000	0.04	0.2	0.4	0.6	0.8	0.9
10,000	1.7	1.7	2.9	4.0	5.8	6.7	600,000	0.03	0.2	0.4	0.5	0.7	0.9
25,000	0.7	0.8	1.8	2.5	3.6	4.2	700,000	0.03	0.2	0.3	0.5	0.7	0.8
50,000	0.4	0.6	1.3	1.8	2.6	3.0	800,000	0.02	0.15	0.3	0.4	0.6	0.7
75,000	0.2	0.5	1.1	1.5	2.1	2.4							

[1]Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.3 for the total SMSA, 1.2 for the central cities, and 1.2 for the balance (not in central cities).

Boston 1981

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1981 survey App-42
Selection of the original AHS-SMSA sample App-43
1977-1981 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-45
1970 Census of Population and Housing App-45
ESTIMATION App-45
1981 housing inventory App-45
1977-1981 lost housing units ... App-47
1977 estimation procedure App-47
Ratio estimation procedure of the 1970 Census of Population and Housing App-47
RELIABILITY OF THE ESTIMATES .. App-47
Nonsampling errors App-47
1970 census App-47
AHS-SMSA App-47
Coverage errors App-47
Rounding errors App-48
Sampling errors for the AHS-SMSA sample App-48
Illustration of the use of the standard error tables App-49
Differences App-49
Illustration of the computation of the standard error of a difference App-50
Medians App-50
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-170-81) are based on data collected from the 1981 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.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1981 were interviewed previously in 1974 and 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

The three largest SMSA's (Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va.) in the 1981 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1974 and 1977 surveys evenly divided between the central city and the balance of the respective SMSA. The 12 remaining SMSA's (Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Forth Worth, Tex.; Madison, Wis.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.) in the 1981 group were represented by a sample of 5,000 designated housing units in the 1974 and 1977 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1974 survey, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panel 3 was dropped from all SMSA's for the 1977 survey. Due to additional budget limitations for the 1981 survey, it also became necessary to drop panels 1, 2, 4, 5, 9, 11, and 12 from the Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va., SMSA's and panels 1 and 2 from the other 12 SMSA's. The designated sample size was reduced by two-thirds for the three largest SMSA's and by one-fourth for the other 12 SMSA's.

The interviewing was done for the three largest SMSA's during the period June 1981 through November 1981 with one-sixth of the sample housing units being visited each month. The interviewing was done for the remaining 12 SMSA's during the period April 1981 through December 1981 with one-ninth of the designated housing units visited each month.

In this SMSA, 5,124 housing units were eligible for interview. Of these sample housing units, 239 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 5,124 housing units eligible for interview, 577 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 1981 survey—The sample housing units designated to be interviewed in the 1981

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 and remained in sample after the 1981 panel reduction.
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 and remained in sample after the 1981 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1981 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 and remained in sample after the 1981 panel reduction. (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 and remained in sample after the 1981 panel reduction. (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 1981 Coverage Improvement Programs. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the original 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 non-permit universe). In 1970, the following five SMSA's were 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 10 SMSA's contain a sample from the non-permit 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 for the 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.

1977-1981 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 from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes 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. Houses that have been moved onto their present site since the 1970 census.
6. Mobile homes placed outside parks since the 1970 census or vacant at the time of the 1970 census.

Each of the 15 SMSA's was previously interviewed in 1974 and 1977. The Coverage Improvement Program was conducted as part of the 1977 AHS with the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., SMSA's receiving some updating and refining as part of the 1981 AHS. The following discussion applies to both the prior year 1977 and the 1981 coverage improvement procedures. For the the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., 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 1981.

Coverage improvement for deficiency 1—A sample of new construction units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 15 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. 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 (regular AHS units), while sample units selected from three-or-more unit structures were sampled at one-half the rate of regular AHS units.

The first stage 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 Fort Worth, Tex., SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second-stage sampling. For the third stage, structures of size three or more units were divided into clusters of an expected size of two units and a sample of clusters was selected. This procedure was employed in the following SMSA's: Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Orlando, Fla.; Spokane, Wash.; and Wichita, Kans.

In the remaining 9 SMSA's, 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, this procedure was also used in parts of the Anaheim-Santa AnaGarden Grove, Calif., SMSA to supplement the sample described above. These procedures added an estimated 13,101 new construction 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 placed in a park that was 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 units

represented the same number of units that the regular AHS sample units represented. However, this procedure did not add any units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiences 3-6—The remaining missed 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 units converted to residential use.
3. Mobile homes 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 a usual residence elsewhere.
4. Houses that had been moved onto their present site since the 1970 census.

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample 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 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 second procedure was designed to represent missed units from structures represented in the AHS. These missed units were:

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential 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 structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply both procedures during the 1977 Coverage Improvement Program to all SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in the Anaheim-Santa Ana-Garden Grove, Calif.; and Spokane, Wash., SMSA's. These SMSA's received the full application of the second procedure at that time. The Dallas, Tex.; Minneapolis-St. Paul, Minn.; and Pittsburgh, Pa., SMSA's did not receive application of either procedure during the 1977 Coverage Improvement Program. The Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and the Spokane, Wash., SMSA's had the first procedure completed as part of the 1981 Coverage Improvement Program. The remaining 10 1981 SMSA's received the full application of both procedures as part of the 1977 Coverage Improvement Program.

The first procedure added an estimated 39,217 housing units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 4,192 housing units.

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 1981 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 1981 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1977 (i.e., 1977-1981 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1981 housing inventory—The AHS estimates of characteristics of the 1981 housing inventory were produced using a one-stage ratio estimation procedure for the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; and Madison, Wis., SMSA's, a two-stage ratio estimation procedure for the Detroit, Mich.; Orlando, Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, and a three-stage ratio estimation procedure for the remaining seven SMSA's.

Prior to the implementation of the ratio estimation procedures, 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 nonpermit 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 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.

The next ratio estimation procedure was applied in the Boston, Mass.; Detroit, Mich.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's. This procedure was utilized 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 ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

The next ratio estimation procedure was applied in 10 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1981 housing inventory in each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1981 housing unit inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

The independent estimates of total housing units that were used as the numerator of this ratio are described below. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight.

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for each of the 15 1981 SMSA's. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1981 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units after the permit new construction ratio estimation (i.e., the existing sample estimate) were compared to the corresponding independent estimates for the central city, balance, and total SMSA for each of the 15 SMSA's and the estimate which showed the most likely level of net growth since the 1980 census in both the central city and balance as well as the total SMSA were used in this ratio estimation. As a result of this analysis, these independent estimates were used in the Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's.

The remaining five SMSA's used a combination of these independent estimates and sample estimates. For the Boston, Mass.; and Minneapolis-St. Paul, Minn., SMSA's, the sample estimate was used for the total SMSA. For the sectors, the sample estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the independent estimates.

For the Orlando Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, the independent estimate of units was used for the total SMSA. For the sectors, the independent estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the sample estimates. Due to the central city

boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimate did not reflect the 1970 central city/balance SMSA definition. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definition were used.

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 in these 10 SMSA's, and the resulting product was used as the final weight for tabulation purposes. For the other five SMSA's, the sample estimates were used as the final weight for tabulation purposes. These SMSA's included Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Detroit, Mich.; Fort Worth, Tex.; and Madison, Wis., SMSA's.

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

1977-1981 lost housing units—The 1977-1981 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1977 housing inventory for 14 of the SMSA's and the corresponding two-stage ratio estimation procedure used for the Pittsburgh, Pa., SMSA as was described in the 1977 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1977-1981 lost housing units existed, by definition, in the 1977 housing inventory, there was a 1977 housing inventory weight associated with each 1977-1981 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1977-1981 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 two-stage ratio estimation process for the Pittsburgh, Pa., SMSA and a one-stage ratio estimation process for the remaining SMSA's. 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; 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 1981 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, and the results 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) 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, the 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 presented.

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 1981 survey, the number of missed housing units may be considerably less for 1981.

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 standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. Approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the average result of all possible samples.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimate would include the average result of all possible samples.

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

The figures presented in the tables that follow (pages App-51 through App-52) 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 1981 housing inventory as well as estimates of characteristics of the 1977-1981 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 35,668 for the total SMSA, 10,252 for the central city of the SMSA, and 33,986 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 through IV (pages App-51 through App-52) present the standard errros of estimated percentages for the 1981 housing inventory as well as estimated percentages of the 1977-1981 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 through 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 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 1981 there were 511,500 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 8,970. 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	8,910
511,500	x
600,000	9,440

The entry of "x" is determined as follows by vertically interpolating between 8,910 and 9,440.

$$511{,}500-500{,}000=11{,}500$$
$$600{,}000-500{,}000=100{,}000$$

$$8{,}910+\frac{11{,}500}{100{,}000}(9{,}440-8{,}910)=8{,}970$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 502,530 to 520,470 housing units. Therefore, a conclusion that the average estimate of 1981 owner-occupied housing units (derived from all possible samples) 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 497,150 to 525,850 housing units with 90 percent confidence; and that the average estimate lies within the interval from 493,560 to 529,440 housing units with 95 percent confidence.

Table A-1 part A also shows that of the 511,500 owner-occupied housing units, 124,300, 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
500,000	0.6	a	0.9
511,500		p	
600,000	0.6	b	0.8

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

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

$$0.6+\frac{14.3}{15.0}(0.9-0.6)=0.9$$

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

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

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

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

$$511{,}500-500{,}000=11{,}500$$
$$600{,}000-500{,}000=100{,}000$$

$$0.9+\frac{11{,}500}{100{,}000}(0.8-0.9)=0.9$$

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; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1977 and 1981 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1977 and 1981 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1981 there were 229,300 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 105,000. Table I shows the standard error of 124,300 is approximately 4,930, and the standard error of 229,300 is approximately 6,540. Therefore, the standard error of the estimated difference of 105,000 is about 8,190.

$$8{,}190 = \sqrt{(4{,}930)^2 + (6{,}540)^2}$$

Consequently, the 68-percent confidence interval for the 105,000 difference is from 96,810 to 113,910 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 91,900 to 118,100 housing units, and the 95-percent confidence interval is from 88,620 to 121,380 housing units. Thus, we can conclude with 95 percent confidence that the number of 1981 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied 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 through 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 into which interval of the distribution the lower percentage limit falls. Similarly, to find the upper endpoint to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.9. The base of the distribution from which this median was determined is 511,500 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 511,500 is approximately 1.0 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.0 and 52.0
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with 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 222,700 housing units or 43.5 percent fall below this interval, and 90,800 housing units or 17.8 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(48.0-43.5)}{17.8} = 2.8$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.0 percent derived in step 2. About 222,700 housing units or 43.5 percent fall below this interval, and 90,800 housing units or 17.8 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(52.0-43.5)}{17.8} = 3.0$$

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

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1981 Housing Inventory and for Estimated Number of 1977-1981 Lost Units for the Boston, 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	210	120	300	150,000	5,410	3,100	6,350
100	210	120	300	200,000	6,170	3,070	7,190
200	210	150	300	250,000	6,800	2,740	7,870
500	320	240	390	300,000	7,340	–	8,430
700	380	280	460	400,000	8,230	–	9,280
1,000	460	340	550	500,000	8,910	–	9,850
2,500	730	540	870	600,000	9,440	–	10,170
5,000	1,030	750	1,230	700,000	9,930	–	10,280
10,000	1,450	1,060	1,730	800,000	10,110	–	10,170
25,000	2,280	1,630	2,720	900,000	10,280	–	–
50,000	3,210	2,220	3,810	1,000,000	10,340	–	–
75,000	3,900	2,600	4,630	1,100,000	10,310	–	–
100,000	4,480	2,850	5,290				

[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 1.1 for the balance (not in central city) estimates. For estimates pertaining to lost units, the standard errors in the table should be multiplied by a factor of 1.2 for the total SMSA, 1.0 for the central city, and 1.4 for the balance (not in central city) estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Boston, Mass., 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	67.8	67.8	67.8	67.8	67.8	72.6	150,000	0.14	0.4	0.8	1.1	1.6	1.9
200	51.3	51.3	51.3	51.3	51.3	51.3	200,000	0.11	0.3	0.7	1.0	1.4	1.6
500	29.7	29.7	29.7	29.7	29.7	32.5	250,000	0.08	0.3	0.6	0.9	1.3	1.5
700	23.1	23.1	23.1	23.1	23.8	27.4	300,000	0.07	0.3	0.6	0.8	1.1	1.3
1,000	17.4	17.4	17.4	17.4	19.9	23.0	400,000	0.05	0.2	0.5	0.7	1.0	1.1
2,500	7.8	7.8	7.8	8.7	12.6	14.5	500,000	0.04	0.2	0.4	0.6	0.9	1.0
5,000	4.0	4.0	4.5	6.2	8.9	10.3	600,000	0.04	0.2	0.4	0.6	0.8	0.9
10,000	2.1	2.1	3.2	4.4	6.3	7.3	700,000	0.03	0.2	0.4	0.5	0.8	0.9
25,000	0.8	0.9	2.0	2.8	4.0	4.6	800,000	0.03	0.2	0.4	0.5	0.7	0.8
50,000	0.4	0.6	1.4	1.9	2.8	3.2	900,000	0.02	0.2	0.3	0.5	0.7	0.8
75,000	0.3	0.5	1.2	1.6	2.3	2.7	1,000,000	0.02	0.14	0.3	0.4	0.6	0.7
100,000	0.2	0.5	1.0	1.4	2.0	2.3	1,100,000	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 and lost units, the standard errors shown in the table should be multiplied by a factor of 1.2.

TABLE III. Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Central City of the Boston, Mass., 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	53.6	53.6	53.6	53.6	53.6	53.7	25,000	0.5	0.7	1.5	2.0	2.9	3.4
200	36.6	36.6	36.6	36.6	36.6	38.0	50,000	0.2	0.5	1.0	1.4	2.1	2.4
500	18.8	18.8	18.8	18.8	20.8	24.0	75,000	0.2	0.4	0.9	1.2	1.7	2.0
700	14.2	14.2	14.2	14.2	17.6	20.3	100,000	0.12	0.3	0.7	1.0	1.5	1.7
1,000	10.4	10.4	10.4	10.4	14.7	17.0	150,000	0.08	0.3	0.6	0.8	1.2	1.4
2,500	4.4	4.4	4.7	6.4	9.3	10.7	200,000	0.06	0.2	0.5	0.7	1.0	1.2
5,000	2.3	2.3	3.3	4.6	6.6	7.6	250,000	0.05	0.2	0.5	0.6	0.9	1.1
10,000	1.1	1.1	2.3	3.2	4.7	5.4							

[1]Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction the standard errors shown in the table should be multiplied by a factor of 1.1.

TABLE IV. Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Balance (Not in Central City) of the Boston, Mass., 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	75.1	75.1	75.1	75.1	75.2	86.8	100,000	0.3	0.5	1.2	1.6	2.4	2.7
200	60.1	60.1	60.1	60.1	60.1	61.4	150,000	0.2	0.4	1.0	1.3	1.9	2.2
500	37.6	37.6	37.6	37.6	37.6	38.8	200,000	0.2	0.4	0.8	1.2	1.7	1.9
700	30.1	30.1	30.1	30.1	30.1	32.8	250,000	0.12	0.3	0.8	1.0	1.5	1.7
1,000	23.2	23.2	23.2	23.2	23.8	27.5	300,000	0.10	0.3	0.7	1.0	1.4	1.6
2,500	10.8	10.8	10.8	10.8	15.0	17.4	400,000	0.08	0.3	0.6	0.8	1.2	1.4
5,000	5.7	5.7	5.7	7.4	10.6	12.3	500,000	0.06	0.2	0.5	0.7	1.1	1.2
10,000	2.9	2.9	3.8	5.2	7.5	8.7	600,000	0.05	0.2	0.5	0.7	1.0	1.1
25,000	1.2	1.2	2.4	3.3	4.8	5.5	700,000	0.04	0.2	0.5	0.6	0.9	1.0
50,000	0.6	0.8	1.7	2.3	3.4	3.9	800,000	0.04	0.2	0.4	0.6	0.8	1.0
75,000	0.4	0.6	1.4	1.9	2.7	3.2							

[1]Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is lessproof 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 estimates pertaining to lost units, the standard errors shown in the table should be multiplied or a factor of 1.4.

Dallas 1981

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1981 survey App-42
Selection of the original AHS-SMSA sample App-43
1977-1981 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-45
1970 Census of Population and Housing App-45
ESTIMATION App-45
1981 housing inventory App-45
1977-1981 lost housing units ... App-47
1977 estimation procedure App-47
Ratio estimation procedure of the 1970 Census of Population and Housing App-47
RELIABILITY OF THE ESTIMATES .. App-47
Nonsampling errors App-47
1970 census App-47
AHS-SMSA App-47
Coverage errors App-47
Rounding errors App-48
Sampling errors for the AHS-SMSA sample App-48
Illustration of the use of the standard error tables App-49
Differences App-49
Illustration of the computation of the standard error of a difference App-50
Medians App-50
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-170-81) are based on data collected from the 1981 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.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1981 were interviewed previously in 1974 (1975 for Madison, Wis.) and 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

The three largest SMSA's (Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va.) in the 1981 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1974 and 1977 surveys evenly divided between the central city and the balance of the respective SMSA. The 12 remaining SMSA's (Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.) in the 1981 group were represented by a sample of 5,000 designated housing units in the 1974 (1975 for Madison, Wis.) and 1977 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1974 survey, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panel 3 was dropped from all SMSA's for the 1977 survey. Due to additional budget limitations for the 1981 survey, it also became necessary to drop panels 1, 2, 4, 5, 9, 11, and 12 from the Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va., SMSA's and panels 1 and 2 from the other 12 SMSA's. The designated sample size was reduced by two-thirds for the three largest SMSA's and by one-fourth for the other 12 SMSA's.

The interviewing was done for the three largest SMSA's during the period June 1981 through November 1981 with one-sixth of the sample housing units being visited each month. The interviewing was done for the remaining 12 SMSA's during the period April 1981 through December 1981 with one-ninth of the designated housing units visited each month.

In this SMSA, 4,449 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 4,449 housing units eligible for interview, 340 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 1981 survey—The sample housing units designated to be interviewed in the 1981

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 and remained in sample after the 1981 panel reduction.
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 and remained in sample after the 1981 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1981 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 and remained in sample after the 1981 panel reduction. (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 and remained in sample after the 1981 panel reduction. (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 1981 Coverage Improvement Programs. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the original 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 five SMSA's were 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 10 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 for the 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.

1977-1981 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 from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes 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. Houses that have been moved onto their present site since the 1970 census.
6. Mobile homes placed outside parks since the 1970 census or vacant at the time of the 1970 census.

Each of the 15 SMSA's was previously interviewed in 1974 (1975 for Madison, Wis.) and 1977. The Coverage Improvement Program was conducted as part of the 1977 AHS with the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., SMSA's receiving some updating and refining as part of the 1981 AHS. The following discussion applies to both the prior year 1977 and the 1981 coverage improvement procedures. For the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., 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 1981.

Coverage improvement for deficiency 1—A sample of new construction units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 15 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. 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 (regular AHS units), while sample units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage 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 Fort Worth, Tex., SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second-stage sampling. For the third stage, structures of size three or more units were divided into clusters of an expected size of two units and a sample of clusters was selected. This procedure was employed in the following SMSA's: Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Orlando, Fla.; Spokane, Wash.; and Wichita, Kans.

In the remaining 9 SMSA's, 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, this procedure was also used in parts of the Anaheim-Santa Ana-Garden Grove, Calif., SMSA to supplement the sample described above. These procedures added an estimated 8,677 new construction 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 placed in a park that was 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 units represented the same number of units that the regular AHS sample units represented. This procedure added an estimated 5,031 units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed 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 units converted to residential use.
3. Mobile homes 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 a usual residence elsewhere.
4. Houses that had been moved onto their present site since the 1970 census.

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample 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 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 second procedure was designed to represent missed units from structures represented in the AHS. These missed units were:

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential 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 structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply both procedures during the 1977 Coverage Improvement Program to all SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in the Anaheim-Santa Ana-Garden Grove, Calif.; and Spokane, Wash., SMSA's. These SMSA's did not receive application of the second procedure. The Dallas, Tex.; Minneapolis-St. Paul, Minn.; and Pittsburgh, Pa., SMSA's did not receive application of either procedure during the 1977 Coverage Improvement Program. The Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and the Spokane, Wash., SMSA's had the first procedure completed as part of the 1981 Coverage Improvement Program. The remaining 10 1981 SMSA's received the full application of both procedures as part of the 1977 Coverage Improvement Program.

The first procedure added an estimated 18,715 housing units to the coverage of the housing inventory of 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 1981 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 1981 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1977 (i.e., 1977-1981 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1981 housing inventory—The AHS estimates of characteristics of the 1981 housing inventory were produced using a one-stage ratio estimation procedure for the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; and Madison, Wis., SMSA's, a two-stage ratio estimation procedure for the Detroit, Mich.; Orlando, Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, and a three-stage ratio estimation procedure for the remaining seven SMSA's.

Prior to the implementation of the ratio estimation procedures, 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 nonpermit 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 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.

The next ratio estimation procedure was applied in the Boston, Mass.; Detroit, Mich.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's. This procedure was utilized 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 ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

The next ratio estimation procedure was applied in 10 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1981 housing inventory in each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1981 housing unit inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

The independent estimates of total housing units that were used as the numerator of this ratio are described below. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight.

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for each of the 15 1981 SMSA's. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1981 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units after the permit new construction ratio estimation (i.e., the existing sample estimate) were compared to the corresponding independent estimates for the central city, balance, and total SMSA for each of the 15 SMSA's and the estimate which showed the most likely level of net growth since the 1980 census in both the central city and balance as well as the total SMSA were used in this ratio estimation. As a result of this analysis, these independent estimates were used in the Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's.

The remaining five SMSA's used a combination of these independent estimates and sample estimates. For the Boston, Mass.; and Minneapolis-St. Paul, Minn., SMSA's, the sample estimate was used for the total SMSA. For the sectors, the sample estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the independent estimates.

For the Orlando Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, the independent estimate of units was used for the total SMSA. For the sectors, the independent estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the sample estimates. Due to the central city

boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimate did not reflect the 1970 central city/balance SMSA definition. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definition were used.

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 in these 10 SMSA's, and the resulting product was used as the final weight for tabulation purposes. For the other five SMSA's, the sample estimates were used as the final weight for tabulation purposes. These SMSA's included Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Detroit, Mich.; Fort Worth, Tex.; and Madison, Wis., SMSA's.

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

1977-1981 lost housing units—The 1977-1981 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1977 housing inventory for 14 of the SMSA's and the corresponding two-stage ratio estimation procedure used for the Pittsburgh, Pa., SMSA as was described in the 1977 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1977-1981 lost housing units existed, by definition, in the 1977 housing inventory, there was a 1977 housing inventory weight associated with each 1977-1981 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1977-1981 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 two-stage ratio estimation process for the Pittsburgh, Pa., SMSA and a one-stage ratio estimation process for the remaining SMSA's. 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; 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 1981 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, and the results 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) 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, the 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 presented.

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 1981 survey, the number of missed housing units may be considerably less for 1981.

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 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 1981 housing inventory as well as estimates of characteristics of the 1977-1981 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 36,481 for the total SMSA, 19,847 for the central city of the SMSA, and 30,610 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 1981 housing inventory as well as estimated percentages of the 1977-1981 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, tables II through IV 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 1981 there were 434,900 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 9,610. The following interpolation procedure was used.

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

Size of estimate	Standard error
400,000	9,190
434,900	x
500,000	10,390

The entry of "x" is determined as follows by vertically interpolating between 9,190 and 10,390.

$$434{,}900-400{,}000 = 34{,}900$$
$$500{,}000-400{,}000 = 100{,}000$$

$$9{,}190 + \frac{34{,}900}{100{,}000}(10{,}390-9{,}190) = 9{,}610$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 425,290 to 444,510 housing units. Therefore, a conclusion that the average estimate of 1981 owner-occupied housing units (derived from all possible samples) 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 419,520 to 450,280 housing units with 90 percent confidence; and that the average estimate lies within the interval from 415,680 to 454,120 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 434,900 owner-occupied housing units, 99,300, or 22.8 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 22.8 percent is approximately 1.0 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	22.8	25 or 75
400,000	0.7	a	1.0
434,900		p	
500,000	0.6	b	0.9

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

$$22.8-10.0 = 12.8$$
$$25.0-10.0 = 15.0$$

$$0.7 + \frac{12.8}{15.0}(1.0-0.7) = 1.0$$

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

$$22.8-10.0 = 12.8$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{12.8}{15.0}(0.9-0.6) = 0.9$$

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

$$434{,}900-400{,}000 = 34{,}900$$
$$500{,}000-400{,}000 = 100{,}000$$

$$1.0 + \frac{34{,}900}{100{,}000}(0.9-1.0) = 1.0$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 21.8 to 23.8 percent; the 90-percent confidence interval is from 21.2 to 24.4 percent; and the 95-percent confidence interval is from 20.8 to 24.8 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; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1977 and 1981 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1977 and 1981 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1981 there were 235,900 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 136,600. Table I shows the standard error of 99,300 is approximately 4,420, and the standard error of 235,900 is approximately 6,920. Therefore, the standard error of the estimated difference of 136,600 is about 8,210.

$$8{,}210 = \sqrt{(4{,}420)^2 + (6{,}920)^2}$$

Consequently, the 68-percent confidence interval for the 136,600 difference is from 128,390 to 144,810 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 123,460 to 149,740 housing units, and the 95-percent confidence interval is from 120,180 to 153,020 housing units. Thus, we can conclude with 95 percent confidence that the number of 1981 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied 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 to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.7. The base of the distribution from which this median was determined is 434,900 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 434,900 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 199,500 housing units or 45.9 percent fall below this interval, and 84,900 housing units or 19.5 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(47.8-45.9)}{19.5} = 2.6$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.2 percent derived in step 2. About 199,500 housing units or 45.9 percent fall below this interval, and 84,900 housing units or 19.5 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(52.2-45.9)}{19.5} = 2.8$$

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 1981 Housing Inventory and for Estimated Number of 1977-1981 Lost Units for the Dallas, Tex., SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	190	190	200	100,000	4,440	4,060	4,970
100	190	190	200	150,000	5,470	4,820	6,370
200	200	190	200	200,000	6,350	5,380	7,650
500	310	300	320	250,000	7,140	5,800	8,890
700	370	360	380	300,000	7,870	6,110	10,080
1,000	440	430	450	400,000	9,190	6,440	12,400
2,500	690	680	710	500,000	10,390	6,450	14,660
5,000	980	960	1,010	600,000	11,500	—	—
10,000	1,390	1,360	1,440	700,000	12,550	—	—
25,000	2,200	2,120	2,310	800,000	13,560	—	—
50,000	3,120	2,960	3,350	900,000	14,520	—	—
75,000	3,830	3,570	4,210				

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

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Dallas, Tex., 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	66.8	66.8	66.8	66.8	66.8	71.0	100,000	0.2	0.4	1.0	1.3	1.9	2.2
200	50.2	50.2	50.2	50.2	50.2	50.2	150,000	0.13	0.4	0.8	1.1	1.6	1.8
500	28.7	28.7	28.7	28.7	28.7	31.7	200,000	0.10	0.3	0.7	1.0	1.4	1.6
700	22.4	22.4	22.4	22.4	23.2	26.8	250,000	0.08	0.3	0.6	0.9	1.2	1.4
1,000	16.8	16.8	16.8	16.8	19.4	22.4	300,000	0.07	0.3	0.6	0.8	1.1	1.3
2,500	7.5	7.5	7.5	8.5	12.3	14.2	400,000	0.05	0.2	0.5	0.7	1.0	1.1
5,000	3.9	3.9	4.4	6.0	8.7	10.0	500,000	0.04	0.2	0.4	0.6	0.9	1.0
10,000	2.0	2.0	3.1	4.3	6.1	7.1	600,000	0.03	0.2	0.4	0.5	0.8	0.9
25,000	0.8	0.9	2.0	2.7	3.9	4.5	700,000	0.03	0.2	0.4	0.5	0.7	0.8
50,000	0.4	0.6	1.4	1.9	2.7	3.2	800,000	0.03	0.2	0.3	0.5	0.7	0.8
75,000	0.3	0.5	1.1	1.6	2.2	2.6	900,000	0.02	0.15	0.3	0.4	0.6	0.7

[1]Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.1.

Detroit 1981

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey . . . App-42
Designation of sample housing units for the 1981 survey . . . App-42
Selection of the original AHS-SMSA sample . . . App-43
1977-1981 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-45
1970 Census of Population and Housing . . . App-45
ESTIMATION . . . App-45
1981 housing inventory . . . App-45
1977-1981 lost housing units . . . App-47
1977 estimation procedure . . . App-47
Ratio estimation procedure of the 1970 Census of Population and Housing . . . App-47
RELIABILITY OF THE ESTIMATES . . . App-47
Nonsampling errors . . . App-47
1970 census . . . App-47
AHS-SMSA . . . App-47
Coverage errors . . . App-47
Rounding errors . . . App-48
Sampling errors for the AHS-SMSA sample . . . App-48
Illustration of the use of the standard error tables . . . App-49
Differences . . . App-49
Illustration of the computation of the standard error of a difference . . . App-50
Medians . . . App-50
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-170-81) are based on data collected from the 1981 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.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1981 were interviewed previously in 1974 and 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

The three largest SMSA's (Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va.) in the 1981 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1974 and 1977 surveys evenly divided between the central city and the balance of the respective SMSA. The 12 remaining SMSA's (Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Forth Worth, Tex.; Madison, Wis.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.) in the 1981 group were represented by a sample of 5,000 designated housing units in the 1974 and 1977 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1974 survey, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panel 3 was dropped from all SMSA's for the 1977 survey. Due to additional budget limitations for the 1981 survey, it also became necessary to drop panels 1, 2, 4, 5, 9, 11, and 12 from the Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va., SMSA's and panels 1 and 2 from the other 12 SMSA's. The designated sample size was reduced by two-thirds for the three largest SMSA's and by one-fourth for the other 12 SMSA's.

The interviewing was done for the three largest SMSA's during the period June 1981 through November 1981 with one-sixth of the sample housing units being visited each month. The interviewing was done for the remaining 12 SMSA's during the period April 1981 through December 1981 with one-ninth of the designated housing units visited each month.

In this SMSA, 4,995 housing units were eligible for interview. Of these sample housing units, 174 interviews were not obtained because, for occupied sample units, the occupants refused to be interviewed, were not at home after repeated visits, or were unavailable for some other reason; or, for vacant units, no informed respondent could be found after repeated visits. In addition to the 4,995 housing units eligible for interview, 294 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 1981 survey—The sample housing units designated to be interviewed in the 1981

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 and remained in sample after the 1981 panel reduction.
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 and remained in sample after the 1981 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1981 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 and remained in sample after the 1981 panel reduction. (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 and remained in sample after the 1981 panel reduction. (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 1981 Coverage Improvement Programs. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the original 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 five SMSA's were 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 10 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 for the 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.

1977-1981 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 from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes 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. Houses that have been moved onto their present site since the 1970 census.
6. Mobile homes placed outside parks since the 1970 census or vacant at the time of the 1970 census.

Each of the 15 SMSA's was previously interviewed in 1974 and 1977. The Coverage Improvement Program was conducted as part of the 1977 AHS with the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., SMSA's receiving some updating and refining as part of the 1981 AHS. The following discussion applies to both the prior year 1977 and the 1981 coverage improvement procedures. For the the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., 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 1981.

Coverage improvement for deficiency 1—A sample of new construction units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 15 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. 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 (regular AHS units), while sample units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage 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 Fort Worth, Tex., SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second-stage sampling. For the third stage, structures of size three or more units were divided into clusters of an expected size of two units and a sample of clusters was selected. This procedure was employed in the following SMSA's: Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Orlando, Fla.; Spokane, Wash.; and Wichita, Kans.

In the remaining 9 SMSA's, 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, this procedure was also used in parts of the Anaheim-Santa Ana Garden Grove, Calif., SMSA to supplement the sample described above. These procedures added an estimated 12,564 new construction 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 placed in a park that was 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 units represented the same number of units that the regular AHS sample units represented. This procedure added an estimated 8,625 units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiences 3-6—The remaining missed 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 units converted to residential use.
3. Mobile homes 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 a usual residence elsewhere.
4. Houses that had been moved onto their present site since the 1970 census.

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample 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 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 second procedure was designed to represent missed units from structures represented in the AHS. These missed units were:

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential 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 structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply both procedures during the 1977 Coverage Improvement Program to all SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in the Anaheim-Santa Ana-Garden Grove, Calif.; and Spokane, Wash., SMSA's. These SMSA's received the full application of the second procedure at that time. The Dallas, Tex.; Minneapolis-St. Paul, Minn.; and Pittsburgh, Pa., SMSA's did not receive application of either procedure during the 1977 Coverage Improvement Program. The Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and the Spokane, Wash., SMSA's had the first procedure completed as part of the 1981 Coverage Improvement Program. The remaining 10 1981 SMSA's received the full application of both procedures as part of the 1977 Coverage Improvement Program.

The first procedure added an estimated 8,361 housing units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 1,268 housing units.

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 1981 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 1981 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1977 (i.e., 1977-1981 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1981 housing inventory—The AHS estimates of characteristics of the 1981 housing inventory were produced using a one-stage ratio estimation procedure for the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; and Madison, Wis., SMSA's, a two-stage ratio estimation procedure for the Detroit, Mich.; Orlando, Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, and a three-stage ratio estimation procedure for the remaining seven SMSA's.

Prior to the implementation of the ratio estimation procedures, 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 nonpermit 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 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.

The next ratio estimation procedure was applied in the Boston, Mass.; Detroit, Mich.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's. This procedure was utilized 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 ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

The next ratio estimation procedure was applied in 10 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1981 housing inventory in each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1981 housing unit inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

The independent estimates of total housing units that were used as the numerator of this ratio are described below. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight.

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for each of the 15 1981 SMSA's. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1981 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units after the permit new construction ratio estimation (i.e., the existing sample estimate) were compared to the corresponding independent estimates for the central city, balance, and total SMSA for each of the 15 SMSA's and the estimate which showed the most likely level of net growth since the 1980 census in both the central city and balance as well as the total SMSA were used in this ratio estimation. As a result of this analysis, these independent estimates were used in the Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.- Md.-Va., SMSA's.

The remaining five SMSA's used a combination of these independent estimates and sample estimates. For the Boston, Mass.; and Minneapolis-St. Paul, Minn., SMSA's, the sample estimate was used for the total SMSA. For the sectors, the sample estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the independent estimates.

For the Orlando Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, the independent estimate of units was used for the total SMSA. For the sectors, the independent estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the sample estimates. Due to the central city

boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimate did not reflect the 1970 central city/balance SMSA definition. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definition were used.

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 in these 10 SMSA's, and the resulting product was used as the final weight for tabulation purposes. For the other five SMSA's, the sample estimates were used as the final weight for tabulation purposes. These SMSA's included Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Detroit, Mich.; Fort Worth, Tex.; and Madison, Wis., SMSA's.

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

1977-1981 lost housing units—The 1977-1981 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1977 housing inventory for 14 of the SMSA's and the corresponding two-stage ratio estimation procedure used for the Pittsburgh, Pa., SMSA as was described in the 1977 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1977-1981 lost housing units existed, by definition, in the 1977 housing inventory, there was a 1977 housing inventory weight associated with each 1977-1981 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1977-1981 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 two-stage ratio estimation process for the Pittsburgh, Pa., SMSA and a one-stage ratio estimation process for the remaining SMSA's. 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; 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 1981 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, and the results 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) 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, the 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 presented.

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 1981 survey, the number of missed housing units may be considerably less for 1981.

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 standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. Approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the average result of all possible samples.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimate would include the average result of all possible samples.

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

The figures presented in the tables that follow (pages App-51 through App-52) 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 1981 housing inventory as well as estimates of characteristics of the 1977-1981 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 54,360 for the total SMSA, 18,600 for the central city of the SMSA, and 50,960 for the balance of the SMSA.

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

Tables II through IV (pages App-51 through App-52) present the standard errros of estimated percentages for the 1981 housing inventory as well as estimated percentages of the 1977-1981 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 through 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 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 1981 there were 995,700 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 17,090. The following interpolation procedure was used.

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

Size of estimate	Standard error
900,000	16,230
995,700	x
1,000,000	17,130

The entry of "x" is determined as follows by vertically interpolating between 16,230 and 17,130.

$$995{,}700-900{,}000=95{,}700$$
$$1{,}000{,}000-900{,}000=100{,}000$$

$$16{,}230+\frac{95{,}700}{100{,}000}(17{,}130-16{,}230)=17{,}090$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 978,610 to 1,012,790 housing units. Therefore, a conclusion that the average estimate of 1981 owner-occupied housing units (derived from all possible samples) 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 968,360 to 1,023,040 housing units with 90 percent confidence; and that the average estimate lies within the interval from 961,520 to 1,029,880 housing units with 95 percent confidence.

Table A-1 part A also shows that of the 995,700 owner-occupied housing units, 234,600, or 23.6 percent, had two bedrooms. Interpolation using table II of this appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 23.6 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	23.6	25 or 75
900,000	0.5	a	0.8
995,700		p	
1,000,000	0.5	b	0.7

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

$$23.6-10.0=13.6$$
$$25.0-10.0=15.0$$

$$0.5+\frac{13.6}{15.0}(0.8-0.5)=0.8$$

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

$$23.6-10.0=13.6$$
$$25.0-10.0=15.0$$

$$0.5+\frac{13.6}{15.0}(0.7-0.5)=0.7$$

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

$$995{,}700-900{,}000=95{,}700$$
$$1{,}000{,}000-900{,}000=100{,}000$$

$$0.8+\frac{95{,}700}{100{,}000}(0.7-0.8)=0.7$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 22.9 to 24.3 percent; the 90-percent confidence interval is from 22.5 to 24.7 percent; and the 95-percent confidence interval is from 22.2 to 25.0 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The

standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. If there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1977 and 1981 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1977 and 1981 characteristics. . .

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1981 there were 557,500 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 322,900. Table I shows the standard error of 234,600 is approximately 8,220, and the standard error of 557,500 is approximately 12,720. Therefore, the standard error of the estimated difference of 322,900 is about 15,140.

$$15{,}140 = \sqrt{(8{,}220)^2 + (12{,}720)^2}$$

Consequently, the 68-percent confidence interval for the 322,900 difference is from 307,760 to 338,040 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 298,680 to 347,120 housing units, and the 95-percent confidence interval is from 292,620 to 353,180 housing units. Thus, we can conclude with 95 percent confidence that the number of 1981 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied 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 through 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 into which interval of the distribution the lower percentage limit falls. Similarly, to find the upper endpoint of the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.8. The base of the distribution from which this median was determined is 995,700 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 995,700 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 435,000 housing units or 43.7 percent fall below this interval, and 185,000 housing units or 18.6 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(48.4-43.7)}{18.6} = 2.8$$

Similarly, the interval for owner-occupied housing units with three persons contains the 51.6 percent derived in step 2. About 435,000 housing units or 43.7 percent fall below this interval, and 185,000 housing units or 18.6 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(51.6-43.7)}{18.6} = 2.9$$

Thus, the 95-percent confidence interval ranges from 2.8 to 2.9 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1981 Housing Inventory and for Estimated Number of 1977-1981 Lost Units for the Detroit, 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	290	220	380	250,000	8,500	5,590	9,720
100	290	220	380	300,000	9,310	5,660	10,660
200	290	220	380	400,000	10,760	5,310	12,320
500	380	330	430	500,000	12,050	4,140	13,790
700	450	390	510	600,000	13,210	—	15,130
1,000	540	460	610	700,000	14,280	—	16,360
2,500	850	730	970	800,000	15,290	—	17,510
5,000	1,200	1,030	1,370	900,000	16,230	—	18,590
10,000	1,700	1,460	1,940	1,000,000	17,130	—	19,620
25,000	2,680	2,270	3,070	1,100,000	17,980	—	20,600
50,000	3,790	3,140	4,340	1,200,000	18,800	—	—
75,000	4,650	3,760	5,320	1,300,000	19,590	—	—
100,000	5,370	4,230	6,140	1,400,000	20,350	—	—
150,000	6,570	4,920	7,520	1,500,000	21,080	—	—
200,000	7,600	5,350	8,690				

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

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Detroit, Mich., 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	74.2	74.2	74.2	74.2	74.2	84.7	250,000	0.11	0.3	0.7	1.0	1.5	1.7
200	59.0	59.0	59.0	59.0	59.0	59.9	300,000	0.10	0.3	0.7	0.9	1.3	1.5
500	36.5	36.5	36.5	36.5	36.5	37.9	400,000	0.07	0.3	0.6	0.8	1.2	1.3
700	29.1	29.1	29.1	29.1	29.1	32.0	500,000	0.06	0.2	0.5	0.7	1.0	1.2
1,000	22.3	22.3	22.3	22.3	23.2	26.8	600,000	0.05	0.2	0.5	0.7	0.9	1.1
2,500	10.3	10.3	10.3	10.3	14.7	16.9	700,000	0.04	0.2	0.4	0.6	0.9	1.0
5,000	5.4	5.4	5.4	7.2	10.4	12.0	800,000	0.04	0.2	0.4	0.6	0.8	0.9
10,000	2.8	2.8	3.7	5.1	7.3	8.5	900,000	0.03	0.2	0.4	0.5	0.8	0.9
25,000	1.1	1.1	2.3	3.2	4.6	5.4	1,000,000	0.03	0.2	0.4	0.5	0.7	0.8
50,000	0.6	0.8	1.7	2.3	3.3	3.8	1,100,000	0.03	0.2	0.4	0.5	0.7	0.8
75,000	0.4	0.6	1.3	1.9	2.7	3.1	1,200,000	0.02	0.2	0.3	0.5	0.7	0.8
100,000	0.3	0.5	1.2	1.6	2.3	2.7	1,300,000	0.02	0.15	0.3	0.4	0.6	0.7
150,000	0.2	0.4	1.0	1.3	1.9	2.2	1,400,000	0.02	0.14	0.3	0.4	0.6	0.7
200,000	0.14	0.4	0.8	1.1	1.6	1.9	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 1.4.

TABLE III. Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Central City of the Detroit, Mich., 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	68.3	68.3	68.3	68.3	68.3	73.4	50,000	0.4	0.7	1.4	2.0	2.8	3.3
200	51.9	51.9	51.9	51.9	51.9	51.9	75,000	0.3	0.5	1.2	1.6	2.3	2.7
500	30.1	30.1	30.1	30.1	30.1	32.8	100,000	0.2	0.5	1.0	1.4	2.0	2.3
700	23.5	23.5	23.5	23.5	24.0	27.7	150,000	0.14	0.4	0.8	1.1	1.6	1.9
1,000	17.7	17.7	17.7	17.7	20.1	23.2	200,000	0.11	0.3	0.7	1.0	1.4	1.6
2,500	7.9	7.9	7.9	8.8	12.7	14.7	250,000	0.09	0.3	0.6	0.9	1.3	1.5
5,000	4.1	4.1	4.5	6.2	9.0	10.4	300,000	0.07	0.3	0.6	0.8	1.2	1.3
10,000	2.1	2.1	3.2	4.4	6.4	7.3	400,000	0.05	0.2	0.5	0.7	1.0	1.2
25,000	0.9	0.9	2.0	2.8	4.0	4.6	500,000	0.04	0.2	0.5	0.6	0.9	1.0

[1]Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2.

TABLE IV. Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Balance (Not in Central City) of the Detroit, Mich., 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	79.0	79.0	79.0	79.0	84.0	97.0	150,000	0.3	0.5	1.1	1.5	2.2	2.5
200	65.3	65.3	65.3	65.3	65.3	68.6	200,000	0.2	0.4	0.9	1.3	1.9	2.2
500	42.9	42.9	42.9	42.9	42.9	43.4	250,000	0.2	0.4	0.8	1.2	1.7	1.9
700	34.9	34.9	34.9	34.9	34.9	36.6	300,000	0.13	0.4	0.8	1.1	1.5	1.8
1,000	27.3	27.3	27.3	27.3	27.3	30.7	400,000	0.09	0.3	0.7	0.9	1.3	1.5
2,500	13.1	13.1	13.1	13.1	16.8	19.4	500,000	0.08	0.3	0.6	0.8	1.2	1.4
5,000	7.0	7.0	7.0	8.2	11.9	13.7	600,000	0.06	0.2	0.5	0.8	1.1	1.3
10,000	3.6	3.6	4.2	5.8	8.4	9.7	700,000	0.05	0.2	0.5	0.7	1.0	1.2
25,000	1.5	1.5	2.7	3.7	5.3	6.1	800,000	0.05	0.2	0.5	0.7	0.9	1.1
50,000	0.7	0.9	1.9	2.6	3.8	4.3	900,000	0.04	0.2	0.4	0.6	0.9	1.0
75,000	0.5	0.7	1.5	2.1	3.1	3.5	1,000,000	0.04	0.2	0.4	0.6	0.8	1.0
100,000	0.4	0.6	1.3	1.8	2.7	3.1	1,100,000	0.03	0.2	0.4	0.6	0.8	0.9

[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 1.1.

Ft. Worth 1981

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1981 survey App-42
Selection of the original AHS-SMSA sample App-43
1977-1981 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-45
1970 Census of Population and Housing App-45
ESTIMATION App-45
1981 housing inventory App-45
1977-1981 lost housing units ... App-47
1977 estimation procedure App-47
Ratio estimation procedure of the 1970 Census of Population and Housing App-47
RELIABILITY OF THE ESTIMATES .. App-47
Nonsampling errors App-47
1970 census App-47
AHS-SMSA App-47
Coverage errors App-47
Rounding errors App-48
Sampling errors for the AHS-SMSA sample App-48
Illustration of the use of the standard error tables App-49
Differences App-49
Illustration of the computation of the standard error of a difference App-50
Medians App-50
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-170-81) are based on data collected from the 1981 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.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1981 were interviewed previously in 1974 (1975 for Madison, Wis.) and 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

The three largest SMSA's (Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va.) in the 1981 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1974 and 1977 surveys evenly divided between the central city and the balance of the respective SMSA. The 12 remaining SMSA's (Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.) in the 1981 group were represented by a sample of 5,000 designated housing units in the 1974 (1975 for Madison, Wis.) and 1977 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1974 survey, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panel 3 was dropped from all SMSA's for the 1977 survey. Due to additional budget limitations for the 1981 survey, it also became necessary to drop panels 1, 2, 4, 5, 9, 11, and 12 from the Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va., SMSA's and panels 1 and 2 from the other 12 SMSA's. The designated sample size was reduced by two-thirds for the three largest SMSA's and by one-fourth for the other 12 SMSA's.

The interviewing was done for the three largest SMSA's during the period June 1981 through November 1981 with one-sixth of the sample housing units being visited each month. The interviewing was done for the remaining 12 SMSA's during the period April 1981 through December 1981 with one-ninth of the designated housing units visited each month.

In this SMSA, 4,780 housing units were eligible for interview. Of these sample housing units, 182 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 4,780 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 1981 survey—The sample housing units designated to be interviewed in the 1981

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 and remained in sample after the 1981 panel reduction.
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 and remained in sample after the 1981 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1981 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 and remained in sample after the 1981 panel reduction. (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 and remained in sample after the 1981 panel reduction. (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 1981 Coverage Improvement Programs. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the original 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 five SMSA's were 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 10 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 for the 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.

1977-1981 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 from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes 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. Houses that have been moved onto their present site since the 1970 census.
6. Mobile homes placed outside parks since the 1970 census or vacant at the time of the 1970 census.

Each of the 15 SMSA's was previously interviewed in 1974 (1975 for Madison, Wis.) and 1977. The Coverage Improvement Program was conducted as part of the 1977 AHS with the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., SMSA's receiving some updating and refining as part of the 1981 AHS. The following discussion applies to both the prior year 1977 and the 1981 coverage improvement procedures. For the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., 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 1981.

Coverage improvement for deficiency 1—A sample of new construction units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 15 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. 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 (regular AHS units), while sample units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage 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 Fort Worth, Tex., SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second-stage sampling. For the third stage, structures of size three or more units were divided into clusters of an expected size of two units and a sample of clusters was selected. This procedure was employed in the following SMSA's: Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Orlando, Fla.; Spokane, Wash.; and Wichita, Kans.

In the remaining 9 SMSA's, 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, this procedure was also used in parts of the Anaheim-Santa Ana Garden Grove, Calif., SMSA to supplement the sample described above. These procedures added an estimated 5,060 new construction 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 placed in a park that was 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 units represented the same number of units that the regular AHS sample units represented. This procedure added an estimated 2,503 units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed 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 units converted to residential use.
3. Mobile homes 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 a usual residence elsewhere.
4. Houses that had been moved onto their present site since the 1970 census.

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample 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 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 second procedure was designed to represent missed units from structures represented in the AHS. These missed units were:

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential 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 structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply both procedures during the 1977 Coverage Improvement Program to all SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in the Anaheim-Santa Ana-Garden Grove, Calif.; and Spokane, Wash., SMSA's. These SMSA's did not receive application of the second procedure. The Dallas, Tex.; Minneapolis-St. Paul, Minn.; and Pittsburgh, Pa., SMSA's did not receive application of either procedure during the 1977 Coverage Improvement Program. The Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and the Spokane, Wash., SMSA's had the first procedure completed as part of the 1981 Coverage Improvement Program. The remaining 10 1981 SMSA's received the full application of both procedures as part of the 1977 Coverage Improvement Program.

The first procedure added an estimated 9,538 housing units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 299 housing units.

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 1981 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 1981 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1977 (i.e., 1977-1981 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1981 housing inventory—The AHS estimates of characteristics of the 1981 housing inventory were produced using a one-stage ratio estimation procedure for the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; and Madison, Wis., SMSA's, a two-stage ratio estimation procedure for the Detroit, Mich.; Orlando, Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, and a three-stage ratio estimation procedure for the remaining seven SMSA's.

Prior to the implementation of the ratio estimation procedures, 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 nonpermit 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 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.

The next ratio estimation procedure was applied in the Boston, Mass.; Detroit, Mich.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's. This procedure was utilized 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 ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

The next ratio estimation procedure was applied in 10 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1981 housing inventory in each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1981 housing unit inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

The independent estimates of total housing units that were used as the numerator of this ratio are described below. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight.

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for each of the 15 1981 SMSA's. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1981 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units after the permit new construction ratio estimation (i.e., the existing sample estimate) were compared to the corresponding independent estimates for the central city, balance, and total SMSA for each of the 15 SMSA's and the estimate which showed the most likely level of net growth since the 1980 census in both the central city and balance as well as the total SMSA were used in this ratio estimation. As a result of this analysis, these independent estimates were used in the Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's.

The remaining five SMSA's used a combination of these independent estimates and sample estimates. For the Boston, Mass.; and Minneapolis-St. Paul, Minn., SMSA's, the sample estimate was used for the total SMSA. For the sectors, the sample estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the independent estimates.

For the Orlando Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, the independent estimate of units was used for the total SMSA. For the sectors, the independent estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the sample estimates. Due to the central city

boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimate did not reflect the 1970 central city/balance SMSA definition. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definition were used.

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 in these 10 SMSA's, and the resulting product was used as the final weight for tabulation purposes. For the other five SMSA's, the sample estimates were used as the final weight for tabulation purposes. These SMSA's included Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Detroit, Mich.; Fort Worth, Tex.; and Madison, Wis., SMSA's.

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

1977-1981 lost housing units—The 1977-1981 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1977 housing inventory for 14 of the SMSA's and the corresponding two-stage ratio estimation procedure used for the Pittsburgh, Pa., SMSA as was described in the 1977 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1977-1981 lost housing units existed, by definition, in the 1977 housing inventory, there was a 1977 housing inventory weight associated with each 1977-1981 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1977-1981 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 two-stage ratio estimation process for the Pittsburgh, Pa., SMSA and a one-stage ratio estimation process for the remaining SMSA's. 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; 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 1981 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, and the results 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) 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, the 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 presented.

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 1981 survey, the number of missed housing units may be considerably less for 1981.

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 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 1981 housing inventory as well as estimates of characteristics of the 1977-1981 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 18,943 for the total SMSA, 7,926 for the central city of the SMSA, and 17,205 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 1981 housing inventory as well as estimated percentages of the 1977-1981 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, tables II through IV 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 1981 there were 232,300 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 4,920. 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	4,510
232,300	x
250,000	5,150

The entry of "x" is determined as follows by vertically interpolating between 4,510 and 5,150.

$$232{,}300 - 200{,}000 = 32{,}300$$
$$250{,}000 - 200{,}000 = 50{,}000$$

$$4{,}510 + \frac{32{,}300}{50{,}000}(5{,}150 - 4{,}510) = 4{,}920$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 227,380 to 237,220 housing units. Therefore, a conclusion that the average estimate of 1981 owner-occupied housing units (derived from all possible samples) 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 224,430 to 240,170 housing units with 90 percent confidence; and that the average estimate lies within the interval from 222,460 to 242,140 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 232,300 owner-occupied housing units, 63,200, or 27.2 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 27.2 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	27.2	50
200,000	0.9	a	1.1
232,300		p	
250,000	0.8	b	1.0

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

$$27.2 - 25.0 = 2.2$$
$$50.0 - 25.0 = 25.0$$

$$0.9 + \frac{2.2}{25.0}(1.1 - 0.9) = 0.9$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.8 and 1.0.

$$27.2 - 25.0 = 2.2$$
$$50.0 - 25.0 = 25.0$$

$$0.8 + \frac{2.2}{25.0}(1.0 - 0.8) = 0.8$$

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

$$232{,}300 - 200{,}000 = 32{,}300$$
$$250{,}000 - 200{,}000 = 50{,}000$$

$$0.9 + \frac{32{,}300}{50{,}000}(0.8 - 0.9) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 26.4 to 28.0 percent; the 90-percent confidence interval is from 25.9 to 28.5 percent; and the 95-percent confidence interval is from 25.6 to 28.8 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; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1977 and 1981 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1977 and 1981 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1981 there were 131,200 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 68,000. Table I shows the standard error of 63,200 is approximately 2,370, and the standard error of 131,200 is approximately 3,530. Therefore, the standard error of the estimated difference of 68,000 is about 4,250.

$$4{,}250 = \sqrt{(2{,}370)^2 + (3{,}530)^2}$$

Consequently, the 68-percent confidence interval for the 68,000 difference is from 63,750 to 72,250 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 61,200 to 74,800 housing units, and the 95-percent confidence interval is from 59,500 to 76,500 housing units. Thus, we can conclude with 95 percent confidence that the number of 1981 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied 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 to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.7. The base of the distribution from which this median was determined is 232,300 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 232,300 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 104,100 housing units or 44.8 percent fall below this interval, and 52,100 housing units or 22.4 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(48.0-44.8)}{22.4} = 2.6$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.0 percent derived in step 2. About 104,100 housing units or 44.8 percent fall below this interval, and 52,100 housing units or 22.4 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(52.0-44.8)}{22.4} = 2.8$$

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 1981 Housing Inventory and for Estimated Number of 1977-1981 Lost Units for the Fort Worth, Tex., SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	80	80	100	25,000	1,470	1,360	1,620
100	90	90	100	50,000	2,100	1,850	2,400
200	130	130	140	75,000	2,610	2,190	3,050
500	200	200	220	100,000	3,050	2,420	3,660
700	240	230	260	150,000	3,820	2,700	4,780
1,000	290	280	310	200,000	4,510	2,780	5,850
2,500	460	440	490	250,000	5,150	—	6,890
5,000	650	620	700	300,000	5,760	—	—
10,000	920	880	1,000	400,000	6,910	—	—

[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 for the balance (not in central city) estimates. For estimates pertaining to lost units, the standard errors should be multiplied by a factor of 1.1.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Fort Worth, Tex., 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	49.0	49.0	49.0	49.0	49.0	49.0	50,000	0.2	0.4	1.0	1.3	1.9	2.2
200	32.4	32.4	32.4	32.4	32.4	34.6	75,000	0.13	0.4	0.8	1.1	1.5	1.8
500	16.1	16.1	16.1	16.1	19.0	21.9	100,000	0.10	0.3	0.7	0.9	1.3	1.5
700	12.1	12.1	12.1	12.1	16.0	18.5	150,000	0.06	0.3	0.6	0.8	1.1	1.3
1,000	8.8	8.8	8.8	9.3	13.4	15.5	200,000	0.05	0.2	0.5	0.7	0.9	1.1
2,500	3.7	3.7	4.3	5.9	8.5	9.8	250,000	0.04	0.2	0.4	0.6	0.8	1.0
5,000	1.9	1.9	3.0	4.2	6.0	6.9	300,000	0.03	0.2	0.4	0.5	0.8	0.9
10,000	1.0	1.0	2.1	2.9	4.2	4.9	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							

[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 for the balance (not in central city). For estimates pertaining to lost units, the standard errors should be multiplied by a factor of 1.1.

Madison 1981

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1981 survey App-42
Selection of the original AHS-SMSA sample App-43
1977-1981 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-45
1970 Census of Population and Housing App-45
ESTIMATION App-45
1981 housing inventory App-45
1977-1981 lost housing units ... App-47
1977 estimation procedure App-47
Ratio estimation procedure of the 1970 Census of Population and Housing App-47
RELIABILITY OF THE ESTIMATES .. App-47
Nonsampling errors App-47
1970 census App-47
AHS-SMSA App-47
Coverage errors App-47
Rounding errors App-48
Sampling errors for the AHS-SMSA sample App-48
Illustration of the use of the standard error tables App-49
Differences App-49
Illustration of the computation of the standard error of a difference App-50
Medians App-50
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-170-81) are based on data collected from the 1981 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.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1981 were interviewed previously in 1974 (1975 for Madison, Wis.) and 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

The three largest SMSA's (Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va.) in the 1981 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1974 and 1977 surveys evenly divided between the central city and the balance of the respective SMSA. The 12 remaining SMSA's (Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Forth Worth, Tex.; Madison, Wis.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.) in the 1981 group were represented by a sample of 5,000 designated housing units in the 1974 (1975 for Madison, Wis.) and 1977 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1975 survey, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panel 3 was dropped from all SMSA's for the 1977 survey. Due to additional budget limitations for the 1981 survey, it also became necessary to drop panels 1, 2, 4, 5, 9, 11, and 12 from the Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va., SMSA's and panels 1 and 2 from the other 12 SMSA's. The designated sample size was reduced by two-thirds for the three largest SMSA's and by one-fourth for the other 12 SMSA's.

The interviewing was done for the three largest SMSA's during the period June 1981 through November 1981 with one-sixth of the sample housing units being visited each month. The interviewing was done for the remaining 12 SMSA's during the period April 1981 through December 1981 with one-ninth of the designated housing units visited each month.

In this SMSA, 4,287 housing units were eligible for interview. Of these sample housing units, 65 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 4,287 housing units eligible for interview, 312 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 1981 survey—The sample housing units designated to be interviewed in the 1981

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 and remained in sample after the 1981 panel reduction.
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 and remained in sample after the 1981 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1981 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 and remained in sample after the 1981 panel reduction. (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 and remained in sample after the 1981 panel reduction. (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 1981 Coverage Improvement Programs. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the original 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 five SMSA's were 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 10 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 for the 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.

1977-1981 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 from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes 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. Houses that have been moved onto their present site since the 1970 census.
6. Mobile homes placed outside parks since the 1970 census or vacant at the time of the 1970 census.

Each of the 15 SMSA's was previously interviewed in 1974 (1975 for Madison, Wis.) and 1977. The Coverage Improvement Program was conducted as part of the 1977 AHS with the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., SMSA's receiving some updating and refining as part of the 1981 AHS. The following discussion applies to both the prior year 1977 and the 1981 coverage improvement procedures. For the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., 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 1981.

Coverage improvement for deficiency 1—A sample of new construction units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 15 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. 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 (regular AHS units), while sample units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage 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 Fort Worth, Tex., SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second-stage sampling. For the third stage, structures of size three or more units were divided into clusters of an expected size of two units and a sample of clusters was selected. This procedure was employed in the following SMSA's: Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Orlando, Fla.; Spokane, Wash.; and Wichita, Kans.

In the remaining 9 SMSA's, 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, this procedure was also used in parts of the Anaheim-Santa Ana Garden Grove, Calif., SMSA to supplement the sample described above. These procedures added an estimated 903 new construction 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 placed in a park that was 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 units represented the same number of units that the regular AHS sample units represented. This procedure added an estimated 134 units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiences 3-6—The remaining missed 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 units converted to residential use.
3. Mobile homes 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 a usual residence elsewhere.
4. Houses that had been moved onto their present site since the 1970 census.

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample 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 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 second procedure was designed to represent missed units from structures represented in the AHS. These missed units were:

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential 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 structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply both procedures during the 1977 Coverage Improvement Program to all SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in the Anaheim-Santa Ana-Garden Grove, Calif.; and Spokane, Wash., SMSA's. These SMSA's did not receive application of the second procedure. The Dallas, Tex.; Minneapolis-St. Paul, Minn.; and Pittsburgh, Pa., SMSA's did not receive application of either procedure during the 1977 Coverage Improvement Program. The Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and the Spokane, Wash., SMSA's had the first procedure completed as part of the 1981 Coverage Improvement Program. The remaining 10 1981 SMSA's received the full application of both procedures as part of the 1977 Coverage Improvement Program.

The first procedure added an estimated 1,682 housing units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 45 housing units.

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 1981 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 1981 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1977 (i.e., 1977-1981 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1981 housing inventory—The AHS estimates of characteristics of the 1981 housing inventory were produced using a one-stage ratio estimation procedure for the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; and Madison, Wis., SMSA's, a two-stage ratio estimation procedure for the Detroit, Mich.; Orlando, Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, and a three-stage ratio estimation procedure for the remaining seven SMSA's.

Prior to the implementation of the ratio estimation procedures, 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 nonpermit 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 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.

The next ratio estimation procedure was applied in the Boston, Mass.; Detroit, Mich.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's. This procedure was utilized 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 ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

The next ratio estimation procedure was applied in 10 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1981 housing inventory in each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1981 housing unit inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

The independent estimates of total housing units that were used as the numerator of this ratio are described below. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight.

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for each of the 15 1981 SMSA's. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1981 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units after the permit new construction ratio estimation (i.e., the existing sample estimate) were compared to the corresponding independent estimates for the central city, balance, and total SMSA for each of the 15 SMSA's and the estimate which showed the most likely level of net growth since the 1980 census in both the central city and balance as well as the total SMSA were used in this ratio estimation. As a result of this analysis, these independent estimates were used in the Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's.

The remaining five SMSA's used a combination of these independent estimates and sample estimates. For the Boston, Mass.; and Minneapolis-St. Paul, Minn., SMSA's, the sample estimate was used for the total SMSA. For the sectors, the sample estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the independent estimates.

For the Orlando Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, the independent estimate of units was used for the total SMSA. For the sectors, the independent estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the sample estimates. Due to the central city

boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimate did not reflect the 1970 central city/balance SMSA definition. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definition were used.

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 in these 10 SMSA's, and the resulting product was used as the final weight for tabulation purposes. For the other five SMSA's, the sample estimates were used as the final weight for tabulation purposes. These SMSA's included Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Detroit, Mich.; Fort Worth, Tex.; and Madison, Wis., SMSA's.

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

1977-1981 lost housing units—The 1977-1981 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1977 housing inventory for 14 of the SMSA's and the corresponding two-stage ratio estimation procedure used for the Pittsburgh, Pa., SMSA as was described in the 1977 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas*. Since the 1977-1981 lost housing units existed, by definition, in the 1977 housing inventory, there was a 1977 housing inventory weight associated with each 1977-1981 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1977-1981 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 two-stage ratio estimation process for the Pittsburgh, Pa., SMSA and a one-stage ratio estimation process for the remaining SMSA's. 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; 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 1981 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, and the results 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) 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, the 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 presented.

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 1981 survey, the number of missed housing units may be considerably less for 1981.

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 standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. Approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the average result of all possible samples.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimate would include the average result of all possible samples.

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

The figures presented in the tables that follow (pages App-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 1981 housing inventory as well as estimates of characteristics of the 1977-1981 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 4,676 for the total SMSA, 2,698 for the central city of the SMSA, and 3,737 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 1981 housing inventory as well as estimated percentages of the 1977-1981 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, tables II through IV 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 1981 there were 68,800 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 1,230. 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
50,000	1,100
68,800	x
75,000	1,270

The entry of "x" is determined as follows by vertically interpolating between 1,100 and 1,270.

$$68{,}800 - 50{,}000 = 18{,}800$$
$$75{,}000 - 50{,}000 = 25{,}000$$

$$1{,}100 + \frac{18{,}800}{25{,}000}\ (1{,}270 - 1{,}100) = 1{,}230$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 67,570 to 70,030 housing units. Therefore, a conclusion that the average estimate of 1981 owner-occupied housing units (derived from all possible samples) 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 66,830 to 70,770 housing units with 90 percent confidence; and that the average estimate lies within the interval from 66,340 to 71,260 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 68,800 owner-occupied housing units, 13,400, or 19.5 percent, had two bedrooms. Interpolation using table II of this appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 19.5 percent is approximately 0.8 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	19.5	25 or 75
50,000	0.7	a	1.0
68,800		p	
75,000	0.6	b	0.9

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

$$19.5 - 10.0 = 9.5$$
$$25.0 - 10.0 = 15.0$$

$$0.7 + \frac{9.5}{15.0}(1.0 - 0.7) = 0.9$$

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

$$19.5 - 10.0 = 9.5$$
$$25.0 - 10.0 = 15.0$$

$$0.6 + \frac{9.5}{15.0}(0.9 - 0.6) = 0.8$$

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

$$68{,}800 - 50{,}000 = 18{,}800$$
$$75{,}000 - 50{,}000 = 25{,}000$$

$$0.9 + \frac{18{,}800}{25{,}000}\ (0.8 - 0.9) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 18.7 to 20.3 percent; the 90-percent confidence interval is from 18.2 to 20.8 percent; and the 95-percent confidence interval is from 17.9 to 21.1 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The

standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. If there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1977 and 1981 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1977 and 1981 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1981 there were 38,600 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 25,200. Table I shows the standard error of 13,400 is approximately 590, and the standard error of 38,600 is approximately 970. Therefore, the standard error of the estimated difference of 25,200 is about 1,140.

$$1{,}140 = \sqrt{(590)^2 \quad (970)^2}$$

Consequently, the 68-percent confidence interval for the 25,200 difference is from 24,060 to 26,340 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 23,380 to 27,020 housing units, and the 95-percent confidence interval is from 22,920 to 27,480 housing units. Thus, we can conclude with 95 percent confidence that the number of 1981 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied 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 to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.8. The base of the distribution from which this median was determined is 68,800 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 68,800 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 31,000 housing units or 45.1 percent fall below this interval, and 13,100 housing units or 19.0 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5) \quad \frac{(48.0-45.1)}{19.0} = 2.7$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.0 percent derived in step 2. About 31,000 housing units or 45.1 percent fall below this interval, and 13,100 housing units or 19.0 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5) \quad \frac{(52.0-45.1)}{19.0} = 2.9$$

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

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1981 Housing Inventory and for Estimated Number of 1977-1981 Lost Units for the Madison, Wis., SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]	Size of estimate	Standard error[1]
0	30	5,000	380
100	50	10,000	530
200	80	25,000	810
500	120	50,000	1,100
700	140	75,000	1,270
1,000	170	100,000	1,380
2,500	270	150,000	1,460

[1]For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Madison, Wis., SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						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	22.5	22.5	22.5	22.5	23.3	26.9	10,000	0.3	0.5	1.2	1.6	2.3	2.7
200	12.7	12.7	12.7	12.7	16.5	19.0	25,000	0.12	0.3	0.7	1.0	1.5	1.7
500	5.5	5.5	5.5	7.2	10.4	12.0	50,000	0.06	0.2	0.5	0.7	1.0	1.2
700	4.0	4.0	4.4	6.1	8.8	10.2	75,000	0.04	0.2	0.4	0.6	0.9	1.0
1,000	2.8	2.8	3.7	5.1	7.4	8.5	100,000	0.03	0.2	0.4	0.5	0.7	0.9
2,500	1.1	1.1	2.3	3.2	4.7	5.4	150,000	0.02	0.14	0.3	0.4	0.6	0.7
5,000	0.6	0.8	1.7	2.3	3.3	3.8							

[1]Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2.

Minneapolis-St. Paul 1981

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1981 survey App-42
Selection of the original AHS-SMSA sample App-43
1977-1981 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-45
1970 Census of Population and Housing App-45
ESTIMATION App-45
1981 housing inventory App-45
1977-1981 lost housing units ... App-47
1977 estimation procedure App-47
Ratio estimation procedure of the 1970 Census of Population and Housing App-47
RELIABILITY OF THE ESTIMATES .. App-47
Nonsampling errors App-47
1970 census App-47
AHS-SMSA App-47
Coverage errors App-47
Rounding errors App-48
Sampling errors for the AHS-SMSA sample App-48
Illustration of the use of the standard error tables App-49
Differences App-49
Illustration of the computation of the standard error of a difference App-50
Medians App-50
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-170-81) are based on data collected from the 1981 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.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1981 were interviewed previously in 1974 (1975 for Madison, Wis.) and 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

The three largest SMSA's (Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va.) in the 1981 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1974 and 1977 surveys evenly divided between the central city and the balance of the respective SMSA. The 12 remaining SMSA's (Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Forth Worth, Tex.; Madison, Wis.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.) in the 1981 group were represented by a sample of 5,000 designated housing units in the 1974 (1975 for Madison, Wis.) and 1977 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1975 survey, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panel 3 was dropped from all SMSA's for the 1977 survey. Due to additional budget limitations for the 1981 survey, it also became necessary to drop panels 1, 2, 4, 5, 9, 11, and 12 from the Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va., SMSA's and panels 1 and 2 from the other 12 SMSA's. The designated sample size was reduced by two-thirds for the three largest SMSA's and by one-fourth for the other 12 SMSA's.

The interviewing was done for the three largest SMSA's during the period June 1981 through November 1981 with one-sixth of the sample housing units being visited each month. The interviewing was done for the remaining 12 SMSA's during the period April 1981 through December 1981 with one-ninth of the designated housing units visited each month.

In this SMSA, 4,287 housing units were eligible for interview. Of these sample housing units, 65 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 4,287 housing units eligible for interview, 312 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 1981 survey—The sample housing units designated to be interviewed in the 1981

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 and remained in sample after the 1981 panel reduction.
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 and remained in sample after the 1981 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1981 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 and remained in sample after the 1981 panel reduction. (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 and remained in sample after the 1981 panel reduction. (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 1981 Coverage Improvement Programs. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the original 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 five SMSA's were 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 10 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 for the 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.

1977-1981 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 from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes 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. Houses that have been moved onto their present site since the 1970 census.
6. Mobile homes placed outside parks since the 1970 census or vacant at the time of the 1970 census.

Each of the 15 SMSA's was previously interviewed in 1974 (1975 for Madison, Wis.) and 1977. The Coverage Improvement Program was conducted as part of the 1977 AHS with the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., SMSA's receiving some updating and refining as part of the 1981 AHS. The following discussion applies to both the prior year 1977 and the 1981 coverage improvement procedures. For the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., 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 1981.

Coverage improvement for deficiency 1—A sample of new construction units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 15 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. 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 (regular AHS units), while sample units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage 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 Fort Worth, Tex., SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second-stage sampling. For the third stage, structures of size three or more units were divided into clusters of an expected size of two units and a sample of clusters was selected. This procedure was employed in the following SMSA's: Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Orlando, Fla.; Spokane, Wash.; and Wichita, Kans.

In the remaining 9 SMSA's, 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, this procedure was also used in parts of the Anaheim-Santa Ana Garden Grove, Calif., SMSA to supplement the sample described above. These procedures added an estimated 903 new construction 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 placed in a park that was 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 units represented the same number of units that the regular AHS sample units represented. This procedure added an estimated 134 units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiences 3-6—The remaining missed 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 units converted to residential use.
3. Mobile homes 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 a usual residence elsewhere.
4. Houses that had been moved onto their present site since the 1970 census.

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample 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 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 second procedure was designed to represent missed units from structures represented in the AHS. These missed units were:

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential 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 structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply both procedures during the 1977 Coverage Improvement Program to all SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in the Anaheim-Santa Ana-Garden Grove, Calif.; and Spokane, Wash., SMSA's. These SMSA's did not receive application of the second procedure. The Dallas, Tex.; Minneapolis-St. Paul, Minn.; and Pittsburgh, Pa., SMSA's did not receive application of either procedure during the 1977 Coverage Improvement Program. The Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and the Spokane, Wash., SMSA's had the first procedure completed as part of the 1981 Coverage Improvement Program. The remaining 10 1981 SMSA's received the full application of both procedures as part of the 1977 Coverage Improvement Program.

The first procedure added an estimated 1,682 housing units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 45 housing units.

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 1981 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 1981 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1977 (i.e., 1977-1981 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1981 housing inventory—The AHS estimates of characteristics of the 1981 housing inventory were produced using a one-stage ratio estimation procedure for the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; and Madison, Wis., SMSA's, a two-stage ratio estimation procedure for the Detroit, Mich.; Orlando, Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, and a three-stage ratio estimation procedure for the remaining seven SMSA's.

Prior to the implementation of the ratio estimation procedures, 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 nonpermit 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 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.

The next ratio estimation procedure was applied in the Boston, Mass.; Detroit, Mich.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's. This procedure was utilized 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 ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

The next ratio estimation procedure was applied in 10 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1981 housing inventory in each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1981 housing unit inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

The independent estimates of total housing units that were used as the numerator of this ratio are described below. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight.

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for each of the 15 1981 SMSA's. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1981 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units after the permit new construction ratio estimation (i.e., the existing sample estimate) were compared to the corresponding independent estimates for the central city, balance, and total SMSA for each of the 15 SMSA's and the estimate which showed the most likely level of net growth since the 1980 census in both the central city and balance as well as the total SMSA were used in this ratio estimation. As a result of this analysis, these independent estimates were used in the Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's.

The remaining five SMSA's used a combination of these independent estimates and sample estimates. For the Boston, Mass.; and Minneapolis-St. Paul, Minn., SMSA's, the sample estimate was used for the total SMSA. For the sectors, the sample estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the independent estimates.

For the Orlando Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, the independent estimate of units was used for the total SMSA. For the sectors, the independent estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the sample estimates. Due to the central city

boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimate did not reflect the 1970 central city/balance SMSA definition. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definition were used.

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 in these 10 SMSA's, and the resulting product was used as the final weight for tabulation purposes. For the other five SMSA's, the sample estimates were used as the final weight for tabulation purposes. These SMSA's included Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Detroit, Mich.; Fort Worth, Tex.; and Madison, Wis., SMSA's.

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

1977-1981 lost housing units—The 1977-1981 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1977 housing inventory for 14 of the SMSA's and the corresponding two-stage ratio estimation procedure used for the Pittsburgh, Pa., SMSA as was described in the 1977 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1977-1981 lost housing units existed, by definition, in the 1977 housing inventory, there was a 1977 housing inventory weight associated with each 1977-1981 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1977-1981 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 two-stage ratio estimation process for the Pittsburgh, Pa., SMSA and a one-stage ratio estimation process for the remaining SMSA's. 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; 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 1981 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, and the results 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) 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, the 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 presented.

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 1981 survey, the number of missed housing units may be considerably less for 1981.

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 standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. Approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the average result of all possible samples.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimate would include the average result of all possible samples.

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

The figures presented in the tables that follow (pages App-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 1981 housing inventory as well as estimates of characteristics of the 1977-1981 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 4,676 for the total SMSA, 2,698 for the central city of the SMSA, and 3,737 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 1981 housing inventory as well as estimated percentages of the 1977-1981 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, tables II through IV 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 1981 there were 68,800 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 1,230. 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
50,000	1,100
68,800	x
75,000	1,270

The entry of "x" is determined as follows by vertically interpolating between 1,100 and 1,270.

$$68{,}800 - 50{,}000 = 18{,}800$$
$$75{,}000 - 50{,}000 = 25{,}000$$

$$1{,}100 + \frac{18{,}800}{25{,}000}\ (1{,}270 - 1{,}100) = 1{,}230$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 67,570 to 70,030 housing units. Therefore, a conclusion that the average estimate of 1981 owner-occupied housing units (derived from all possible samples) 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 66,830 to 70,770 housing units with 90 percent confidence; and that the average estimate lies within the interval from 66,340 to 71,260 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 68,800 owner-occupied housing units, 13,400, or 19.5 percent, had two bedrooms. Interpolation using table II of this appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 19.5 percent is approximately 0.8 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	19.5	25 or 75
50,000	0.7	a	1.0
68,800		p	
75,000	0.6	b	0.9

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

$$19.5 - 10.0 = 9.5$$
$$25.0 - 10.0 = 15.0$$

$$0.7 + \frac{9.5}{15.0}\ (1.0 - 0.7) = 0.9$$

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

$$19.5 - 10.0 = 9.5$$
$$25.0 - 10.0 = 15.0$$

$$0.6 + \frac{9.5}{15.0}\ (0.9 - 0.6) = 0.8$$

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

$$68{,}800 - 50{,}000 = 18{,}800$$
$$75{,}000 - 50{,}000 = 25{,}000$$

$$0.9 + \frac{18{,}800}{25{,}000}\ (0.8 - 0.9) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 18.7 to 20.3 percent; the 90-percent confidence interval is from 18.2 to 20.8 percent; and the 95-percent confidence interval is from 17.9 to 21.1 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The

standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. If there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1977 and 1981 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1977 and 1981 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1981 there were 38,600 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 25,200. Table I shows the standard error of 13,400 is approximately 590, and the standard error of 38,600 is approximately 970. Therefore, the standard error of the estimated difference of 25,200 is about 1,140.

$$1{,}140 = \sqrt{(590)^2 \quad (970)^2}$$

Consequently, the 68-percent confidence interval for the 25,200 difference is from 24,060 to 26,340 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 23,380 to 27,020 housing units, and the 95-percent confidence interval is from 22,920 to 27,480 housing units. Thus, we can conclude with 95 percent confidence that the number of 1981 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied 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 to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.8. The base of the distribution from which this median was determined is 68,800 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 68,800 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 31,000 housing units or 45.1 percent fall below this interval, and 13,100 housing units or 19.0 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5) \quad \frac{(48.0-45.1)}{19.0} = 2.7$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.0 percent derived in step 2. About 31,000 housing units or 45.1 percent fall below this interval, and 13,100 housing units or 19.0 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5) \quad \frac{(52.0-45.1)}{19.0} = 2.9$$

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

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1981 Housing Inventory and for Estimated Number of 1977-1981 Lost Units for the Madison, Wis., SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]	Size of estimate	Standard error[1]
0	30	5,000	380
100	50	10,000	530
200	80	25,000	810
500	120	50,000	1,100
700	140	75,000	1,270
1,000	170	100,000	1,380
2,500	270	150,000	1,460

[1]For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Madison, Wis., SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						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	22.5	22.5	22.5	22.5	23.3	26.9	10,000	0.3	0.5	1.2	1.6	2.3	2.7
200	12.7	12.7	12.7	12.7	16.5	19.0	25,000	0.12	0.3	0.7	1.0	1.5	1.7
500	5.5	5.5	5.5	7.2	10.4	12.0	50,000	0.06	0.2	0.5	0.7	1.0	1.2
700	4.0	4.0	4.4	6.1	8.8	10.2	75,000	0.04	0.2	0.4	0.6	0.9	1.0
1,000	2.8	2.8	3.7	5.1	7.4	8.5	100,000	0.03	0.2	0.4	0.5	0.7	0.9
2,500	1.1	1.1	2.3	3.2	4.7	5.4	150,000	0.02	0.14	0.3	0.4	0.6	0.7
5,000	0.6	0.8	1.7	2.3	3.3	3.8							

[1]Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2.

Newark 1981

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN . . . App-42
Annual Housing Survey . . . App-42
Designation of sample housing units for the 1981 survey . . . App-42
Selection of the original AHS-SMSA sample . . . App-43
1977-1981 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-45
1970 Census of Population and Housing . . . App-45
ESTIMATION . . . App-45
1981 housing inventory . . . App-45
1977-1981 lost housing units . . . App-47
1977 estimation procedure . . . App-47
Ratio estimation procedure of the 1970 Census of Population and Housing . . . App-47
RELIABILITY OF THE ESTIMATES . . . App-47
Nonsampling errors . . . App-47
1970 census . . . App-47
AHS-SMSA . . . App-47
Coverage errors . . . App-47
Rounding errors . . . App-48
Sampling errors for the AHS-SMSA sample . . . App-48
Illustration of the use of the standard error tables . . . App-49
Differences . . . App-49
Illustration of the computation of the standard error of a difference . . . App-50
Medians . . . App-50
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-170-81) are based on data collected from the 1981 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.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1981 were interviewed previously in 1974 (1975 for Madison, Wis.) and 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

The three largest SMSA's (Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va.) in the 1981 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1974 and 1977 surveys evenly divided between the central city and the balance of the respective SMSA. The 12 remaining SMSA's (Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.) in the 1981 group were represented by a sample of 5,000 designated housing units in the 1974 (1975 for Madison, Wis.) and 1977 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1974 survey, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panel 3 was dropped from all SMSA's for the 1977 survey. Due to additional budget limitations for the 1981 survey, it also became necessary to drop panels 1, 2, 4, 5, 9, 11, and 12 from the Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va., SMSA's and panels 1 and 2 from the other 12 SMSA's. The designated sample size was reduced by two-thirds for the three largest SMSA's and by one-fourth for the other 12 SMSA's.

The interviewing was done for the three largest SMSA's during the period June 1981 through November 1981 with one-sixth of the sample housing units being visited each month. The interviewing was done for the remaining 12 SMSA's during the period April 1981 through December 1981 with one-ninth of the designated housing units visited each month.

In this SMSA, 3,736 housing units were eligible for interview. Of these sample housing units, 202 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 3,736 housing units eligible for interview, 209 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 1981 survey—The sample housing units designated to be interviewed in the 1981

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 and remained in sample after the 1981 panel reduction.
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 and remained in sample after the 1981 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1981 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 and remained in sample after the 1981 panel reduction. (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 and remained in sample after the 1981 panel reduction. (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 1981 Coverage Improvement Programs. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the original 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 five SMSA's were 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 10 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 for the 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.

1977-1981 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 from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes 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. Houses that have been moved onto their present site since the 1970 census.
6. Mobile homes placed outside parks since the 1970 census or vacant at the time of the 1970 census.

Each of the 15 SMSA's was previously interviewed in 1974 (1975 for Madison, Wis.) and 1977. The Coverage Improvement Program was conducted as part of the 1977 AHS with the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., SMSA's receiving some updating and refining as part of the 1981 AHS. The following discussion applies to both the prior year 1977 and the 1981 coverage improvement procedures. For the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., 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 1981.

Coverage improvement for deficiency 1—A sample of new construction units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 15 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. 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 (regular AHS units), while sample units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage 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 Fort Worth, Tex., SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second-stage sampling. For the third stage, structures of size three or more units were divided into clusters of an expected size of two units and a sample of clusters was selected. This procedure was employed in the following SMSA's: Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Orlando, Fla.; Spokane, Wash.; and Wichita, Kans.

In the remaining 9 SMSA's, 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, this procedure was also used in parts of the Anaheim-Santa Ana-Garden Grove, Calif., SMSA to supplement the sample described above. These procedures added an estimated 3,759 new construction 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 placed in a park that was 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 units represented the same number of units that the regular AHS sample units represented. However, this procedure did not add any units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed 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 units converted to residential use.
3. Mobile homes 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 a usual residence elsewhere.
4. Houses that had been moved onto their present site since the 1970 census.

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample 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 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 second procedure was designed to represent missed units from structures represented in the AHS. These missed units were:

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential 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 structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply both procedures during the 1977 Coverage Improvement Program to all SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in the Anaheim-Santa Ana-Garden Grove, Calif.; and Spokane, Wash., SMSA's. These SMSA's did not receive application of the second procedure. The Dallas, Tex.; Minneapolis-St. Paul, Minn.; and Pittsburgh, Pa., SMSA's did not receive application of either procedure during the 1977 Coverage Improvement Program. The Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and the Spokane, Wash., SMSA's had the first procedure completed as part of the 1981 Coverage Improvement Program. The remaining 10 1981 SMSA's received the full application of both procedures as part of the 1977 Coverage Improvement Program.

The first procedure added an estimated 12,788 housing units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 4,106 housing units.

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 1981 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 1981 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1977 (i.e., 1977-1981 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1981 housing inventory—The AHS estimates of characteristics of the 1981 housing inventory were produced using a one-stage ratio estimation procedure for the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; and Madison, Wis., SMSA's, a two-stage ratio estimation procedure for the Detroit, Mich.; Orlando, Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, and a three-stage ratio estimation procedure for the remaining seven SMSA's.

Prior to the implementation of the ratio estimation procedures, 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 nonpermit 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 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.

The next ratio estimation procedure was applied in the Boston, Mass.; Detroit, Mich.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's. This procedure was utilized 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 ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

The next ratio estimation procedure was applied in 10 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1981 housing inventory in each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1981 housing unit inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

The independent estimates of total housing units that were used as the numerator of this ratio are described below. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight.

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for each of the 15 1981 SMSA's. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1981 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units after the permit new construction ratio estimation (i.e., the existing sample estimate) were compared to the corresponding independent estimates for the central city, balance, and total SMSA for each of the 15 SMSA's and the estimate which showed the most likely level of net growth since the 1980 census in both the central city and balance as well as the total SMSA were used in this ratio estimation. As a result of this analysis, these independent estimates were used in the Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's.

The remaining five SMSA's used a combination of these independent estimates and sample estimates. For the Boston, Mass.; and Minneapolis-St. Paul, Minn., SMSA's, the sample estimate was used for the total SMSA. For the sectors, the sample estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the independent estimates.

For the Orlando Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, the independent estimate of units was used for the total SMSA. For the sectors, the independent estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the sample estimates. Due to the central city

boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimate did not reflect the 1970 central city/balance SMSA definition. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definition were used.

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 in these 10 SMSA's, and the resulting product was used as the final weight for tabulation purposes. For the other five SMSA's, the sample estimates were used as the final weight for tabulation purposes. These SMSA's included Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Detroit, Mich.; Fort Worth, Tex.; and Madison, Wis., SMSA's.

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

1977-1981 lost housing units—The 1977-1981 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1977 housing inventory for 14 of the SMSA's and the corresponding two-stage ratio estimation procedure used for the Pittsburgh, Pa., SMSA as was described in the 1977 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1977-1981 lost housing units existed, by definition, in the 1977 housing inventory, there was a 1977 housing inventory weight associated with each 1977-1981 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1977-1981 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 two-stage ratio estimation process for the Pittsburgh, Pa., SMSA and a one-stage ratio estimation process for the remaining SMSA's. 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; 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 1981 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, and the results 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) 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, the 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 presented.

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 1981 survey, the number of missed housing units may be considerably less for 1981.

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 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 1981 housing inventory as well as estimates of characteristics of the 1977-1981 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 20,957 for the total SMSA, 12,772 for the central city of the SMSA, and 16,563 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 1981 housing inventory as well as estimated percentages of the 1977-1981 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, tables II through IV 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 1981 there were 342,100 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 5,460. 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
300,000	5,530
342,100	x
400,000	5,370

The entry of "x" is determined as follows by vertically interpolating between 5,530 and 5,370.

$$342{,}100-300{,}000=42{,}100$$
$$400{,}000-300{,}000=100{,}000$$

$$5{,}530+\frac{42{,}100}{100{,}000}(5{,}370-5{,}530)=5{,}460$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 336,640 to 347,560 housing units. Therefore, a conclusion that the average estimate of 1981 owner-occupied housing units (derived from all possible samples) 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 333,360 to 350,840 housing units with 90 percent confidence; and that the average estimate lies within the interval from 331,180 to 353,020 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 342,100 owner-occupied housing units, 76,300, or 22.3 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 22.3 percent is approximately 1.0 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	22.3	25 or 75
300,000	0.8	a	1.2
342,100		p	
400,000	0.7	b	1.0

1. The entry for cell "a" is determined by horizontal interpolation between 0.8 and 1.2.

$$22.3-10.0=12.3$$
$$25.0-10.0=15.0$$

$$0.8+\frac{12.3}{15.0}(1.2-0.8)=1.1$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.7 and 1.0.

$$22.3-10.0=12.3$$
$$25.0-10.0=15.0$$

$$0.7+\frac{12.3}{15.0}(1.0-0.7)=0.9$$

3. The entry for "p" is then determined by vertical interpolation between 1.1 and 0.9.

$$342{,}100-300{,}000=42{,}100$$
$$400{,}000-300{,}000=100{,}000$$

$$1.1+\frac{42{,}100}{100{,}000}(0.9-1.1)=1.0$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 21.3 to 23.3 percent; the 90-percent confidence interval is from 20.7 to 23.9 percent; and the 95-percent confidence interval is from 20.3 to 24.3 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 1977 and 1981 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1977 and 1981 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1981 there were 155,900 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 79,600. Table I shows the standard error of 76,300 is approximately 3,590, and the standard error of 155,900 is approximately 4,750. Therefore, the standard error of the estimated difference of 79,600 is about 5,950.

$$5{,}950 = \sqrt{(3{,}590)^2 + (4{,}750)^2}$$

Consequently, the 68-percent confidence interval for the 79,600 difference is from 73,650 to 85,550 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 70,080 to 89,120 housing units, and the 95-percent confidence interval is from 67,700 to 91,500 housing units. Thus, we can conclude with 95 percent confidence that the number of 1981 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied 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 to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 3.0. The base of the distribution from which this median was determined is 342,100 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 342,100 is approximately 1.3 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 47.4 and 52.6
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 47.4 percent derived in step 2. About 132,900 housing units or 38.8 percent fall below this interval, and 71,700 housing units or 21.0 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(47.4-38.8)}{21.0} = 2.9$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.6 percent derived in step 2. About 132,900 housing units or 38.8 percent fall below this interval, and 71,700 housing units or 21.0 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(52.6-38.8)}{21.0} = 3.2$$

Thus, the 95-percent confidence interval ranges from 2.9 to 3.2 persons.

TABLE I. **Standard Errors for Estimated Number of Housing Units in the 1981 Housing Inventory and for Estimated Number of 1977-1981 Lost Units for the Newark, 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	190	220	180	100,000	4,030	1,930	3,810
100	190	220	180	120,800	4,340	—	4,080
200	200	220	190	150,000	4,700	—	4,380
500	310	330	300	200,000	5,140	—	4,700
700	370	390	350	250,000	5,410	—	4,820
1,000	440	460	420	300,000	5,530	—	4,770
2,500	690	730	670	400,000	5,370	—	4,060
5,000	980	1,020	940	500,000	4,590	—	1,830
10,000	1,380	1,410	1,330	519,300	4,340	—	—
25,000	2,150	2,070	2,070	600,000	2,690	—	—
50,000	2,980	2,510	2,850	640,200	—	—	—
75,000	3,570	2,480	3,390				

[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, 1.0 for the central city, and 1.4 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 where the estimates for the total SMSA were derived from an independent estimate.

TABLE II. **Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Newark, N.J., 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	68.3	68.3	68.3	68.3	68.3	73.4	75,000	0.3	0.5	1.2	1.6	2.3	2.7
200	51.9	51.9	51.9	51.9	51.9	51.9	100,000	0.2	0.5	1.0	1.4	2.0	2.3
500	30.1	30.1	30.1	30.1	30.1	32.8	150,000	0.14	0.4	0.8	1.1	1.6	1.9
700	23.5	23.5	23.5	23.5	24.0	27.7	200,000	0.11	0.3	0.7	1.0	1.4	1.6
1,000	17.7	17.7	17.7	17.7	20.1	23.2	250,000	0.09	0.3	0.6	0.9	1.3	1.5
2,500	7.9	7.9	7.9	8.8	12.7	14.7	300,000	0.07	0.3	0.6	0.8	1.2	1.3
5,000	4.1	4.1	4.5	6.2	9.0	10.4	400,000	0.05	0.2	0.5	0.7	1.0	1.2
10,000	2.1	2.1	3.2	4.4	6.4	7.3	500,000	0.04	0.2	0.5	0.6	0.9	1.0
25,000	0.9	0.9	2.0	2.8	4.0	4.6	600,000	0.04	0.2	0.4	0.6	0.8	0.9
50,000	0.4	0.7	1.4	2.0	2.8	3.3	700,000	0.03	0.2	0.4	0.5	0.8	0.9

[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 1.3 for the total SMSA, 1.0 for the central city, and 1.4 for the balance (not in central city).

Orlando 1981

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey . . . App-42
Designation of sample housing units for the 1981 survey . . App-42
Selection of the original AHS-SMSA sample . . App-43
1977-1981 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-45
1970 Census of Population and Housing . . App-45
ESTIMATION . . App-45
1981 housing inventory . . App-45
1977-1981 lost housing units . . App-47
1977 estimation procedure . . App-47
Ratio estimation procedure of the 1970 Census of Population and Housing . . App-47
RELIABILITY OF THE ESTIMATES . . App-47
Nonsampling errors . . App-47
1970 census . . App-47
AHS-SMSA . . App-47
Coverage errors . . App-47
Rounding errors . . App-48
Sampling errors for the AHS-SMSA sample . . App-48
Illustration of the use of the standard error tables . . App-49
Differences . . App-49
Illustration of the computation of the standard error of a difference . . App-50
Medians . . App-50
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-170-81) are based on data collected from the 1981 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.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1981 were interviewed previously in 1974 (1975 for Madison, Wis.) and 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

The three largest SMSA's (Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va.) in the 1981 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1974 and 1977 surveys evenly divided between the central city and the balance of the respective SMSA. The 12 remaining SMSA's (Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.) in the 1981 group were represented by a sample of 5,000 designated housing units in the 1974 (1975 for Madison, Wis.) and 1977 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1974 survey, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panel 3 was dropped from all SMSA's for the 1977 survey. Due to additional budget limitations for the 1981 survey, it also became necessary to drop panels 1, 2, 4, 5, 9, 11, and 12 from the Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va., SMSA's and panels 1 and 2 from the other 12 SMSA's. The designated sample size was reduced by two-thirds for the three largest SMSA's and by one-fourth for the other 12 SMSA's.

The interviewing was done for the three largest SMSA's during the period June 1981 through November 1981 with one-sixth of the sample housing units being visited each month. The interviewing was done for the remaining 12 SMSA's during the period April 1981 through December 1981 with one-ninth of the designated housing units visited each month.

In this SMSA, 4,417 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 4,417 housing units eligible for interview, 216 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 1981 survey—The sample housing units designated to be interviewed in the 1981

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 and remained in sample after the 1981 panel reduction.
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 and remained in sample after the 1981 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1981 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 and remained in sample after the 1981 panel reduction. (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 and remained in sample after the 1981 panel reduction. (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 1981 Coverage Improvement Programs. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the original 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 five SMSA's were 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 10 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 for the 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.

1977-1981 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 from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes 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. Houses that have been moved onto their present site since the 1970 census.
6. Mobile homes placed outside parks since the 1970 census or vacant at the time of the 1970 census.

Each of the 15 SMSA's was previously interviewed in 1974 (1975 for Madison, Wis.) and 1977. The Coverage Improvement Program was conducted as part of the 1977 AHS with the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., SMSA's receiving some updating and refining as part of the 1981 AHS. The following discussion applies to both the prior year 1977 and the 1981 coverage improvement procedures. For the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., 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 1981.

Coverage improvement for deficiency 1—A sample of new construction units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 15 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. 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 (regular AHS units), while sample units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage 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 Fort Worth, Tex., SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second-stage sampling. For the third stage, structures of size three or more units were divided into clusters of an expected size of two units and a sample of clusters was selected. This procedure was employed in the following SMSA's: Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Orlando, Fla.; Spokane, Wash.; and Wichita, Kans.

In the remaining 9 SMSA's, 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, this procedure was also used in parts of the Anaheim-Santa Ana-Garden Grove, Calif., SMSA to supplement the sample described above. These procedures added an estimated 990 new construction 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 placed in a park that was 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 units represented the same number of units that the regular AHS sample units represented. This procedure added an estimated 3,433 units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed 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 units converted to residential use.
3. Mobile homes 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 a usual residence elsewhere.
4. Houses that had been moved onto their present site since the 1970 census.

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample 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 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 second procedure was designed to represent missed units from structures represented in the AHS. These missed units were:

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential 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 structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply both procedures during the 1977 Coverage Improvement Program to all SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in the Anaheim-Santa Ana-Garden Grove, Calif.; and Spokane, Wash., SMSA's. These SMSA's did not receive application of the second procedure. The Dallas, Tex.; Minneapolis-St. Paul, Minn.; and Pittsburgh, Pa., SMSA's did not receive application of either procedure during the 1977 Coverage Improvement Program. The Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and the Spokane, Wash., SMSA's had the first procedure completed as part of the 1981 Coverage Improvement Program. The remaining 10 1981 SMSA's received the full application of both procedures as part of the 1977 Coverage Improvement Program.

The first procedure added an estimated 12,949 housing units to the coverage of the housing inventory of this SMSA, while the second procedure did not add any housing units.

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 1981 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 1981 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1977 (i.e., 1977-1981 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1981 housing inventory—The AHS estimates of characteristics of the 1981 housing inventory were produced using a one-stage ratio estimation procedure for the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; and Madison, Wis., SMSA's, a two-stage ratio estimation procedure for the Detroit, Mich.; Orlando, Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, and a three-stage ratio estimation procedure for the remaining seven SMSA's.

Prior to the implementation of the ratio estimation procedures, 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 nonpermit 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 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.

The next ratio estimation procedure was applied in the Boston, Mass.; Detroit, Mich.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's. This procedure was utilized 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 ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

The next ratio estimation procedure was applied in 10 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1981 housing inventory in each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1981 housing unit inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

The independent estimates of total housing units that were used as the numerator of this ratio are described below. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight.

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for each of the 15 1981 SMSA's. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1981 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units after the permit new construction ratio estimation (i.e., the existing sample estimate) were compared to the corresponding independent estimates for the central city, balance, and total SMSA for each of the 15 SMSA's and the estimate which showed the most likely level of net growth since the 1980 census in both the central city and balance as well as the total SMSA were used in this ratio estimation. As a result of this analysis, these independent estimates were used in the Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's.

The remaining five SMSA's used a combination of these independent estimates and sample estimates. For the Boston, Mass.; and Minneapolis-St. Paul, Minn., SMSA's, the sample estimate was used for the total SMSA. For the sectors, the sample estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the independent estimates.

For the Orlando Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, the independent estimate of units was used for the total SMSA. For the sectors, the independent estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the sample estimates. Due to the central city

boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimate did not reflect the 1970 central city/balance SMSA definition. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definition were used.

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 in these 10 SMSA's, and the resulting product was used as the final weight for tabulation purposes. For the other five SMSA's, the sample estimates were used as the final weight for tabulation purposes. These SMSA's included Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Detroit, Mich.; Fort Worth, Tex.; and Madison, Wis., SMSA's.

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

1977-1981 lost housing units—The 1977-1981 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1977 housing inventory for 14 of the SMSA's and the corresponding two-stage ratio estimation procedure used for the Pittsburgh, Pa., SMSA as was described in the 1977 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1977-1981 lost housing units existed, by definition, in the 1977 housing inventory, there was a 1977 housing inventory weight associated with each 1977-1981 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1977-1981 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 two-stage ratio estimation process for the Pittsburgh, Pa., SMSA and a one-stage ratio estimation process for the remaining SMSA's. 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; 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 1981 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, and the results 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) 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, the 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 presented.

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 1981 survey, the number of missed housing units may be considerably less for 1981.

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 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 1981 housing inventory as well as estimates of characteristics of the 1977-1981 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 8,651 for the total SMSA, 2,934 for the central city of the SMSA, and 8,064 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 1981 housing inventory as well as estimated percentages of the 1977-1981 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 1981 there were 159,300 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 2,100. 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	2,150
159,300	x
200,000	1,880

The entry of "x" is determined as follows by vertically interpolating between 2,150 and 1,880.

$$159{,}300 - 150{,}000 = 9{,}300$$
$$200{,}000 - 150{,}000 = 50{,}000$$

$$2{,}150 + \frac{9{,}300}{50{,}000}(1{,}880 - 2{,}150) = 2{,}100$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 157,200 to 161,400 housing units. Therefore, a conclusion that the average estimate of 1981 owner-occupied housing units (derived from all possible samples) 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 155,940 to 162,660 housing units with 90 percent confidence; and that the average estimate lies within the interval from 155,100 to 163,500 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 159,300 owner-occupied housing units, 39,000, or 24.5 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 24.5 percent is approximately 1.0 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	24.5	25 or 75
150,000	0.7	a	1.0
159,300		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.

$$24.5 - 10.0 = 14.5$$
$$25.0 - 10.0 = 15.0$$

$$0.7 + \frac{14.5}{15.0}(1.0 - 0.7) = 1.0$$

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

$$24.5 - 10.0 = 14.5$$
$$25.0 - 10.0 = 15.0$$

$$0.6 + \frac{14.5}{15.0}(0.8 - 0.6) = 0.8$$

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

$$159{,}300 - 150{,}000 = 9{,}300$$
$$200{,}000 - 150{,}000 = 50{,}000$$

$$1.0 + \frac{9{,}300}{50{,}000}(0.8 - 1.0) = 1.0$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 23.5 to 25.5 percent; the 90-percent confidence interval is from 22.9 to 26.1 percent; and the 95-percent confidence interval is from 22.5 to 26.5 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; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1977 and 1981 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1977 and 1981 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1981 there were 86,100 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 47,100. Table I shows the standard error of 39,000 is approximately 1,530, and the standard error of 86,100 is approximately 2,020. Therefore, the standard error of the estimated difference of 47,100 is about 2,530:

$$2{,}530 = \sqrt{(1{,}530)^2 + (2{,}020)^2}$$

Consequently, the 68-percent confidence interval for the 47,100 difference is from 44,570 to 49,630 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 43,050 to 51,150 housing units, and the 95-percent confidence interval is from 42,040 to 52,160 housing units. Thus, we can conclude with 95 percent confidence that the number of 1981 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied 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 to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.5. The base of the distribution from which this median was determined is 159,300 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 159,300 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 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 47.8 percent derived in step 2. About 24,200 housing units or 15.2 percent fall below this interval, and 53,800 housing units or 33.8 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$1.5 + (2.5-1.5)\ \frac{(47.8-15.2)}{33.8} = 2.5$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.2 percent derived in step 2. About 78,000 housing units or 49.0 percent fall below this interval, and 31,200 housing units or 19.6 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(52.2-49.0)}{19.6} = 2.7$$

Thus, the 95-percent confidence interval ranges from 2.5 to 2.7 persons. Although it appears that this confidence interval has the same estimate as the lower limit it actually is a refelction of the rounding error associated with this median (see the paragraph on rounding error in the nonsampling error section of this appendix).

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1981 Housing Inventory and for Estimated Number of 1977-1981 Lost Units for the Orlando, Fla., SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]			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	43,500	1,610	—	1,620
100	80	80	90	50,000	1,700	—	1,700
200	120	110	120	75,000	1,960	—	1,930
500	190	170	190	100,000	2,100	—	2,030
700	220	210	230	150,000	2,150	—	1,910
1,000	270	250	270	200,000	1,880	—	1,230
2,500	420	380	430	222,500	1,610	—	—
5,000	590	520	600	250,000	1,030	—	—
10,000	830	690	840	265,900	—	—	—
25,000	1,270	810	1,290				

[1]For estimates pertaining to new construction and lost units, the standard errors shown in the table should be multiplied by a factor of 1.1. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero where the estimate for the total SMSA was derived from an independent estimate.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Orlando, Fla., SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						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	42.7	42.7	42.7	42.7	42.7	43.2	25,000	0.3	0.5	1.2	1.6	2.4	2.7
200	27.2	27.2	27.2	27.2	27.2	30.5	50,000	0.15	0.4	0.8	1.2	1.7	1.9
500	13.0	13.0	13.0	13.0	16.7	19.3	75,000	0.10	0.3	0.7	0.9	1.4	1.6
700	9.6	9.6	9.6	9.8	14.1	16.3	100,000	0.07	0.3	0.6	0.8	1.2	1.4
1,000	6.9	6.9	6.9	8.2	11.8	13.7	150,000	0.05	0.2	0.5	0.7	1.0	1.1
2,500	2.9	2.9	3.8	5.2	7.5	8.6	200,000	0.04	0.2	0.4	0.6	0.8	1.0
5,000	1.5	1.5	2.7	3.7	5.3	6.1	250,000	0.03	0.2	0.4	0.5	0.7	0.9
10,000	0.7	0.9	1.9	2.6	3.7	4.3	300,000	0.02	0.2	0.3	0.5	0.7	0.8

[1]Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction and lost units, the standard errors shown in the table should be multiplied by a factor of 1.1.

Phoenix 1981

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN . . . App-42
Annual Housing Survey . . . App-42
Designation of sample housing units for the 1981 survey . . . App-42
Selection of the original AHS-SMSA sample . . . App-43
1977-1981 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-45
1970 Census of Population and Housing . . . App-45
ESTIMATION . . . App-45
1981 housing inventory . . . App-45
1977-1981 lost housing units . . . App-47
1977 estimation procedure . . . App-47
Ratio estimation procedure of the 1970 Census of Population and Housing . . . App-47
RELIABILITY OF THE ESTIMATES . . . App-47
Nonsampling errors . . . App-47
1970 census . . . App-47
AHS-SMSA . . . App-47
Coverage errors . . . App-47
Rounding errors . . . App-48
Sampling errors for the AHS-SMSA sample . . . App-48
Illustration of the use of the standard error tables . . . App-49
Differences . . . App-49
Illustration of the computation of the standard error of a difference . . . App-50
Medians . . . App-50
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-170-81) are based on data collected from the 1981 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.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1981 were interviewed previously in 1974 (1975 for Madison, Wis.) and 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

The three largest SMSA's (Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va.) in the 1981 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1974 and 1977 surveys evenly divided between the central city and the balance of the respective SMSA. The 12 remaining SMSA's (Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.) in the 1981 group were represented by a sample of 5,000 designated housing units in the 1974 (1975 for Madison, Wis.) and 1977 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1974 survey, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panel 3 was dropped from all SMSA's for the 1977 survey. Due to additional budget limitations for the 1981 survey, it also became necessary to drop panels 1, 2, 4, 5, 9, 11, and 12 from the Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va., SMSA's and panels 1 and 2 from the other 12 SMSA's. The designated sample size was reduced by two-thirds for the three largest SMSA's and by one-fourth for the other 12 SMSA's.

The interviewing was done for the three largest SMSA's during the period June 1981 through November 1981 with one-sixth of the sample housing units being visited each month. The interviewing was done for the remaining 12 SMSA's during the period April 1981 through December 1981 with one-ninth of the designated housing units visited each month.

In this SMSA, 5,134 housing units were eligible for interview. Of these sample housing units, 137 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 5,134 housing units eligible for interview, 362 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 1981 survey—The sample housing units designated to be interviewed in the 1981

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 and remained in sample after the 1981 panel reduction.
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 and remained in sample after the 1981 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1981 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 and remained in sample after the 1981 panel reduction. (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 and remained in sample after the 1981 panel reduction. (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 1981 Coverage Improvement Programs. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the original 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 five SMSA's were 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 10 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 for the 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.

1977-1981 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 from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes 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. Houses that have been moved onto their present site since the 1970 census.
6. Mobile homes placed outside parks since the 1970 census or vacant at the time of the 1970 census.

Each of the 15 SMSA's was previously interviewed in 1974 (1975 for Madison, Wis.) and 1977. The Coverage Improvement Program was conducted as part of the 1977 AHS with the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., SMSA's receiving some updating and refining as part of the 1981 AHS. The following discussion applies to both the prior year 1977 and the 1981 coverage improvement procedures. For the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., 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 1981.

Coverage improvement for deficiency 1—A sample of new construction units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 15 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. 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 (regular AHS units), while sample units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage 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 Fort Worth, Tex., SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second-stage sampling. For the third stage, structures of size three or more units were divided into clusters of an expected size of two units and a sample of clusters was selected. This procedure was employed in the following SMSA's: Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Orlando, Fla.; Spokane, Wash.; and Wichita, Kans.

In the remaining 9 SMSA's, 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, this procedure was also used in parts of the Anaheim-Santa Ana-Garden Grove, Calif., SMSA to supplement the sample described above. These procedures added an estimated 4,544 new construction 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 placed in a park that was 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 units represented the same number of units that the regular AHS sample units represented. This procedure added an estimated 10,615 units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed 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 units converted to residential use.
3. Mobile homes 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 a usual residence elsewhere.
4. Houses that had been moved onto their present site since the 1970 census.

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample 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 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 second procedure was designed to represent missed units from structures represented in the AHS. These missed units were:

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential 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 structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply both procedures during the 1977 Coverage Improvement Program to all SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in the Anaheim-Santa Ana-Garden Grove, Calif.; and Spokane, Wash., SMSA's. These SMSA's did not receive application of the second procedure. The Dallas, Tex.; Minneapolis-St. Paul, Minn.; and Pittsburgh, Pa., SMSA's did not receive application of either procedure during the 1977 Coverage Improvement Program. The Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and the Spokane, Wash., SMSA's had the first procedure completed as part of the 1981 Coverage Improvement Program. The remaining 10 1981 SMSA's received the full application of both procedures as part of the 1977 Coverage Improvement Program.

The first procedure added an estimated 13,937 housing units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 488 housing units.

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 1981 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 1981 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1977 (i.e., 1977-1981 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1981 housing inventory—The AHS estimates of characteristics of the 1981 housing inventory were produced using a one-stage ratio estimation procedure for the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; and Madison, Wis., SMSA's, a two-stage ratio estimation procedure for the Detroit, Mich.; Orlando, Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, and a three-stage ratio estimation procedure for the remaining seven SMSA's.

Prior to the implementation of the ratio estimation procedures, 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 nonpermit 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 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.

The next ratio estimation procedure was applied in the Boston, Mass.; Detroit, Mich.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's. This procedure was utilized 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 ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

The next ratio estimation procedure was applied in 10 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1981 housing inventory in each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1981 housing unit inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

The independent estimates of total housing units that were used as the numerator of this ratio are described below. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight.

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for each of the 15 1981 SMSA's. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1981 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units after the permit new construction ratio estimation (i.e., the existing sample estimate) were compared to the corresponding independent estimates for the central city, balance, and total SMSA for each of the 15 SMSA's and the estimate which showed the most likely level of net growth since the 1980 census in both the central city and balance as well as the total SMSA were used in this ratio estimation. As a result of this analysis, these independent estimates were used in the Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's.

The remaining five SMSA's used a combination of these independent estimates and sample estimates. For the Boston, Mass.; and Minneapolis-St. Paul, Minn., SMSA's, the sample estimate was used for the total SMSA. For the sectors, the sample estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the independent estimates.

For the Orlando Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, the independent estimate of units was used for the total SMSA. For the sectors, the independent estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the sample estimates. Due to the central city

boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimate did not reflect the 1970 central city/balance SMSA definition. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definition were used.

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 in these 10 SMSA's, and the resulting product was used as the final weight for tabulation purposes. For the other five SMSA's, the sample estimates were used as the final weight for tabulation purposes. These SMSA's included Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Detroit, Mich.; Fort Worth, Tex.; and Madison, Wis., SMSA's.

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

1977-1981 lost housing units—The 1977-1981 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1977 housing inventory for 14 of the SMSA's and the corresponding two-stage ratio estimation procedure used for the Pittsburgh, Pa., SMSA as was described in the 1977 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas*. Since the 1977-1981 lost housing units existed, by definition, in the 1977 housing inventory, there was a 1977 housing inventory weight associated with each 1977-1981 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1977-1981 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 two-stage ratio estimation process for the Pittsburgh, Pa., SMSA and a one-stage ratio estimation process for the remaining SMSA's. 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; 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 1981 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, and the results 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) 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, the 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 presented.

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 1981 survey, the number of missed housing units may be considerably less for 1981.

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 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 1981 housing inventory as well as estimates of characteristics of the 1977-1981 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 27,555 for the total SMSA, 19,050 for the central city of the SMSA, and 19,822 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 1981 housing inventory as well as estimated percentages of the 1977-1981 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 1981 there were 378,100 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 4,740. 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
345,900	4,820
378,100	x
400,000	4,690

The entry of "x" is determined as follows by vertically interpolating between 4,820 and 4,690.

$$378{,}100 - 345{,}900 = 32{,}200$$
$$400{,}000 - 345{,}900 = 54{,}100$$

$$4{,}820 + \frac{32{,}200}{54{,}100}(4{,}690 - 4{,}820) = 4{,}740$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 373,360 to 382,840 housing units. Therefore, a conclusion that the average estimate of 1981 owner-occupied housing units (derived from all possible samples) 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 370,520 to 385,680 housing units with 90 percent confidence; and that the average estimate lies within the interval from 368,620 to 387,580 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 378,100 owner-occupied housing units, 101,500, or 26.8 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 26.8 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	26.8	50
300,000	1.0	a	1.1
378,100		p	
400,000	0.8	b	1.0

1. The entry for cell "a" is determined by horizontal interpolation between 1.0 and 1.1.

$$26.8 - 25.0 = 1.8$$
$$50.0 - 25.0 = 25.0$$

$$1.0 + \frac{1.8}{25.0}(1.1 - 1.0) = 1.0$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.8 and 1.0.

$$26.8 - 25.0 = 1.8$$
$$50.0 - 25.0 = 25.0$$

$$0.8 + \frac{1.8}{25.0}(1.0 - 0.8) = 0.8$$

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

$$378{,}100 - 300{,}000 = 78{,}100$$
$$400{,}000 - 300{,}000 = 100{,}000$$

$$1.0 + \frac{78{,}100}{100{,}000}(0.8 - 1.0) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 26.0 to 27.6 percent; the 90-percent confidence interval is from 25.5 to 28.1 percent; and the 95-percent confidence interval is from 25.2 to 28.4 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The

standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. If there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1977 and 1981 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1977 and 1981 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1981 there were 183,300 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 81,800. Table I shows the standard error of 101,500 is approximately 3,530, and the standard error of 183,300 is approximately 4,350. Therefore, the standard error of the estimated difference of 81,800 is about 5,600:

$$5{,}600 = \sqrt{(3{,}530)^2 + (4{,}350)^2}$$

Consequently, the 68-percent confidence interval for the 81,800 difference is from 76,200 to 87,400 housing units. Therefore, a conclusion that the average estimate derived from all possible samples, of this difference, lies within a range computed in this way, would be correct for roughly 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 72,840 to 90,760 housing units, and the 95-percent confidence interval is from 70,600 to 93,000 housing units. Thus, we can conclude with 95 percent confidence that the number of 1981 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied 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 to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.4. The base of the distribution from which this median was determined is 378,100 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 378,100 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 60,300 housing units or 15.9 percent fall below this interval, and 137,000 housing units or 36.2 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$1.5 + (2.5-1.5)\ \frac{(48.0-15.9)}{36.2} = 2.4$$

Similarly, the interval for owner-occupied housing units with two persons contains the 52.0 percent derived in step 2. About 60,300 housing units or 15.9 percent fall below this interval, and 137,000 housing units or 36.2 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$1.5 + (2.5-1.5)\ \frac{(52.0-15.9)}{36.2} = 2.5$$

Thus, the 95-percent confidence interval ranges from 2.4 to 2.5 persons. Although it appears that this confidence interval has the same estimate as the lower limit it actually is a refelction of the rounding error associated with this median (see the paragraph on rounding error in the nonsampling error section of this appendix).

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1981 Housing Inventory and for Estimated Number of 1977-1981 Lost Units for the Phoenix, Ariz., 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	150	140	150	100,000	3,510	3,030	3,210
100	150	140	150	150,000	4,090	3,200	3,510
200	170	170	170	200,000	4,480	3,000	3,500
500	270	260	270	250,000	4,720	2,320	3,170
700	320	310	320	296,300	4,820	—	2,480
1,000	380	370	380	300,000	4,830	—	2,400
2,500	600	590	600	345,900	4,820	—	—
5,000	850	820	850	400,000	4,690	—	—
10,000	1,200	1,160	1,190	500,000	4,020	—	—
25,000	1,870	1,780	1,830	600,000	2,400	—	—
50,000	2,590	2,400	2,490	642,200	—	—	—
75,000	3,110	2,780	2,920				

[1]For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.1. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero where the estimate for the total SMSA was derived from an independent estimate.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Phoenix, Ariz., 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	59.3	59.3	59.3	59.3	59.3	60.4	75,000	0.2	0.4	1.0	1.3	1.9	2.2
200	42.2	42.2	42.2	42.2	42.2	42.7	100,000	0.15	0.4	0.8	1.1	1.7	1.9
500	22.6	22.6	22.6	22.6	23.4	27.0	150,000	0.10	0.3	0.7	0.9	1.3	1.6
700	17.2	17.2	17.2	17.2	19.8	22.8	200,000	0.07	0.3	0.6	0.8	1.2	1.3
1,000	12.7	12.7	12.7	12.7	16.5	19.1	250,000	0.06	0.2	0.5	0.7	1.0	1.2
2,500	5.5	5.5	5.5	7.2	10.5	12.1	300,000	0.05	0.2	0.5	0.7	1.0	1.1
5,000	2.8	2.8	3.7	5.1	7.4	8.5	400,000	0.04	0.2	0.4	0.6	0.8	1.0
10,000	1.4	1.4	2.6	3.6	5.2	6.0	500,000	0.03	0.2	0.4	0.5	0.7	0.9
25,000	0.6	0.8	1.7	2.3	3.3	3.8	600,000	0.02	0.2	0.3	0.5	0.7	0.8
50,000	0.3	0.5	1.2	1.6	2.3	2.7	700,000	0.02	0.14	0.3	0.4	0.6	0.7

[1]Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.1.

Pittsburgh 1981

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN . . . App-42
Annual Housing Survey . . . App-42
Designation of sample housing units for the 1981 survey . . . App-42
Selection of the original AHS-SMSA sample . . . App-43
1977-1981 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-45
1970 Census of Population and Housing . . . App-45
ESTIMATION . . . App-45
1981 housing inventory . . . App-45
1977-1981 lost housing units . . . App-47
1977 estimation procedure . . . App-47
Ratio estimation procedure of the 1970 Census of Population and Housing . . . App-47
RELIABILITY OF THE ESTIMATES . . . App-47
Nonsampling errors . . . App-47
1970 census . . . App-47
AHS-SMSA . . . App-47
Coverage errors . . . App-47
Rounding errors . . . App-48
Sampling errors for the AHS-SMSA sample . . . App-48
Illustration of the use of the standard error tables . . . App-49
Differences . . . App-49
Illustration of the computation of the standard error of a difference . . . App-50
Medians . . . App-50
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-170-81) are based on data collected from the 1981 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.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1981 were interviewed previously in 1974 (1975 for Madison, Wis.) and 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

The three largest SMSA's (Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va.) in the 1981 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1974 and 1977 surveys evenly divided between the central city and the balance of the respective SMSA. The 12 remaining SMSA's (Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.) in the 1981 group were represented by a sample of 5,000 designated housing units in the 1974 (1975 for Madison, Wis.) and 1977 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1974 survey, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panel 3 was dropped from all SMSA's for the 1977 survey. Due to additional budget limitations for the 1981 survey, it also became necessary to drop panels 1, 2, 4, 5, 9, 11, and 12 from the Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va., SMSA's and panels 1 and 2 from the other 12 SMSA's. The designated sample size was reduced by two-thirds for the three largest SMSA's and by one-fourth for the other 12 SMSA's.

The interviewing was done for the three largest SMSA's during the period June 1981 through November 1981 with one-sixth of the sample housing units being visited each month. The interviewing was done for the remaining 12 SMSA's during the period April 1981 through December 1981 with one-ninth of the designated housing units visited each month.

In this SMSA, 3,885 housing units were eligible for interview. Of these sample housing units, 100 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 3,885 housing units eligible for interview, 258 units were visited but were not eligible for interview because they were condemned, unfit, demolished, converted to group quarters use, etc.

Designation of sample housing units for the 1981 survey—The sample housing units designated to be interviewed in the 1981

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 and remained in sample after the 1981 panel reduction.
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 and remained in sample after the 1981 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1981 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 and remained in sample after the 1981 panel reduction. (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 and remained in sample after the 1981 panel reduction. (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 1981 Coverage Improvement Programs. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the original 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 five SMSA's were 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 10 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 for the 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.

1977-1981 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 from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes 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. Houses that have been moved onto their present site since the 1970 census.
6. Mobile homes placed outside parks since the 1970 census or vacant at the time of the 1970 census.

Each of the 15 SMSA's was previously interviewed in 1974 (1975 for Madison, Wis.) and 1977. The Coverage Improvement Program was conducted as part of the 1977 AHS with the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., SMSA's receiving some updating and refining as part of the 1981 AHS. The following discussion applies to both the prior year 1977 and the 1981 coverage improvement procedures. For the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., 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 1981.

Coverage improvement for deficiency 1—A sample of new construction units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 15 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. 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 (regular AHS units), while sample units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage 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 Fort Worth, Tex., SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second-stage sampling. For the third stage, structures of size three or more units were divided into clusters of an expected size of two units and a sample of clusters was selected. This procedure was employed in the following SMSA's: Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Orlando, Fla.; Spokane, Wash.; and Wichita, Kans.

In the remaining 9 SMSA's, 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, this procedure was also used in parts of the Anaheim-Santa Ana-Garden Grove, Calif., SMSA to supplement the sample described above. These procedures added an estimated 2,056 new construction 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 placed in a park that was 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 units represented the same number of units that the regular AHS sample units represented. This procedure added an estimated 3,236 units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed 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 units converted to residential use.
3. Mobile homes 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 a usual residence elsewhere.
4. Houses that had been moved onto their present site since the 1970 census.

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample 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 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 second procedure was designed to represent missed units from structures represented in the AHS. These missed units were:

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential 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 structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply both procedures during the 1977 Coverage Improvement Program to all SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in the Anaheim-Santa Ana-Garden Grove, Calif.; and Spokane, Wash., SMSA's. These SMSA's did not receive application of the second procedure. The Dallas, Tex.; Minneapolis-St. Paul, Minn.; and Pittsburgh, Pa., SMSA's did not receive application of either procedure during the 1977 Coverage Improvement Program. The Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and the Spokane, Wash., SMSA's had the first procedure completed as part of the 1981 Coverage Improvement Program. The remaining 10 1981 SMSA's received the full application of both procedures as part of the 1977 Coverage Improvement Program.

The first procedure added an estimated 18,896 housing units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 488 housing units.

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 1981 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 1981 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1977 (i.e., 1977-1981 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1981 housing inventory—The AHS estimates of characteristics of the 1981 housing inventory were produced using a one-stage ratio estimation procedure for the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; and Madison, Wis., SMSA's, a two-stage ratio estimation procedure for the Detroit, Mich.; Orlando, Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, and a three-stage ratio estimation procedure for the remaining seven SMSA's.

Prior to the implementation of the ratio estimation procedures, 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 nonpermit 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 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.

The next ratio estimation procedure was applied in the Boston, Mass.; Detroit, Mich.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's. This procedure was utilized 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 ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

The next ratio estimation procedure was applied in 10 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1981 housing inventory in each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1981 housing unit inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

The independent estimates of total housing units that were used as the numerator of this ratio are described below. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight.

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for each of the 15 1981 SMSA's. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1981 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units after the permit new construction ratio estimation (i.e., the existing sample estimate) were compared to the corresponding independent estimates for the central city, balance, and total SMSA for each of the 15 SMSA's and the estimate which showed the most likely level of net growth since the 1980 census in both the central city and balance as well as the total SMSA were used in this ratio estimation. As a result of this analysis, these independent estimates were used in the Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's.

The remaining five SMSA's used a combination of these independent estimates and sample estimates. For the Boston, Mass.; and Minneapolis-St. Paul, Minn., SMSA's, the sample estimate was used for the total SMSA. For the sectors, the sample estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the independent estimates.

For the Orlando Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, the independent estimate of units was used for the total SMSA. For the sectors, the independent estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the sample estimates. Due to the central city

boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimate did not reflect the 1970 central city/balance SMSA definition. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definition were used.

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 in these 10 SMSA's, and the resulting product was used as the final weight for tabulation purposes. For the other five SMSA's, the sample estimates were used as the final weight for tabulation purposes. These SMSA's included Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Detroit, Mich.; Fort Worth, Tex.; and Madison, Wis., SMSA's.

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

1977-1981 lost housing units—The 1977-1981 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1977 housing inventory for 14 of the SMSA's and the corresponding two-stage ratio estimation procedure used for the Pittsburgh, Pa., SMSA as was described in the 1977 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1977-1981 lost housing units existed, by definition, in the 1977 housing inventory, there was a 1977 housing inventory weight associated with each 1977-1981 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1977-1981 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 two-stage ratio estimation process for the Pittsburgh, Pa., SMSA and a one-stage ratio estimation process for the remaining SMSA's. 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; 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 1981 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, and the results 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) 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, the 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 presented.

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 1981 survey, the number of missed housing units may be considerably less for 1981.

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 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 1981 housing inventory as well as estimates of characteristics of the 1977-1981 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 32,009 for the total SMSA, 14,659 for the central city of the SMSA, and 28,390 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 1981 housing inventory as well as estimated percentages of the 1977-1981 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 1981 there were 580,800 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 6,930. 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	7,280
580,800	x
600,000	6,850

The entry of "x" is determined as follows by vertically interpolating between 7,280 and 6,850.

$$580{,}800 - 500{,}000 = 80{,}800$$
$$600{,}000 - 500{,}000 = 100{,}000$$

$$7{,}280 + \frac{80{,}800}{100{,}000}(6{,}850 - 7{,}280) = 6{,}930$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 573,870 to 587,730 housing units. Therefore, a conclusion that the average estimate of 1981 owner-occupied housing units (derived from all possible samples) 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 569,710 to 591,890 housing units with 90 percent confidence; and that the average estimate lies within the interval from 566,940 to 594,660 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 580,800 owner-occupied housing units, 169,800, or 29.2 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 29.2 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	29.2	50
500,000	1.0	a	1.1
580,000		p	
600,000	0.9	b	1.0

1. The entry for cell "a" is determined by horizontal interpolation between 1.0 and 1.1.

$$29.2 - 25.0 = 4.2$$
$$50.0 - 25.0 = 25.0$$

$$1.0 + \frac{4.2}{25.0}(1.1 - 1.0) = 1.0$$

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

$$29.2 - 25.0 = 4.2$$
$$50.0 - 25.0 = 25.0$$

$$0.9 + \frac{4.2}{25.0}(1.0 - 0.9) = 0.9$$

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

$$580{,}800 - 500{,}000 = 80{,}800$$
$$600{,}000 - 500{,}000 = 100{,}000$$

$$1.0 + \frac{80{,}800}{100{,}000}(0.9 - 1.0) = 0.9$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 28.3 to 30.1 percent; the 90-percent confidence interval is from 27.8 to 30.6 percent; and the 95-percent confidence interval is from 27.4 to 31.0 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The

standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. If there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1977 and 1981 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1977 and 1981 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1981 there were 291,400 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 121,600. Table I shows the standard error of 169,800 is approximately 5,790, and the standard error of 291,400 is approximately 6,900. Therefore, the standard error of the estimated difference of 121,600 is about 9,010.

$$9{,}010 = \sqrt{(5{,}790)^2 + (6{,}900)^2}$$

Consequently, the 68-percent confidence interval for the 121,600 difference is from 112,590 to 130,610 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 107,180 to 136,020 housing units, and the 95-percent confidence interval is from 103,580 to 139,620 housing units. Thus, we can conclude with 95 percent confidence that the number of 1981 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied 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 to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.6. The base of the distribution from which this median was determined is 580,800 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 580,800 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 274,400 housing units or 47.2 percent fall below this interval, and 116,000 housing units or 20.0 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(48.0-47.2)}{20.0} = 2.5$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.0 percent derived in step 2. About 274,400 housing units or 47.2 percent fall below this interval, and 116,000 housing units or 20.0 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(52.0-47.2)}{20.0} = 2.7$$

Thus, the 95-percent confidence interval ranges from 2.5 to 2.7 persons. Although it appears that this confidence interval has the same estimate as the lower limit it actually is a reflection of the rounding error associated with this median (see the paragraph on rounding error in the nonsampling error section of this appendix).

TABLE I. **Standard Errors for Estimated Number of Housing Units in the 1981 Housing Inventory and for Estimated Number of 1977-1981 Lost Units for the Pittsburgh, Pa., 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	240	230	240	150,000	5,520	2,400	5,370
100	240	230	240	179,800	5,920	—	5,710
200	240	230	240	200,000	6,150	—	5,910
500	350	340	350	250,000	6,620	—	6,270
700	410	400	410	300,000	6,960	—	6,480
1,000	490	480	490	400,000	7,310	—	6,490
2,500	780	760	780	500,000	7,280	—	5,930
5,000	1,100	1,060	1,100	600,000	6,850	—	4,610
10,000	1,560	1,480	1,550	700,000	5,940	—	600
25,000	2,440	2,230	2,430	701,500	5,920	—	—
50,000	3,400	2,890	3,370	800,000	4,280	—	—
75,000	4,100	3,180	4,050	881,300	—	—	—
100,000	4,660	3,210	4,580				

[1]For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero where the estimate for the total SMSA was derived from an independent estimate.

TABLE II. **Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Pittsburgh, Pa., 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	71.0	71.0	71.0	71.0	71.0	78.3	100,000	0.2	0.5	1.1	1.5	2.1	2.5
200	55.1	55.1	55.1	55.1	55.1	55.3	150,000	0.2	0.4	0.9	1.2	1.7	2.0
500	32.9	32.9	32.9	32.9	32.9	35.0	200,000	0.12	0.3	0.8	1.0	1.5	1.7
700	25.9	25.9	25.9	25.9	25.9	29.6	250,000	0.10	0.3	0.7	0.9	1.4	1.6
1,000	19.7	19.7	19.7	19.7	21.4	24.7	300,000	0.08	0.3	0.6	0.9	1.2	1.4
2,500	8.9	8.9	8.9	9.4	13.6	15.7	400,000	0.06	0.2	0.5	0.7	1.1	1.2
5,000	4.7	4.7	4.8	6.6	9.6	11.1	500,000	0.05	0.2	0.5	0.7	1.0	1.1
10,000	2.4	2.4	3.4	4.7	6.8	7.8	600,000	0.04	0.2	0.4	0.6	0.9	1.0
25,000	1.0	1.0	2.2	3.0	4.3	4.9	700,000	0.03	0.2	0.4	0.6	0.8	0.9
50,000	0.5	0.7	1.5	2.1	3.0	3.5	800,000	0.03	0.2	0.4	0.5	0.8	0.9
75,000	0.3	0.6	1.2	1.7	2.5	2.9	900,000	0.03	0.2	0.4	0.5	0.7	0.8

[1]Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2.

Spokane 1981

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN . . . App-42
Annual Housing Survey . . . App-42
Designation of sample housing units for the 1981 survey . . . App-42
Selection of the original AHS-SMSA sample . . . App-43
1977-1981 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-45
1970 Census of Population and Housing . . . App-45
ESTIMATION . . . App-45
1981 housing inventory . . . App-45
1977-1981 lost housing units . . . App-47
1977 estimation procedure . . . App-47
Ratio estimation procedure of the 1970 Census of Population and Housing . . . App-47
RELIABILITY OF THE ESTIMATES . . . App-47
Nonsampling errors . . . App-47
1970 census . . . App-47
AHS-SMSA . . . App-47
Coverage errors . . . App-47
Rounding errors . . . App-48
Sampling errors for the AHS-SMSA sample . . . App-48
Illustration of the use of the standard error tables . . . App-49
Differences . . . App-49
Illustration of the computation of the standard error of a difference . . . App-50
Medians . . . App-50
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-170-81) are based on data collected from the 1981 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.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1981 were interviewed previously in 1974 (1975 for Madison, Wis.) and 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

The three largest SMSA's (Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va.) in the 1981 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1974 and 1977 surveys evenly divided between the central city and the balance of the respective SMSA. The 12 remaining SMSA's (Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.) in the 1981 group were represented by a sample of 5,000 designated housing units in the 1974 (1975 for Madison, Wis.) and 1977 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1974 survey, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panel 3 was dropped from all SMSA's for the 1977 survey. Due to additional budget limitations for the 1981 survey, it also became necessary to drop panels 1, 2, 4, 5, 9, 11, and 12 from the Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va., SMSA's and panels 1 and 2 from the other 12 SMSA's. The designated sample size was reduced by two-thirds for the three largest SMSA's and by one-fourth for the other 12 SMSA's.

The interviewing was done for the three largest SMSA's during the period June 1981 through November 1981 with one-sixth of the sample housing units being visited each month. The interviewing was done for the remaining 12 SMSA's during the period April 1981 through December 1981 with one-ninth of the designated housing units visited each month.

In this SMSA, 4,402 housing units were eligible for interview. Of these sample housing units, 127 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 4,402 housing units eligible for interview, 337 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 1981 survey—The sample housing units designated to be interviewed in the 1981

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 and remained in sample after the 1981 panel reduction.
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 and remained in sample after the 1981 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1981 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 and remained in sample after the 1981 panel reduction. (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 and remained in sample after the 1981 panel reduction. (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 1981 Coverage Improvement Programs. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the original 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 five SMSA's were 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 10 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 for the 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.

1977-1981 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 from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes 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. Houses that have been moved onto their present site since the 1970 census.
6. Mobile homes placed outside parks since the 1970 census or vacant at the time of the 1970 census.

Each of the 15 SMSA's was previously interviewed in 1974 (1975 for Madison, Wis.) and 1977. The Coverage Improvement Program was conducted as part of the 1977 AHS with the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., SMSA's receiving some updating and refining as part of the 1981 AHS. The following discussion applies to both the prior year 1977 and the 1981 coverage improvement procedures. For the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., 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 1981.

Coverage improvement for deficiency 1—A sample of new construction units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 15 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. 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 (regular AHS units), while sample units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage 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 Fort Worth, Tex., SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second-stage sampling. For the third stage, structures of size three or more units were divided into clusters of an expected size of two units and a sample of clusters was selected. This procedure was employed in the following SMSA's: Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Orlando, Fla.; Spokane, Wash.; and Wichita, Kans.

In the remaining 9 SMSA's, 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, this procedure was also used in parts of the Anaheim-Santa Ana-Garden Grove, Calif., SMSA to supplement the sample described above. These procedures added an estimated 499 new construction 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 placed in a park that was 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 units represented the same number of units that the regular AHS sample units represented. This procedure added an estimated 405 units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed 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 units converted to residential use.
3. Mobile homes 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 a usual residence elsewhere.
4. Houses that had been moved onto their present site since the 1970 census.

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample 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 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 second procedure was designed to represent missed units from structures represented in the AHS. These missed units were:

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential 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 structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply both procedures during the 1977 Coverage Improvement Program to all SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in the Anaheim-Santa Ana-Garden Grove, Calif.; and Spokane, Wash., SMSA's. These SMSA's did not receive application of the second procedure. The Dallas, Tex.; Minneapolis-St. Paul, Minn.; and Pittsburgh, Pa., SMSA's did not receive application of either procedure during the 1977 Coverage Improvement Program. The Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and the Spokane, Wash., SMSA's had the first procedure completed as part of the 1981 Coverage Improvement Program. The remaining 10 1981 SMSA's received the full application of both procedures as part of the 1977 Coverage Improvement Program.

The first procedure added an estimated 2,670 housing units to the coverage of the housing inventory of 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 1981 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 1981 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1977 (i.e., 1977-1981 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1981 housing inventory—The AHS estimates of characteristics of the 1981 housing inventory were produced using a one-stage ratio estimation procedure for the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; and Madison, Wis., SMSA's, a two-stage ratio estimation procedure for the Detroit, Mich.; Orlando, Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, and a three-stage ratio estimation procedure for the remaining seven SMSA's.

Prior to the implementation of the ratio estimation procedures, 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 nonpermit 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 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.

The next ratio estimation procedure was applied in the Boston, Mass.; Detroit, Mich.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's. This procedure was utilized 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 ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

The next ratio estimation procedure was applied in 10 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1981 housing inventory in each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1981 housing unit inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

The independent estimates of total housing units that were used as the numerator of this ratio are described below. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight.

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for each of the 15 1981 SMSA's. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1981 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units after the permit new construction ratio estimation (i.e., the existing sample estimate) were compared to the corresponding independent estimates for the central city, balance, and total SMSA for each of the 15 SMSA's and the estimate which showed the most likely level of net growth since the 1980 census in both the central city and balance as well as the total SMSA were used in this ratio estimation. As a result of this analysis, these independent estimates were used in the Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's.

The remaining five SMSA's used a combination of these independent estimates and sample estimates. For the Boston, Mass.; and Minneapolis-St. Paul, Minn., SMSA's, the sample estimate was used for the total SMSA. For the sectors, the sample estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the independent estimates.

For the Orlando Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, the independent estimate of units was used for the total SMSA. For the sectors, the independent estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the sample estimates. Due to the central city

boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimate did not reflect the 1970 central city/balance SMSA definition. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definition were used.

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 in these 10 SMSA's, and the resulting product was used as the final weight for tabulation purposes. For the other five SMSA's, the sample estimates were used as the final weight for tabulation purposes. These SMSA's included Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Detroit, Mich.; Fort Worth, Tex.; and Madison, Wis., SMSA's.

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

1977-1981 lost housing units—The 1977-1981 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1977 housing inventory for 14 of the SMSA's and the corresponding two-stage ratio estimation procedure used for the Pittsburgh, Pa., SMSA as was described in the 1977 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas*. Since the 1977-1981 lost housing units existed, by definition, in the 1977 housing inventory, there was a 1977 housing inventory weight associated with each 1977-1981 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1977-1981 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 two-stage ratio estimation process for the Pittsburgh, Pa., SMSA and a one-stage ratio estimation process for the remaining SMSA's. 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; 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 1981 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, and the results 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) 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, the 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 presented.

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 1981 survey, the number of missed housing units may be considerably less for 1981.

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 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 1981 housing inventory as well as estimates of characteristics of the 1977-1981 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 5,238 for the total SMSA, 2,963 for the central city of the SMSA, and 4,319 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 1981 housing inventory as well as estimated percentages of the 1977-1981 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 1981 there were 87,700 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 1,040. 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
77,000	1,090
87,700	x
100,000	990

The entry of "x" is determined as follows by vertically interpolating between 1,090 and 990.

$$87{,}700 - 77{,}000 = 10{,}700$$
$$100{,}000 - 77{,}000 = 23{,}000$$

$$1{,}090 + \frac{10{,}700}{23{,}000}(990 - 1{,}090) = 1{,}040$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 86,660 to 88,740 housing units. Therefore, a conclusion that the average estimate of 1981 owner-occupied housing units (derived from all possible samples) 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 86,040 to 89,360 housing units with 90 percent confidence; and that the average estimate lies within the interval from 85,620 to 89,780 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 87,700 owner-occupied housing units, 24,200, or 27.6 percent, had two bedrooms. Interpolation using table II of this appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 27.6 percent is approximately 0.8 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	25 or 75	27.6	50
75,000	0.9	a	1.1
87,700		p	
100,000	0.8	b	0.9

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

$$27.6 - 25.0 = 2.6$$
$$50.0 - 25.0 = 25.0$$

$$0.9 + \frac{2.6}{25.0}(1.1 - 0.9) = 0.9$$

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

$$27.6 - 25.0 = 2.6$$
$$50.0 - 25.0 = 25.0$$

$$0.8 + \frac{2.6}{25.0}(0.9 - 0.8) = 0.8$$

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

$$87{,}700 - 75{,}000 = 12{,}700$$
$$100{,}000 - 75{,}000 = 25{,}000$$

$$0.9 + \frac{12{,}700}{25{,}000}(0.8 - 0.9) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 26.8 to 28.4 percent; the 90-percent confidence interval is from 26.3 to 28.9 percent; and the 95-percent confidence interval is from 26.0 to 29.2 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 1977 and 1981 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1977 and 1981 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1981 there were 36,500 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 12,300. Table I shows the standard error of 24,200 is approximately 830, and the standard error of 36,500 is approximately 940. Therefore, the standard error of the estimated difference of 12,300 is about 1,250.

$$1{,}250 = \sqrt{(830)^2 + (940)^2}$$

Consequently, the 68-percent confidence interval for the 12,300 difference is from 11,050 to 13,550 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 10,300 to 14,300 housing units, and the 95-percent confidence interval is from 9,800 to 14,800 housing units. Thus, we can conclude with 95 percent confidence that the number of 1981 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied 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 to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.5. The base of the distribution from which this median was determined is 87,700 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 87,700 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 14,700 housing units or 16.8 percent fall below this interval, and 30,000 housing units or 34.2 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$1.5 + (2.5-1.5)\ \frac{(48.0-16.8)}{34.2} = 2.4$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.0 percent derived in step 2. About 44,700 housing units or 51.0 percent fall below this interval, and 14,100 housing units or 16.1 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(52.0-51.0)}{16.1} = 2.6$$

Thus, the 95-percent confidence interval ranges from 2.4 to 2.6 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1981 Housing Inventory and for Estimated Number of 1977-1981 Lost Units for the Spokane, Wash., SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	30	30	30	10,000	560	530	540
100	60	60	60	25,000	840	740	720
200	80	80	80	50,000	1,050	750	590
500	130	130	130	62,700	1,090	610	—
700	150	150	160	75,000	1,090	250	—
1,000	180	180	190	77,000	1,090	—	—
2,500	290	280	290	100,000	990	—	—
5,000	410	390	400	139,700	—	—	—

[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 and 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 where the estimate for the total SMSA was derived from an independent estimate.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Spokane, Wash., SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						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	25.8	25.8	25.8	25.8	25.8	29.5	10,000	0.3	0.6	1.3	1.8	2.6	2.9
200	14.8	14.8	14.8	14.8	18.1	20.9	25,000	0.14	0.4	0.8	1.1	1.6	1.9
500	6.5	6.5	6.5	7.9	11.4	13.2	50,000	0.07	0.3	0.6	0.8	1.1	1.3
700	4.7	4.7	4.9	6.7	9.7	11.1	75,000	0.05	0.2	0.5	0.6	0.9	1.1
1,000	3.4	3.4	4.1	5.6	8.1	9.3	100,000	0.03	0.2	0.4	0.6	0.8	0.9
2,500	1.4	1.4	2.6	3.5	5.1	5.9	150,000	0.02	0.2	0.3	0.5	0.7	0.8
5,000	0.7	0.8	1.8	2.5	3.6	4.2							

[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. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2. for the total SMSA and 1.1 for the central city and for the balance (not in central city).

Tacoma 1981

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN . . . App-42
Annual Housing Survey . . . App-42
Designation of sample housing units for the 1981 survey . . . App-42
Selection of the original AHS-SMSA sample . . . App-43
1977-1981 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-45
1970 Census of Population and Housing . . . App-45
ESTIMATION . . . App-45
1981 housing inventory . . . App-45
1977-1981 lost housing units . . . App-47
1977 estimation procedure . . . App-47
Ratio estimation procedure of the 1970 Census of Population and Housing . . . App-47
RELIABILITY OF THE ESTIMATES . . . App-47
Nonsampling errors . . . App-47
1970 census . . . App-47
AHS-SMSA . . . App-47
Coverage errors . . . App-47
Rounding errors . . . App-48
Sampling errors for the AHS-SMSA sample . . . App-48
Illustration of the use of the standard error tables . . . App-49
Differences . . . App-49
Illustration of the computation of the standard error of a difference . . . App-50
Medians . . . App-50
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-170-81) are based on data collected from the 1981 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.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1981 were interviewed previously in 1974 (1975 for Madison, Wis.) and 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

The three largest SMSA's (Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va.) in the 1981 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1974 and 1977 surveys evenly divided between the central city and the balance of the respective SMSA. The 12 remaining SMSA's (Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.) in the 1981 group were represented by a sample of 5,000 designated housing units in the 1974 (1975 for Madison, Wis.) and 1977 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1974 survey, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panel 3 was dropped from all SMSA's for the 1977 survey. Due to additional budget limitations for the 1981 survey, it also became necessary to drop panels 1, 2, 4, 5, 9, 11, and 12 from the Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va., SMSA's and panels 1 and 2 from the other 12 SMSA's. The designated sample size was reduced by two-thirds for the three largest SMSA's and by one-fourth for the other 12 SMSA's.

The interviewing was done for the three largest SMSA's during the period June 1981 through November 1981 with one-sixth of the sample housing units being visited each month. The interviewing was done for the remaining 12 SMSA's during the period April 1981 through December 1981 with one-ninth of the designated housing units visited each month.

In this SMSA, 4,636 housing units were eligible for interview. Of these sample housing units, 188 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 4,636 housing units eligible for interview, 225 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 1981 survey—The sample housing units designated to be interviewed in the 1981

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 and remained in sample after the 1981 panel reduction.
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 and remained in sample after the 1981 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1981 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 and remained in sample after the 1981 panel reduction. (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 and remained in sample after the 1981 panel reduction. (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 1981 Coverage Improvement Programs. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the original 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 five SMSA's were 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 10 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 for the 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.

1977-1981 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 from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes 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. Houses that have been moved onto their present site since the 1970 census.
6. Mobile homes placed outside parks since the 1970 census or vacant at the time of the 1970 census.

Each of the 15 SMSA's was previously interviewed in 1974 (1975 for Madison, Wis.) and 1977. The Coverage Improvement Program was conducted as part of the 1977 AHS with the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., SMSA's receiving some updating and refining as part of the 1981 AHS. The following discussion applies to both the prior year 1977 and the 1981 coverage improvement procedures. For the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., 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 1981.

Coverage improvement for deficiency 1—A sample of new construction units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 15 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. 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 (regular AHS units), while sample units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage 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 Fort Worth, Tex., SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second-stage sampling. For the third stage, structures of size three or more units were divided into clusters of an expected size of two units and a sample of clusters was selected. This procedure was employed in the following SMSA's: Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Orlando, Fla.; Spokane, Wash.; and Wichita, Kans.

In the remaining 9 SMSA's, 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, this procedure was also used in parts of the Anaheim-Santa Ana-Garden Grove, Calif., SMSA to supplement the sample described above. These procedures added an estimated 2,413 new construction 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 placed in a park that was 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 units represented the same number of units that the regular AHS sample units represented. This procedure added an estimated 659 units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed 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 units converted to residential use.
3. Mobile homes 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 a usual residence elsewhere.
4. Houses that had been moved onto their present site since the 1970 census.

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample 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 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 second procedure was designed to represent missed units from structures represented in the AHS. These missed units were:

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential 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 structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply both procedures during the 1977 Coverage Improvement Program to all SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in the Anaheim-Santa Ana-Garden Grove, Calif.; and Spokane, Wash., SMSA's. These SMSA's did not receive application of the second procedure. The Dallas, Tex.; Minneapolis-St. Paul, Minn.; and Pittsburgh, Pa., SMSA's did not receive application of either procedure during the 1977 Coverage Improvement Program. The Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and the Spokane, Wash., SMSA's had the first procedure completed as part of the 1981 Coverage Improvement Program. The remaining 10 1981 SMSA's received the full application of both procedures as part of the 1977 Coverage Improvement Program.

The first procedure added an estimated 7,382 housing units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 246 housing units.

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 1981 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 1981 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1977 (i.e., 1977-1981 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1981 housing inventory—The AHS estimates of characteristics of the 1981 housing inventory were produced using a one-stage ratio estimation procedure for the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; and Madison, Wis., SMSA's, a two-stage ratio estimation procedure for the Detroit, Mich.; Orlando, Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, and a three-stage ratio estimation procedure for the remaining seven SMSA's.

Prior to the implementation of the ratio estimation procedures, 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 nonpermit 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 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.

The next ratio estimation procedure was applied in the Boston, Mass.; Detroit, Mich.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's. This procedure was utilized 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 ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

The next ratio estimation procedure was applied in 10 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1981 housing inventory in each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1981 housing unit inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

The independent estimates of total housing units that were used as the numerator of this ratio are described below. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight.

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for each of the 15 1981 SMSA's. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1981 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units after the permit new construction ratio estimation (i.e., the existing sample estimate) were compared to the corresponding independent estimates for the central city, balance, and total SMSA for each of the 15 SMSA's and the estimate which showed the most likely level of net growth since the 1980 census in both the central city and balance as well as the total SMSA were used in this ratio estimation. As a result of this analysis, these independent estimates were used in the Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's.

The remaining five SMSA's used a combination of these independent estimates and sample estimates. For the Boston, Mass.; and Minneapolis-St. Paul, Minn., SMSA's, the sample estimate was used for the total SMSA. For the sectors, the sample estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the independent estimates.

For the Orlando Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, the independent estimate of units was used for the total SMSA. For the sectors, the independent estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the sample estimates. Due to the central city

boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimate did not reflect the 1970 central city/balance SMSA definition. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definition were used.

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 in these 10 SMSA's, and the resulting product was used as the final weight for tabulation purposes. For the other five SMSA's, the sample estimates were used as the final weight for tabulation purposes. These SMSA's included Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Detroit, Mich.; Fort Worth, Tex.; and Madison, Wis., SMSA's.

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

1977-1981 lost housing units—The 1977-1981 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1977 housing inventory for 14 of the SMSA's and the corresponding two-stage ratio estimation procedure used for the Pittsburgh, Pa., SMSA as was described in the 1977 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1977-1981 lost housing units existed, by definition, in the 1977 housing inventory, there was a 1977 housing inventory weight associated with each 1977-1981 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1977-1981 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 two-stage ratio estimation process for the Pittsburgh, Pa., SMSA and a one-stage ratio estimation process for the remaining SMSA's. 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; 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 1981 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, and the results 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) 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, the 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 presented.

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 1981 survey, the number of missed housing units may be considerably less for 1981.

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 standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. Approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the average result of all possible samples.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimate would include the average result of all possible samples.

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

The figures presented in the tables that follow (pages App-51 and App-52) 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 1981 housing inventory as well as estimates of characteristics of the 1977-1981 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 5,989 for the total SMSA, 3,661 for the central city of the SMSA, and 4,740 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.

Tables II and III (pages App-51 and App-52) present the standard errors of estimated percentages for the 1981 housing inventory as well as estimated percentages of the 1977-1981 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 1981 there were 111,700 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 1,550. The following interpolation procedure was used.

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

Size of estimate	Standard error
100,000	1,580
111,700	x
122,800	1,520

The entry of "x" is determined as follows by vertically interpolating between 1,580 and 1,520.

$$111{,}700 - 100{,}000 = 11{,}700$$
$$122{,}800 - 100{,}000 = 22{,}800$$

$$1{,}580 + \frac{11{,}700}{22{,}800}(1{,}520 - 1{,}580) = 1{,}550$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 110,150 to 113,250 housing units. Therefore, a conclusion that the average estimate of 1981 owner-occupied housing units (derived from all possible samples) 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 109,220 to 114,180 housing units with 90 percent confidence; and that the average estimate lies within the interval from 108,600 to 114,800 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 111,700 owner-occupied housing units, 27,800, or 24.9 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 24.9 percent is approximately 1.0 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	24.9	25 or 75
100,000	0.7	a	1.0
111,700		p	
150,000	0.6	b	0.8

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

$$24.9 - 10.0 = 14.9$$
$$25.0 - 10.0 = 15.0$$

$$0.7 + \frac{14.9}{15.0}(1.0 - 0.7) = 1.0$$

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

$$24.9 - 10.0 = 14.9$$
$$25.0 - 10.0 = 15.0$$

$$0.6 + \frac{14.9}{15.0}(0.8 - 0.6) = 0.8$$

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

$$111{,}700 - 100{,}000 = 11{,}700$$
$$150{,}000 - 100{,}000 = 50{,}000$$

$$1.0 + \frac{11{,}700}{50{,}000}(0.8 - 1.0) = 1.0$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 23.9 to 25.9 percent; the 90-percent confidence interval is from 23.3 to 26.5 percent; and the 95-percent confidence interval is from 22.9 to 26.9 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The

standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. If there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1977 and 1981 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1977 and 1981 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1981 there were 57,100 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 29,300. Table I shows the standard error of 27,800 is approximately 1,110, and the standard error of 57,100 is approximately 1,440. Therefore, the standard error of the estimated difference of 29,300 is about 1,820.

$$1{,}820 = \sqrt{(1{,}110)^2 + (1{,}440)^2}$$

Consequently, the 68-percent confidence interval for the 29,300 difference is from 27,480 to 31,120 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 21,390 to 32,210 housing units, and the 95-percent confidence interval is from 25,660 to 32,940 housing units. Thus, we can conclude with 95 percent confidence that the number of 1981 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied 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 to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.6. The base of the distribution from which this median was determined is 111,700 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 111,700 is approximately 1.2 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 47.6 and 52.4.
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 47.6 percent derived in step 2. About 15,800 housing units or 14.1 percent fall below this interval, and 38,500 housing units or 34.5 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$1.5 + (2.5-1.5)\ \frac{(47.6-14.1)}{34.5} = 2.5$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.4 percent derived in step 2. About 54,300 housing units or 48.6 percent fall below this interval, and 21,100 housing units or 18.9 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(52.4-48.6)}{18.9} = 2.7$$

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 1981 Housing Inventory and for Estimated Number of 1977-1981 Lost Units for the Tacoma, Wash., SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]			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	1,070	810	1,040
100	70	60	70	50,000	1,390	750	1,270
200	100	90	100	69,200	1,520	—	1,280
500	160	140	160	75,000	1,550	—	1,260
700	190	170	190	100,000	1,580	—	1,010
1,000	230	200	230	122,800	1,520	—	—
2,500	360	310	370	150,000	1,310	—	—
5,000	510	440	510	192,000	—	—	—
10,000	700	590	710				

[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, 1.0 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 where the estimate for the total SMSA was derived from an independent estimate.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Tacoma, Wash., 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	35.3	35.3	35.3	35.3	35.3	36.9	10,000	0.5	0.7	1.6	2.2	3.2	3.7
200	21.4	21.4	21.4	21.4	22.6	26.1	25,000	0.2	0.5	1.0	1.4	2.0	2.3
500	9.8	9.8	9.8	9.9	14.3	16.5	50,000	0.11	0.3	0.7	1.0	1.4	1.7
700	7.2	7.2	7.2	8.4	12.1	13.9	75,000	0.07	0.3	0.6	0.8	1.2	1.3
1,000	5.2	5.2	5.2	7.0	10.1	11.7	100,000	0.05	0.2	0.5	0.7	1.0	1.2
2,500	2.1	2.1	3.2	4.4	6.4	7.4	150,000	0.04	0.2	0.4	0.6	0.8	1.0
5,000	1.1	1.1	2.3	3.1	4.5	5.2	200,000	0.03	0.2	0.4	0.5	0.7	0.8

[1]Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2.

TABLE III. Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Central City of the Tacoma, Wash., 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	29.0	29.0	29.0	29.0	29.0	31.9	5,000	0.8	0.9	2.0	2.7	3.9	4.5
200	16.9	16.9	16.9	16.9	19.6	22.6	10,000	0.4	0.6	1.4	1.9	2.8	3.2
500	7.5	7.5	7.5	8.6	12.4	14.3	25,000	0.2	0.4	0.9	1.2	1.7	2.0
700	5.5	5.5	5.5	7.2	10.5	12.1	50,000	0.08	0.3	0.6	0.9	1.2	1.4
1,000	3.9	3.9	4.4	6.1	8.7	10.1	75,000	0.05	0.2	0.5	0.7	1.0	1.2
2,500	1.6	1.6	2.8	3.8	5.5	6.4							

[1]Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point.

Washington 1981

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1981 survey App-42
Selection of the original AHS-SMSA sample App-43
1977-1981 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 3-8 ..... App-45
1970 Census of Population and Housing App-45
ESTIMATION App-45
1981 housing inventory App-45
1977-1981 lost housing units ... App-47
1977 estimation procedure App-47
Ratio estimation procedure of the 1970 Census of Population and Housing App-47
RELIABILITY OF THE ESTIMATES .. App-47
Nonsampling errors App-47
1970 census App-47
AHS-SMSA App-47
Coverage errors App-47
Rounding errors App-48
Sampling errors for the AHS-SMSA sample App-48
Illustration of the use of the standard error tables App-49
Differences App-49
Illustration of the computation of the standard error of a difference App-50
Medians App-50
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-170-81) are based on data collected from the 1981 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.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1981 were interviewed previously in 1974 and 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

The three largest SMSA's (Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va.) in the 1981 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1974 and 1977 surveys evenly divided between the central city and the balance of the respective SMSA. The 12 remaining SMSA's (Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Forth Worth, Tex.; Madison, Wis.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.) in the 1981 group were represented by a sample of 5,000 designated housing units in the 1974 and 1977 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1974 survey, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panel 3 was dropped from all SMSA's for the 1977 survey. Due to additional budget limitations for the 1981 survey, it also became necessary to drop panels 1, 2, 4, 5, 9, 11, and 12 from the Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va., SMSA's and panels 1 and 2 from the other 12 SMSA's. The designated sample size was reduced by two-thirds for the three largest SMSA's and by one-fourth for the other 12 SMSA's.

The interviewing was done for the three largest SMSA's during the period June 1981 through November 1981 with one-sixth of the sample housing units being visited each month. The interviewing was done for the remaining 12 SMSA's during the period April 1981 through December 1981 with one-ninth of the designated housing units visited each month.

In this SMSA, 5,288 housing units were eligible for interview. Of these sample housing units, 269 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 5,288 housing units eligible for interview, 509 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 1981 survey—The sample housing units designated to be interviewed in the 1981

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 and remained in sample after the 1981 panel reduction.
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 and remained in sample after the 1981 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1981 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 and remained in sample after the 1981 panel reduction. (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 and remained in sample after the 1981 panel reduction. (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 1981 Coverage Improvement Programs. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the original 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 five SMSA's were 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 10 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 for the 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.

1977-1981 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 from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes 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. Houses that have been moved onto their present site since the 1970 census.
6. Mobile homes placed outside parks since the 1970 census or vacant at the time of the 1970 census.

Each of the 15 SMSA's was previously interviewed in 1974 and 1977. The Coverage Improvement Program was conducted as part of the 1977 AHS with the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., SMSA's receiving some updating and refining as part of the 1981 AHS. The following discussion applies to both the prior year 1977 and the 1981 coverage improvement procedures. For the the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., 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 1981.

Coverage improvement for deficiency 1—A sample of new construction units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 15 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. 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 (regular AHS units), while sample units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage 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 Fort Worth, Tex., SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second-stage sampling. For the third stage, structures of size three or more units were divided into clusters of an expected size of two units and a sample of clusters was selected. This procedure was employed in the following SMSA's: Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Orlando, Fla.; Spokane, Wash.; and Wichita, Kans.

In the remaining 9 SMSA's, 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, this procedure was also used in parts of the Anaheim-Santa Ana Garden Grove, Calif., SMSA to supplement the sample described above. These procedures added an estimated 9,726 new construction 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 placed in a park that was 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 units represented the same number of units that the regular AHS sample units represented. This procedure added an estimated 1,184 units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiences 3-6—The remaining missed 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 units converted to residential use.
3. Mobile homes 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 a usual residence elsewhere.
4. Houses that had been moved onto their present site since the 1970 census.

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample 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 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 second procedure was designed to represent missed units from structures represented in the AHS. These missed units were:

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential 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 structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply both procedures during the 1977 Coverage Improvement Program to all SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in the Anaheim-Santa Ana-Garden Grove, Calif.; and Spokane, Wash., SMSA's. These SMSA's did not receive application of the second procedure. The Dallas, Tex.; Minneapolis-St. Paul, Minn.; and Pittsburgh, Pa., SMSA's did not receive application of either procedure during the 1977 Coverage Improvement Program. The Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and the Spokane, Wash., SMSA's had the first procedure completed as part of the 1981 Coverage Improvement Program. The remaining 10 1981 SMSA's received the full application of both procedures as part of the 1977 Coverage Improvement Program.

The first procedure added an estimated 22,391 housing units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 842 housing units.

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 1981 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 1981 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1977 (i.e., 1977-1981 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1981 housing inventory—The AHS estimates of characteristics of the 1981 housing inventory were produced using a one-stage ratio estimation procedure for the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; and Madison, Wis., SMSA's, a two-stage ratio estimation procedure for the Detroit, Mich.; Orlando, Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, and a three-stage ratio estimation procedure for the remaining seven SMSA's.

Prior to the implementation of the ratio estimation procedures, 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 nonpermit 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 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.

The next ratio estimation procedure was applied in the Boston, Mass.; Detroit, Mich.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's. This procedure was utilized 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 ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

The next ratio estimation procedure was applied in 10 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1981 housing inventory in each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1981 housing unit inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

The independent estimates of total housing units that were used as the numerator of this ratio are described below. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight.

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for each of the 15 1981 SMSA's. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1981 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units after the permit new construction ratio estimation (i.e., the existing sample estimate) were compared to the corresponding independent estimates for the central city, balance, and total SMSA for each of the 15 SMSA's and the estimate which showed the most likely level of net growth since the 1980 census in both the central city and balance as well as the total SMSA were used in this ratio estimation. As a result of this analysis, these independent estimates were used in the Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's.

The remaining five SMSA's used a combination of these independent estimates and sample estimates. For the Boston, Mass.; and Minneapolis-St. Paul, Minn., SMSA's, the sample estimate was used for the total SMSA. For the sectors, the sample estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the independent estimates.

For the Orlando Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, the independent estimate of units was used for the total SMSA. For the sectors, the independent estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the sample estimates. Due to the central city

boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimate did not reflect the 1970 central city/balance SMSA definition. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definition were used.

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 in these 10 SMSA's, and the resulting product was used as the final weight for tabulation purposes. For the other five SMSA's, the sample estimates were used as the final weight for tabulation purposes. These SMSA's included Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Detroit, Mich.; Fort Worth, Tex.; and Madison, Wis., SMSA's.

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

1977-1981 lost housing units—The 1977-1981 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1977 housing inventory for 14 of the SMSA's and the corresponding two-stage ratio estimation procedure used for the Pittsburgh, Pa., SMSA as was described in the 1977 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas*. Since the 1977-1981 lost housing units existed, by definition, in the 1977 housing inventory, there was a 1977 housing inventory weight associated with each 1977-1981 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1977-1981 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 two-stage ratio estimation process for the Pittsburgh, Pa., SMSA and a one-stage ratio estimation process for the remaining SMSA's. 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; 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 1981 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, and the results 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) 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, the 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 presented.

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 1981 survey, the number of missed housing units may be considerably less for 1981.

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 standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.
2. Approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the average result of all possible samples.
3. Approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimate would include the average result of all possible samples.

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

The figures presented in the tables that follow (pages App-51 through App-52) 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 1981 housing inventory as well as estimates of characteristics of the 1977-1981 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 46,150 for the total SMSA, 13,660 for the central city of the SMSA, and 43,890 for the balance of the SMSA.

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

Tables II through IV (pages App-51 through App-52) present the standard errors of estimated percentages for the 1981 housing inventory as well as estimated percentages of the 1977-1981 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 through 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 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 1981 there were 596,300 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 8,770. 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	8,670
596,300	x
600,000	8,770

The entry of "x" is determined as follows by vertically interpolating between 8,670 and 8,770.

$$596{,}300-500{,}000=96{,}300$$
$$600{,}000-500{,}000=100{,}000$$

$$8{,}670+\frac{96{,}300}{100{,}000}(8{,}770-8{,}670)=8{,}770$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 587,530 to 605,070 housing units. Therefore, a conclusion that the average estimate of 1981 owner-occupied housing units (derived from all possible samples) 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 582,270 to 610,330 housing units with 90 percent confidence; and that the average estimate lies within the interval from 578,760 to 613,840 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 596,300 owner-occupied housing units, 84,500, or 14.2 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 14.2 percent is approximately 0.7 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	14.2	25 or 75
500,000	0.7	a	1.0
596,300		p	
600,000	0.6	b	0.9

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

$$14.2-10.0=4.2$$
$$25.0-10.0=15.0$$

$$0.7+\frac{4.2}{15.0}(1.0-0.7)=0.8$$

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

$$14.2-10.0=4.2$$
$$25.0-10.0=15.0$$

$$0.6+\frac{4.2}{15.0}(0.9-0.6)=0.7$$

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

$$596{,}300-500{,}000=96{,}300$$
$$600{,}000-500{,}000=100{,}000$$

$$0.8+\frac{96{,}300}{100{,}000}(0.7-0.8)=0.7$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 13.5 to 14.9 percent; the 90-percent confidence interval is from 13.1 to 15.3 percent; and the 95-percent confidence interval is from 12.8 to 15.6 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The

standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. If there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1977 and 1981 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1977 and 1981 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1981 there were 264,000 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 179,500. Table I shows the standard error of 84,500 is approximately 4,520, and the standard error of 264,000 is approximately 7,320. Therefore, the standard error of the estimated difference of 179,500 is about 8,600.

$$8{,}600 = \sqrt{(4{,}520)^2 + (7{,}320)^2}$$

Consequently, the 68-percent confidence interval for the 179,500 difference is from 170,900 to 188,100 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 165,740 to 193,260 housing units, and the 95-percent confidence interval is from 162,300 to 196,700 housing units. Thus, we can conclude with 95 percent confidence that the number of 1981 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied 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 through 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 into which interval of the distribution the lower percentage limit falls. Similarly, to find the upper endpoint to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.8. The base of the distribution from which this median was determined is 596,300 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 596,300 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 260,200 housing units or 43.6 percent fall below this interval, and 117,200 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:

$$2.5 + (3.5-2.5)\ \frac{(48.0-43.6)}{19.7} = 2.7$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.0 percent derived in step 2. About 260,000 housing units or 43.6 percent fall below this interval, and 117,200 housing units or 19.7 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(52.0-43.6)}{19.7} = 2.9$$

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

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1981 Housing Inventory and for Estimated Number of 1977-1981 Lost Units for the Washington, D.C.-Md.-Va., 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	260	130	330	200,000	6,590	2,770	7,180
100	260	130	330	250,000	7,180	1,900	7,730
200	260	160	330	280,400	7,480	—	7,990
500	360	260	410	300,000	7,650	—	8,130
700	430	310	480	400,000	8,310	—	8,570
1,000	510	360	580	500,000	8,670	—	8,560
2,500	810	580	910	600,000	8,770	—	8,100
5,000	1,140	810	1,280	700,000	8,610	—	7,120
10,000	1,610	1,130	1,810	800,000	8,180	—	5,310
25,000	2,530	1,740	2,840	895,300	7,480	—	—
50,000	3,540	2,340	3,960	900,000	7,430	—	—
75,000	4,290	2,710	4,770	1,000,000	6,250	—	—
100,000	4,890	2,930	5,430	1,100,000	4,310	—	—
150,000	5,850	3,050	6,440	1,175,700	—	—	—

[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, 1.1 for the characteristics pertaining to renters in the central city, 2.1 for all other characteristics in the central city, and 1.3 for the balance (not in central city) estimates. Standard errors of estimates pertaining to total housing units for each of the central city, balance, and total SMSA are assumed to be equal to zero where the estimate for the total SMSA was derived from an independent estimate.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Washington, D.C.-Md.-Va., 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	56.7	56.7	56.7	56.7	56.7	57.2	200,000	0.13	0.4	0.8	1.1	1.6	1.8
500	34.4	34.4	34.4	34.4	34.4	36.2	250,000	0.10	0.3	0.7	1.0	1.4	1.6
700	27.2	27.2	27.2	27.2	27.2	30.6	300,000	0.09	0.3	0.6	0.9	1.3	1.5
1,000	20.7	20.7	20.7	20.7	22.2	25.6	400,000	0.07	0.3	0.6	0.8	1.1	1.3
2,500	9.5	9.5	9.5	9.7	14.0	16.2	500,000	0.05	0.2	0.5	0.7	1.0	1.1
5,000	5.0	5.0	5.0	6.9	9.9	11.4	600,000	0.04	0.2	0.5	0.6	0.9	1.0
10,000	2.6	2.6	3.5	4.9	7.0	8.1	700,000	0.04	0.2	0.4	0.6	0.8	1.0
25,000	1.0	1.0	2.2	3.1	4.4	5.1	800,000	0.03	0.2	0.4	0.5	0.8	0.9
50,000	0.5	0.7	1.6	2.2	3.1	3.6	900,000	0.03	0.2	0.4	0.5	0.7	0.9
75,000	0.3	0.6	1.3	1.8	2.6	3.0	1,000,000	0.03	0.2	0.4	0.5	0.7	0.8
100,000	0.3	0.5	1.1	1.5	2.2	2.6	1,100,000	0.02	0.2	0.3	0.5	0.7	0.8
150,000	0.2	0.4	0.9	1.3	1.8	2.1	1,200,000	0.02	0.15	0.3	0.4	0.6	0.7

[1]Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.4.

TABLE III. Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Central City of the Washington, D.C.-Md.-Va., 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	57.2	57.2	57.2	57.2	57.2	57.8	25,000	0.5	0.7	1.6	2.2	3.2	3.7
200	40.0	40.0	40.0	40.0	40.0	40.9	50,000	0.3	0.5	1.1	1.6	2.2	2.6
500	21.1	21.1	21.1	21.1	22.4	25.8	75,000	0.2	0.4	0.9	1.3	1.8	2.1
700	16.0	16.0	16.0	16.0	18.9	21.8	100,000	0.13	0.4	0.8	1.1	1.6	1.8
1,000	11.8	11.8	11.8	11.8	15.8	18.3	150,000	0.09	0.3	0.7	0.9	1.3	1.5
2,500	5.1	5.1	5.1	6.9	10.0	11.6	200,000	0.07	0.3	0.6	0.8	1.1	1.3
5,000	2.6	2.6	3.6	4.9	7.1	8.2	250,000	0.05	0.2	0.5	0.7	1.0	1.2
10,000	1.3	1.3	2.5	3.5	5.0	5.8	300,000	0.04	0.2	0.5	0.6	0.9	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 1.1 for the characteristics pertaining to renters and 2.1 for all other characteristics.

TABLE IV. Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Balance (Not in Central City) of the Washington, D.C.-Md.-Va., 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	62.4	62.4	62.4	62.4	62.4	64.4	150,000	0.2	0.5	1.0	1.4	2.0	2.4
500	39.9	39.9	39.9	39.9	39.9	40.7	200,000	0.2	0.4	0.9	1.2	1.8	2.0
700	32.2	32.2	32.2	32.2	32.2	34.4	250,000	0.13	0.4	0.8	1.1	1.6	1.8
1,000	24.9	24.9	24.9	24.9	24.9	28.8	300,000	0.11	0.3	0.7	1.0	1.4	1.7
2,500	11.7	11.7	11.7	11.7	15.8	18.2	400,000	0.08	0.3	0.6	0.9	1.2	1.4
5,000	6.2	6.2	6.2	7.7	11.2	12.9	500,000	0.07	0.3	0.6	0.8	1.1	1.3
10,000	3.2	3.2	4.0	5.5	7.9	9.1	600,000	0.06	0.2	0.5	0.7	1.0	1.2
25,000	1.3	1.3	2.5	3.5	5.0	5.8	700,000	0.05	0.2	0.5	0.7	0.9	1.1
50,000	0.7	0.8	1.8	2.4	3.5	4.1	800,000	0.04	0.2	0.4	0.6	0.9	1.0
75,000	0.4	0.7	1.4	2.0	2.9	3.3	900,000	0.04	0.2	0.4	0.6	0.8	1.0
100,000	0.3	0.6	1.3	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 1.3.

Wichita 1981

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1981 survey App-42
Selection of the original AHS-SMSA sample App-43
1977-1981 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-45
1970 Census of Population and Housing App-45
ESTIMATION App-45
1981 housing inventory App-45
1977-1981 lost housing units ... App-47
1977 estimation procedure App-47
Ratio estimation procedure of the 1970 Census of Population and Housing App-47
RELIABILITY OF THE ESTIMATES .. App-47
Nonsampling errors App-47
1970 census App-47
AHS-SMSA App-47
Coverage errors App-47
Rounding errors App-48
Sampling errors for the AHS-SMSA sample App-48
Illustration of the use of the standard error tables App-49
Differences App-49
Illustration of the computation of the standard error of a difference App-50
Medians App-50
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-170-81) are based on data collected from the 1981 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.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 15 SMSA's selected for interview during 1981 were interviewed previously in 1974 (1975 for Madison, Wis.) and 1977 (see the list of SMSA reports from the AHS in the introduction of this report).

The three largest SMSA's (Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va.) in the 1981 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1974 and 1977 surveys evenly divided between the central city and the balance of the respective SMSA. The 12 remaining SMSA's (Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.) in the 1981 group were represented by a sample of 5,000 designated housing units in the 1974 (1975 for Madison, Wis.) and 1977 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1974 survey, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panel 3 was dropped from all SMSA's for the 1977 survey. Due to additional budget limitations for the 1981 survey, it also became necessary to drop panels 1, 2, 4, 5, 9, 11, and 12 from the Boston, Mass.; Detroit, Mich.; and Washington, D.C.-Md.-Va., SMSA's and panels 1 and 2 from the other 12 SMSA's. The designated sample size was reduced by two-thirds for the three largest SMSA's and by one-fourth for the other 12 SMSA's.

The interviewing was done for the three largest SMSA's during the period June 1981 through November 1981 with one-sixth of the sample housing units being visited each month. The interviewing was done for the remaining 12 SMSA's during the period April 1981 through December 1981 with one-ninth of the designated housing units visited each month.

In this SMSA, 4,339 housing units were eligible for interview. Of these sample housing units, 103 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 4,339 housing units eligible for interview, 229 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 1981 survey—The sample housing units designated to be interviewed in the 1981

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 and remained in sample after the 1981 panel reduction.
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 and remained in sample after the 1981 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1981 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 and remained in sample after the 1981 panel reduction. (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 and remained in sample after the 1981 panel reduction. (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 1981 Coverage Improvement Programs. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the original 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 five SMSA's were 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 10 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 for the 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.

1977-1981 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 from building permits issued prior to January 1970, but completed after April 1, 1970.
2. Mobile homes 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. Houses that have been moved onto their present site since the 1970 census.
6. Mobile homes placed outside parks since the 1970 census or vacant at the time of the 1970 census.

Each of the 15 SMSA's was previously interviewed in 1974 (1975 for Madison, Wis.) and 1977. The Coverage Improvement Program was conducted as part of the 1977 AHS with the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., SMSA's receiving some updating and refining as part of the 1981 AHS. The following discussion applies to both the prior year 1977 and the 1981 coverage improvement procedures. For the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and Spokane, Wash., 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 1981.

Coverage improvement for deficiency 1—A sample of new construction units, whose permits were issued before January 1970, but completed after April 1970, was selected for each of the 15 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. 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 (regular AHS units), while sample units selected from three-or-more-unit structures were sampled at one-half the rate of regular AHS units.

The first stage 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 Fort Worth, Tex., SMSA, an additional sample of 1968 permits for three-or-more-unit structures was included in the second-stage sampling. For the third stage, structures of size three or more units were divided into clusters of an expected size of two units and a sample of clusters was selected. This procedure was employed in the following SMSA's: Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Orlando, Fla.; Spokane, Wash.; and Wichita, Kans.

In the remaining 9 SMSA's, 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, this procedure was also used in parts of the Anaheim-Santa Ana-Garden Grove, Calif., SMSA to supplement the sample described above. These procedures added an estimated 702 new construction 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 placed in a park that was 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 units represented the same number of units that the regular AHS sample units represented. This procedure added an estimated 1,934 units to the coverage of the housing inventory of this SMSA.

Coverage improvement for deficiencies 3-6—The remaining missed 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 units converted to residential use.
3. Mobile homes 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 a usual residence elsewhere.
4. Houses that had been moved onto their present site since the 1970 census.

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample 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 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 second procedure was designed to represent missed units from structures represented in the AHS. These missed units were:

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential 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 structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply both procedures during the 1977 Coverage Improvement Program to all SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in the Anaheim-Santa Ana-Garden Grove, Calif.; and Spokane, Wash., SMSA's. These SMSA's did not receive application of the second procedure. The Dallas, Tex.; Minneapolis-St. Paul, Minn.; and Pittsburgh, Pa., SMSA's did not receive application of either procedure during the 1977 Coverage Improvement Program. The Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Minneapolis-St. Paul, Minn.; Pittsburgh, Pa.; and the Spokane, Wash., SMSA's had the first procedure completed as part of the 1981 Coverage Improvement Program. The remaining 10 1981 SMSA's received the full application of both procedures as part of the 1977 Coverage Improvement Program.

The first procedure added an estimated 2,504 housing units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 152 housing units.

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 1981 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 1981 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1977 (i.e., 1977-1981 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1981 housing inventory—The AHS estimates of characteristics of the 1981 housing inventory were produced using a one-stage ratio estimation procedure for the Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; and Madison, Wis., SMSA's, a two-stage ratio estimation procedure for the Detroit, Mich.; Orlando, Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, and a three-stage ratio estimation procedure for the remaining seven SMSA's.

Prior to the implementation of the ratio estimation procedures, 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 nonpermit 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 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.

The next ratio estimation procedure was applied in the Boston, Mass.; Detroit, Mich.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's. This procedure was utilized 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 ratio estimation factor was then applied to the existing weight for all sample housing units classified within the corresponding ratio estimation cell.

The next ratio estimation procedure was applied in 10 SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1981 housing inventory in each sector (central city and balance) for each SMSA to an independent estimate of total housing units for the corresponding sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1981 housing unit inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

The independent estimates of total housing units that were used as the numerator of this ratio are described below. The denominator of this ratio was obtained from the weighted estimate of the AHS-SMSA sample housing units using the existing weight.

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for each of the 15 1981 SMSA's. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1981 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units after the permit new construction ratio estimation (i.e., the existing sample estimate) were compared to the corresponding independent estimates for the central city, balance, and total SMSA for each of the 15 SMSA's and the estimate which showed the most likely level of net growth since the 1980 census in both the central city and balance as well as the total SMSA were used in this ratio estimation. As a result of this analysis, these independent estimates were used in the Newark, N.J.; Pittsburgh, Pa.; Spokane, Wash.; Tacoma, Wash.; and Washington, D.C.-Md.-Va., SMSA's.

The remaining five SMSA's used a combination of these independent estimates and sample estimates. For the Boston, Mass.; and Minneapolis-St. Paul, Minn., SMSA's, the sample estimate was used for the total SMSA. For the sectors, the sample estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the independent estimates.

For the Orlando Fla.; Phoenix, Ariz.; and Wichita, Kans., SMSA's, the independent estimate of units was used for the total SMSA. For the sectors, the independent estimate of the total SMSA was proportioned between the central city and balance according to the central city/balance distribution of total housing units given by the sample estimates. Due to the central city

boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimate did not reflect the 1970 central city/balance SMSA definition. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definition were used.

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 in these 10 SMSA's, and the resulting product was used as the final weight for tabulation purposes. For the other five SMSA's, the sample estimates were used as the final weight for tabulation purposes. These SMSA's included Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Detroit, Mich.; Fort Worth, Tex.; and Madison, Wis., SMSA's.

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

1977-1981 lost housing units—The 1977-1981 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1977 housing inventory for 14 of the SMSA's and the corresponding two-stage ratio estimation procedure used for the Pittsburgh, Pa., SMSA as was described in the 1977 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas*. Since the 1977-1981 lost housing units existed, by definition, in the 1977 housing inventory, there was a 1977 housing inventory weight associated with each 1977-1981 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1977-1981 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 two-stage ratio estimation process for the Pittsburgh, Pa., SMSA and a one-stage ratio estimation process for the remaining SMSA's. 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; 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 1981 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, and the results 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) 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, the 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 presented.

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 1981 survey, the number of missed housing units may be considerably less for 1981.

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 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 1981 housing inventory as well as estimates of characteristics of the 1977-1981 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 8,159 for the total SMSA, 4,724 for the central city of the SMSA, and 6,562 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 1981 housing inventory as well as estimated percentages of the 1977-1981 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 1981 there were 103,200 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 1,260. 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
100,000	1,270
103,200	x
112,800	1,210

The entry of "x" is determined as follows by vertically interpolating between 1,270 and 1,210.

$$103{,}200 - 100{,}000 = 3{,}200$$
$$112{,}800 - 100{,}000 = 12{,}800$$

$$1{,}270 + \frac{3{,}200}{12{,}800}(1{,}210 - 1{,}270) = 1{,}260$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 101,940 to 104,460 housing units. Therefore, a conclusion that the average estimate of 1981 owner-occupied housing units (derived from all possible samples) 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 101,180 to 105,220 housing units with 90 percent confidence; and that the average estimate lies within the interval from 100,680 to 105,720 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 103,200 owner-occupied housing units, 32,000, or 31.0 percent, had two bedrooms. Interpolation using table II of this appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 31.0 percent is approximately 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	31.0	50
100,000	0.9	a	1.0
103,200		p	
150,000	0.7	b	0.8

1. The entry for cell "a" is determined by horizontal interpolation between 0.9 and 1.0.

$$31.0 - 25.0 = 6.0$$
$$50.0 - 25.0 = 25.0$$

$$0.9 + \frac{6.0}{25.0}(1.0 - 0.9) = 0.9$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.7 and 0.8.

$$31.0 - 25.0 = 6.0$$
$$50.0 - 25.0 = 25.0$$

$$0.7 + \frac{6.0}{25.0}(0.8 - 0.7) = 0.7$$

3. The entry for "p" is then determined by vertical interpolation between 0.9 and 0.7..

$$103{,}200 - 100{,}000 = 3{,}200$$
$$150{,}000 - 100{,}000 = 50{,}000$$

$$0.9 + \frac{3{,}200}{50{,}000}(0.7 - 0.9) = 0.9$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 30.1 to 31.9 percent; the 90-percent confidence interval is from 29.6 to 32.4 percent; and the 95-percent confidence interval is from 29.2 to 32.8 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; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1977 and 1981 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1977 and 1981 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1981 there were 51,600 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 19,600. Table I shows the standard error of 32,000 is approximately 1,000, and the standard error of 51,600 is approximately 1,200. Therefore, the standard error of the estimated difference of 19,600 is about 1,560.

$$1{,}560 = \sqrt{(1{,}000)^2 + (1{,}200)^2}$$

Consequently, the 68-percent confidence interval for the 19,600 difference is from 18,040 to 21,160 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 17,100 to 22,100 housing units, and the 95-percent confidence interval is from 16,480 to 22,720 housing units. Thus, we can conclude with 95 percent confidence that the number of 1981 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied 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 to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.5. The base of the distribution from which this median was determined is 103,200 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 103,200 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 15,100 housing units or 14.6 percent fall below this interval, and 37,600 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:

$$1.5 + (2.5-1.5)\ \frac{(48.0-14.6)}{36.4} = 2.4$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.0 percent derived in step 2. About 52,700 housing units or 51.1 percent fall below this interval, and 19,200 housing units or 18.6 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(52.0-51.1)}{18.6} = 2.5$$

Thus, the 95-percent confidence interval ranges from 2.4 to 2.5 persons. Although it appears that this confidence interval has the same estimate as the upper limit it actually is a reflection of the rounding error associated with this median (see the paragraph on rounding errors in the nonsampling errors section of this appendix).

TABLE I. **Standard Errors for Estimated Number of Housing Units in the 1981 Housing Inventory and for Estimated Number of 1977-1981 Lost Units for the Wichita, Kans., 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	40	40	40	25,000	920	880	710
100	60	60	60	50,000	1,190	1,050	380
200	90	90	90	54,200	1,210	1,060	—
500	140	140	140	75,000	1,290	1,000	—
700	170	170	160	100,000	1,270	670	—
1,000	200	200	190	112,800	1,210	—	—
2,500	310	310	300	150,000	780	—	—
5,000	440	440	410	167,000	—	—	—
10,000	610	600	560				

[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 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 where the estimate for the total SMSA was derived from an independent estimate.

TABLE II. **Standard Errors for Estimated Percentages of Housing Units in the 1981 Housing Inventory and for Estimated Percentages of 1977-1981 Lost Housing Units for the Wichita, Kans., 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	28.7	28.7	28.7	28.7	28.7	31.7	10,000	0.4	0.6	1.4	1.9	2.7	3.2
200	16.7	16.7	16.7	16.7	19.4	22.4	25,000	0.2	0.4	0.9	1.2	1.7	2.0
500	7.4	7.4	7.4	8.5	12.3	14.2	50,000	0.08	0.3	0.6	0.9	1.2	1.4
700	5.4	5.4	5.4	7.2	10.4	12.0	75,000	0.05	0.2	0.5	0.7	1.0	1.2
1,000	3.9	3.9	4.4	6.0	8.7	10.0	100,000	0.04	0.2	0.4	0.6	0.9	1.0
2,500	1.6	1.6	2.8	3.8	5.5	6.3	150,000	0.03	0.2	0.4	0.5	0.7	0.8
5,000	0.8	0.9	2.0	2.7	3.9	4.5	200,000	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 the central city and 1.1 for the balance (not in central city).