1977 Metropolitan Appendix B: Source and Accuracy

Table of Contents

Albany-Schenectady-Troy 3

Anaheim-Santa Ana-Garden Grove 14

Boston 25

Dallas 37

Detroit 48

Ft. Worth 60

Los Angeles-Long Beach 71

Madison 82

Memphis 93

Minneapolis-St Paul 104

Newark 115

Orlando 126

Phoenix 137

Pittsburgh 148

Saginaw 159

Salt Lake City 170

Spokane 181

Tacoma 192

Washington 203

Wichita 215

Albany-Schenectady-Troy 1977

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1977 survey App-42
Selection of the 1974 sample App-42
1977 sample reduction App-43
1974-1977 additions to the housing inventory App-43
Coverage improvement sample selection App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1970 Census of Population and Housing App-45
ESTIMATION App-45
1977 housing inventory App-45
1974-1977 lost units App-46
1974 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES App-46
Nonsampling errors App-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-50
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-51

SAMPLE DESIGN

Annual Housing Survey.—The estimates for each of the 20 SMSA's are based on data collected from the 1977 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 20 SMSA's, the data were collected for the 11-month period from April 1977 through February 1978 with one-eleventh of the sample units visited each month.

Each group of SMSA's is interviewed for the AHS on a rotating basis, and this group (A-2 SMSA's) was the first to be revisited. Nineteen of the 20 SMSA's were interviewed for the first time in 1974, while the Madison, Wis., SMSA was interviewed for the first time in 1975.

In each group of SMSA's, the largest SMSA from each of the 4 geographic regions is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All remaining SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the A-2 group (1977-1978) are: Boston, Mass.; Detroit, Mich.; Los Angeles-Long Beach, Calif.; and Washington, D.C.-Md.-Va.

The remaining SMSA's in the A-2 group are: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.

In this SMSA, 4,680 housing units were eligible for interview. Of these sample units, 218 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 units eligible for interview, 502 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 1977 survey.—The sample housing units designated for interview in the 1977 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 1974 survey which were not part of the 1977 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 1974 survey which were not part of the 1977 reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1977 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from the list of new construction building permits issued since the 1974 survey. (This sample represented the housing units built in permit-issuing areas, since the 1974 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1974 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1974 survey.)
5. All sample housing units that were selected as part of the 1977 Coverage Improvement Program. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1974 sample.—The sample for the SMSA's which are 100-percent permit-issuing was selected from two sample frames—units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which are

not 100-percent permit-issuing included a sample selected from a third frame—those units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). The following SMSA's are 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Los Angeles-Long Beach, Calif.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 14 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 four largest SMSA's, the overall sampling rate differed by central city and 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 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 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 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-$5,999. . .		
$6,000-$9,999. . .		
$10,000-$14,999 .		
$15,000 and over .		

Thus, for this SMSA, the occupied housing unit records from this 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 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 are 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 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 units in existence at the time of interview in these selected segments are in sample. Thus, units enumerated in the 1970 census as well as units built since the 1970 census are included.

1977 sample reduction.—In previous years, AHS sample units have been divided among 12 panels which were interviewed each month. Due to budget limitations, it became necessary to drop one of these panels from the 1977 survey. The designated sample size for 1977 was thus reduced by one-twelfth.

1974-1977 additions to the housing inventory.—In the permit-issuing universe, a sample of new construction building permits, issued since the 1974 survey, was

selected to represent housing units built in permit-issuing areas since the 1974 survey. Sampling procedures were identical to those used in selecting the 1970-1974 new construction sample. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1974, to identify any units missed in the 1974 survey or any units added since the 1974 survey.

Coverage improvement sample selection.—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 census.
3. Units missed in the 1970 census.
4. 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.

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 A-2 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 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: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Orlando, Fla.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; and Wichita, Kans.

In the remaining 10 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 and the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 913 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, for each 1977-1978 SMSA, 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 159 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 no 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 units in multi-unit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multi-unit structures

selected, all units were listed and matched to the 1970 census. Any missed units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply these two procedures in all 20 A-2 SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in three SMSA's. In these SMSA's, a string of 10 structures (instead of 8) that were eligible to be selected for the AHS, were listed. The full application of this technique was employed in 12 SMSA's, with 5 SMSA's not receiving any application of the first procedure. The second procedure was instituted in 13 SMSA's. The following table shows which of these two techniques were applied to the A-2 SMSA's.

The first procedure added an estimated 318 units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 1,365 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 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 1977 housing inventory) and estimates pertaining to characteristics of units removed from the housing inventory since 1974 (i.e., 1974-1977 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1977 housing inventory.—The AHS estimates of characteristics of the 1977 housing inventory employed a 1-stage ratio estimation procedure in 19 SMSA's, with a second-stage ratio estimation procedure being employed in 1 SMSA. Prior to the implementation of the ratio estimation procedure, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of 1 or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample 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:

SMSA	Coverage improvement technique employed			
	First		Second	Neither
	Partial	Full		
Albany-Schenectady-Troy, N.Y	X		X	
Anaheim-Santa Ana-Garden Grove, Calif	X			
Boston, Mass		X	X	
Dallas, Tex				X
Detroit, Mich		X	X	
Fort Worth, Tex		X	X	
Los Angeles-Long Beach, Calif				X
Madison, Wis		X	X	
Memphis, Tenn.-Ark		X	X	
Minneapolis-St. Paul, Minn				X
Newark, N.J		X	X	
Orlando, Fla		X	X	
Phoenix, Ariz		X	X	
Pittsburgh, Pa				X
Saginaw, Mich		X	X	
Salt Lake City, Utah				X
Spokane, Wash	X			
Tacoma, Wash		X	X	
Washington, D.C.-Md.-Va		X	X	
Wichita, Kans		X	X	

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of 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 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 unit within the corresponding ratio estimation category.

The effect of this ratio estimation procedure was to somewhat reduce the variance due to the variation in sample size of strata used in the sample selection of the permit-issuing universe, thereby reducing the sampling error below what would have been obtained by simply weighting the results of the sample by the inverse of the probability of selection. Ordinarily, this would have been controlled by sampling within the strata during the sample selection process. However, prior to the AHS sample selection within each SMSA, units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. Thus, some variation in sample size was introduced during the AHS sample selection process.

For the Pittsburgh, Pa., SMSA, a second-stage ratio estimation procedure was also employed. This procedure involved the ratio estimation of the AHS weighted sample estimate of the October 1977 housing inventory to an independent estimate of the SMSA's October 1977 housing inventory. This ratio estimate factor equaled the following:

$$\frac{\text{SMSA independent estimate of October 1977 housing inventory}}{\text{AHS sample estimate of the housing inventory}}$$

The numerator of the ratio was derived by using the 1970 census estimate of total housing units for the SMSA in conjunction with the estimate of the change in the housing inventory since the census. The estimate of change was based on administrative records from utility companies. The denominator of that ratio was obtained from the weighted estimate of AHS sample units, using the existing weight after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for each sample unit and the resulting product was used as the final weight for tabulation purposes.

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

The second-stage ratio estimation procedure was not employed in all A-2 SMSA's, since it was felt that coverage improvement procedures adequately corrected known deficiencies in the sample where full coverage improvement procedures were employed. The independent estimates were only employed in SMSA's where full coverage improvements were not used and the estimated relative bias of the independent estimate was thought to be low enough (i.e., 2 percent or less over a 10-year period) to preclude the possibility of doing more harm than good to the survey results. A measure of the relative bias of the independent estimate was available based on a comparison between the 1970 census estimate of the April 1970 housing inventory and the independent estimate of the April 1970 housing inventory which was projected over a 10-year period from the 1960 census using the same procedure that produced the October 1977 independent estimates.

1974-1977 lost units.—The 1974-1977 lost unit estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1974 housing inventory, as was described in the 1974 Current Housing Report, Series H-170, **Housing Characteristics for Selected Metropolitan Areas.** Since the 1974-1977 lost units existed, by definition, in the 1974 housing inventory, there was a 1974 housing inventory weight associated with each 1974-1977 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1974-1977 lost units.

1974 estimation procedure.—This report presents data on the housing characteristics of the 1974 housing inventory from the 1974 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1974.

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 to provide correct information on the part of respondents, mistakes in recording or coding the data, and other errors of collection, response, processing, coverage, and estimation for missing data. Nonsampling errors are not unique to sample surveys since they can, and do, occur in complete censuses as well.

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing, the 1977 AHS-SMSA sample, and the 1974 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 included space errors, definitional errors, and occupancy errors. The "content" errors measured the accuracy of the data collected for surveyed 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.—For the 1977 AHS-SMSA sample, a reinterview program was conducted to obtain a measurement of some of the components of the nonsampling error associated with the AHS estimates. A sample of households was revisited and answers to some of the questions on the AHS questionnaire were obtained again. The original interview and the reinterview were assumed to be two independent readings and, thus, were the basis for the measurement of the accuracy of the AHS data collected from interviewed households.

As part of the reinterview, a check was made at each of these households to determine if the following was done during the original interview:

1. The correct unit was visited.
2. The correct number of housing units were interviewed at that address.
3. The correct information on "Year Built" was obtained.
4. The correct information on "Tenure" was obtained.
5. The correct information on "Household Composition" was obtained.
6. The correct information on "Type of Housing Unit" was obtained.
7. The correct information on "Occupancy Status" was obtained.

This check was made for interviewer evaluation and control. That is, tolerance limits were derived to determine which interviewers passed or failed this reinterview check. The results of this study are available and are presented in the following Census Bureau memorandum, "Response Error in the Annual Housing Survey as Measured by Reinterviews—Year IV, 1977-1978 SMSA Sample."

For the A-2 reinterview program, measures of response inconsistency were taken on two subsamples—reconciled and unreconciled. After the question is answered in the reconciled subsample, the interviewer presented the original response and then asked the respondent to decide upon the best answer. Past experience has shown that the level of inconsistency is lower when the interviewer has the response available, and thus, it is believed that there is a downward bias in these estimates. In the unreconciled sample, the interviewer did not have access to previous responses. For the reconciled items, three-fourths showed low levels of response inconsistency while the remaining one-fourth were in the moderate range. For the unreconciled items, approximately one-half showed low levels of response disagreement, with the remaining one-half of the categories in the moderate and high inconsistency range. Although attitudinal items were excluded from the A-2 reinterview program, other reinterview programs have shown that, in general, attitudinal items such as items 102-104 (see facsimile of the 1977 AHS questionnaire, page App-34) have moderate to high levels of inconsistency. Moderate levels of inconsistency indicate that there is some problem with inconsistent reporting, and high levels indicate that improvements are needed in the method used to collect data or that the category concepts themselves are ambiguous.

A second measure of response error, the net difference rate, which is a measure of net error (or bias) in a particular category, was also provided by this reinterview program. None of the categories tested showed any significant evidence of bias.

The results of this study were based on sample data, so there is a sampling error associated with these estimates of nonsampling error. Therefore, the possibility of such errors should be taken into account when considering the results of this study.

A similar study was conducted for the 1974 AHS-SMSA sample. The results of which are presented in the following Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1974."

Coverage errors.—With respect to errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units, since, due to the relatively short time span involved, it is possible that construction of these units was not completed at the time the survey was conducted. In which case, they would not have been eligible for interview. In addi-

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

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit. Also, additional coverage errors may exist in SMSA's where the complete Coverage Improvement Program was not applied (see the table on page App-45 of this appendix).

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all units located inside these ED's would be represented in the sample. However, it has been estimated that the 1974 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. Since these ED's were recanvassed for the 1977 survey, the number of missed housing units may be considerably less for 1977.

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 schedules, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by the standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

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

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

The figures presented in the following tables 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 1974 housing inventory can be found in the AHS Series H-170 reports for 1974.

Table I presents the standard errors applicable to estimates of characteristics of the 1977 housing inventory as well as estimates of characteristics of the

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1977 Housing Inventory and for Estimated Number of 1974-1977 Lost Units for the Albany-Schenectady-Troy, N.Y., SMSA, for the Central City of the SMSA, and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]	
	SMSA and not in central city	Central city
0	60	60
100	80	80
200	110	110
500	180	170
700	210	210
1,000	250	250
2,500	400	390
5,000	560	540
10,000	790	760
25,000	1,240	1,130
50,000	1,730	1,430
75,000	2,080	1,510
100,000	2,360	1,420
150,000	2,790	—
200,000	3,100	—
250,000	3,320	—
300,000	3,480	—

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2 for the total SMSA, 1.1 for the central city, and 1.2 for the balance (not in central city) estimates.

1974-1977 lost units. Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 11,820 for the total SMSA, 4,190 for the central city of the SMSA, and 10,950 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 presents the standard errors of estimated percentages of the 1977 housing inventory as well as estimated percentages of the 1974-1977 lost units. 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 1977 there were 167,100 owner-occupied housing units in this SMSA. Interpolation in table I of this appendix shows that the standard error of an estimate of this size is approximately 2,900. 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	2,790
167,100	x
200,000	3,100

The entry for x is determined as follows by vertically interpolating between 2,790 and 3,100:

167,100−150,000 = 17,100
200,000−150,000 = 50,000

$$2{,}790+\frac{17{,}100}{50{,}000}(3{,}100-2{,}790)$$
$$=2{,}900$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 164,200 to 170,000 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1977 owner-occupied housing units lies within a range computed in this way would be correct for roughly 68 percent

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 Lost Units for the Albany-Schenectady-Troy, N.Y., SMSA, for the Central City of the SMSA, and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	38.8	38.8	38.8	38.8	38.8	39.8
200	24.1	24.1	24.1	24.1	24.4	28.2
500	11.3	11.3	11.3	11.3	15.4	17.8
700	8.3	8.3	8.3	9.0	13.0	15.0
1,000	6.0	6.0	6.0	7.6	10.9	12.6
2,500	2.5	2.5	3.5	4.8	6.9	8.0
5,000	1.3	1.3	2.5	3.4	4.9	5.6
10,000	0.6	0.8	1.7	2.4	3.4	4.0
25,000	0.3	0.5	1.1	1.5	2.2	2.5
50,000	0.13	0.4	0.8	1.1	1.5	1.8
75,000	0.08	0.3	0.6	0.9	1.3	1.5
100,000	0.06	0.3	0.5	0.8	1.1	1.3
150,000	0.04	0.2	0.4	0.6	0.9	1.0
200,000	0.03	0.2	0.4	0.5	0.8	0.9
250,000	0.03	0.2	0.3	0.5	0.7	0.8
300,000	0.02	0.14	0.3	0.4	0.6	0.7

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2 for the total SMSA, 1.1 for the central city, and 1.2 for the balance (not in central city) estimates.

of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 162,460 to 171,740 housing units with 90 percent confidence; and that the average estimate lies within the interval from 161,300 to 172,900 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 167,100 owner-occupied housing units, 34,700, or 20.8 percent, had two bedrooms. Interpolation in table II of the appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 20.8 percent is approximately 0.8 percentage points. The following interpolation procedure was used:

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

Base of percentage	Estimated percentage		
	10 or 90	20.8	25.0
150,000. . . .	0.6	a	0.9
167,100. . . .		p	
200,000. . . .	0.5	b	0.8

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

$$20.8-10.0 = 10.8$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{10.8}{15.0}(0.9-0.6) = 0.82$$

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

$$20.8-10.0 = 10.8$$
$$25.0-10.0 = 15.0$$

$$0.5 + \frac{10.8}{15.0}(0.8-0.5) = 0.72$$

3. The entry for "p" was then determined by vertical interpolation between 0.82 and 0.72.

$$167{,}100-150{,}000 = 17{,}100$$
$$200{,}000-150{,}000 = 50{,}000$$

$$0.82 + \frac{17{,}100}{50{,}000}(0.72-0.82) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 20.0 to 21.6 percent; the 90-percent confidence interval is from 19.5 to 22.1 percent; and the 95-percent confidence interval is from 19.2 to 22.4 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. However, 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 1974 and 1977 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1974 and 1977 characteristics.

Illustration of the computation of the standard error of a difference.—Table A-1 of part A of this report shows that in 1977 there were 83,300 owner-occupied units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied units with two bedrooms and owner-occupied units with three bedrooms is 48,600. Table I shows the standard error of 83,300 is approximately 2,170 and the standard error of 34,700 is approximately 1,430. Therefore the standard error of the estimated difference of 48,600 is about:

$$2{,}600 = \sqrt{(2{,}170)^2 + (1{,}430)^2}$$

Consequently, the 68-percent confidence interval for the 48,600 difference is from 46,000 to 51,200 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 44,440 to 52,760 housing units, and the 95-percent confidence interval is from 43,400 to 53,800. Thus, we can conclude with 95 percent confidence that the number of 1977 owner-occupied 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, read off the confidence interval corresponding to the two points established in step 2.

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 167,100 housing units.

1. Interpolation in table II shows that the standard error of 50 percent on a base of 167,100 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 table A-1 of part A, it can be seen by cumulating the frequencies for the first two categories that 71,800 owner-occupied housing units, or 43.0 percent, had one or two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) and that an additional 32,200 owner-occupied housing units, or 19.3 percent, had three persons (i.e., 2.5 to 3.5 persons).

By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{48.0-43.0}{19.3}\right) = 2.8$$

Similarly, the upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{52.0-43.0}{19.3}\right) = 3.0$$

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

Anaheim-Santa Ana-Garden Grove 1977

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1977 survey App-42
Selection of the 1974 sample App-42
1977 sample reduction App-43
1974-1977 additions to the housing inventory App-43
Coverage improvement sample selection App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1970 Census of Population and Housing App-45
ESTIMATION App-45
1977 housing inventory App-45
1974-1977 lost units App-46
1974 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES App-46
Nonsampling errors App-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-50
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-51

SAMPLE DESIGN

Annual Housing Survey.—The estimates for each of the 20 SMSA's are based on data collected from the 1977 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 20 SMSA's, the data were collected for the 11-month period from April 1977 through February 1978 with one-eleventh of the sample units visited each month.

Each group of SMSA's is interviewed for the AHS on a rotating basis, and this group (A-2 SMSA's) was the first to be revisited. Nineteen of the 20 SMSA's were interviewed for the first time in 1974, while the Madison, Wis., SMSA was interviewed for the first time in 1975.

In each group of SMSA's, the largest SMSA from each of the 4 geographic regions is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All remaining SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the A-2 group (1977-1978) are: Boston, Mass.; Detroit, Mich.; Los Angeles-Long Beach, Calif.; and Washington, D.C.-Md.-Va.

The remaining SMSA's in the A-2 group are: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.

In this SMSA, 5,250 housing units were eligible for interview. Of these sample units, 236 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 units eligible for interview, 281 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 1977 survey.—The sample housing units designated for interview in the 1977 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 1974 survey which were not part of the 1977 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 1974 survey which were not part of the 1977 reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1977 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from the list of new construction building permits issued since the 1974 survey. (This sample represented the housing units built in permit-issuing areas since the 1974 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1974 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1974 survey.)
5. All sample housing units that were selected as part of the 1977 Coverage Improvement Program. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1974 sample.—The sample for the SMSA's which are 100-percent permit-issuing was selected from two sample frames—units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which are not 100-percent permit-issuing in-

cluded a sample selected from a third frame—those units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). The following SMSA's are 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Los Angeles-Long Beach, Calif.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 14 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 four largest SMSA's, the overall sampling rate differed by central city and 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 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 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 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-$5,999. . .		
$6,000-$9,999. . .		
$10,000-$14,999 .		
$15,000 and over .		

Thus, for this SMSA, the occupied housing unit records from this 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 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 are 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 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 units in existence at the time of interview in these selected segments are in sample. Thus, units enumerated in the 1970 census as well as units built since the 1970 census are included.

1977 sample reduction.—In previous years, AHS sample units have been divided among 12 panels which were interviewed each month. Due to budget limitations, it became necessary to drop one of these panels from the 1977 survey. The designated sample size for 1977 was thus reduced by one-twelfth.

1974-1977 additions to the housing inventory.—In the permit-issuing universe, a sample of new construction building permits, issued since the 1974 survey, was

selected to represent housing units built in permit-issuing areas since the 1974 survey. Sampling procedures were identical to those used in selecting the 1970-1974 new construction sample. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1974, to identify any units missed in the 1974 survey or any units added since the 1974 survey.

Coverage improvement sample selection.—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 census.
3. Units missed in the 1970 census.
4. 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.

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 A-2 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 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: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Orlando, Fla.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; and Wichita, Kans.

In the remaining 10 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 and the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 18,103 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, for each 1977-1978 SMSA, 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 14,458 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 no 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 units in multi-unit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multi-unit structures

selected, all units were listed and matched to the 1970 census. Any missed units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply these two procedures in all 20 A-2 SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in three SMSA's. In these SMSA's, a string of 10 structures (instead of 8) that were eligible to be selected for the AHS, were listed. The full application of this technique was employed in 12 SMSA's, with 5 SMSA's not receiving any application of the first procedure. The second procedure was instituted in 13 SMSA's. The following table shows which of these two techniques were applied to the A-2 SMSA's.

The first procedure added an estimated 1,061 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 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 1977 housing inventory) and estimates pertaining to characteristics of units removed from the housing inventory since 1974 (i.e., 1974-1977 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1977 housing inventory.—The AHS estimates of characteristics of the 1977 housing inventory employed a 1-stage ratio estimation procedure in 19 SMSA's, with a second-stage ratio estimation procedure being employed in 1 SMSA. Prior to the implementation of the ratio estimation procedure, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of 1 or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample 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 esti-

SMSA	Coverage improvement technique employed			
	First		Second	Neither
	Partial	Full		
Albany-Schenectady-Troy, N.Y	X		X	
Anaheim-Santa Ana-Garden Grove, Calif	X			
Boston, Mass		X	X	
Dallas, Tex				X
Detroit, Mich		X	X	
Fort Worth, Tex		X	X	
Los Angeles-Long Beach, Calif				X
Madison, Wis		X	X	
Memphis, Tenn.-Ark		X	X	
Minneapolis-St. Paul, Minn				X
Newark, N.J		X	X	
Orlando, Fla		X	X	
Phoenix, Ariz		X	X	
Pittsburgh, Pa				X
Saginaw, Mich		X	X	
Salt Lake City, Utah				X
Spokane, Wash	X			
Tacoma, Wash		X	X	
Washington, D.C.-Md.-Va		X	X	
Wichita, Kans		X	X	

mation factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of 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 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 unit within the corresponding ratio estimation category.

The effect of this ratio estimation procedure was to somewhat reduce the variance due to the variation in sample size of strata used in the sample selection of the permit-issuing universe, thereby reducing the sampling error below what would have been obtained by simply weighting the results of the sample by the inverse of the probability of selection. Ordinarily, this would have been controlled by sampling within the strata during the sample selection process. However, prior to the AHS sample selection within each SMSA, units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. Thus, some variation in sample size was introduced during the AHS sample selection process.

For the Pittsburgh, Pa., SMSA, a second-stage ratio estimation procedure was also employed. This procedure involved the ratio estimation of the AHS weighted sample estimate of the October 1977 housing inventory to an independent estimate of the SMSA's October 1977 housing inventory. This ratio estimate factor equaled the following:

$$\frac{\text{SMSA independent estimate of October 1977 housing inventory}}{\text{AHS sample estimate of the housing inventory}}$$

The numerator of the ratio was derived by using the 1970 census estimate of total housing units for the SMSA in conjunction with the estimate of the change in the housing inventory since the census. The estimate of change was based on administrative records from utility companies. The denominator of that ratio was obtained from the weighted estimate of AHS sample units, using the existing weight after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for each sample unit and the resulting product was used as the final weight for tabulation purposes.

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

The second-stage ratio estimation procedure was not employed in all A-2 SMSA's, since it was felt that coverage improvement procedures adequately corrected known deficiencies in the sample where full coverage improvement procedures were employed. The independent estimates were only employed in SMSA's where full coverage improvements were not used and the estimated relative bias of the independent estimate was thought to be low enough (i.e., 2 percent or less over a 10-year period) to preclude the possibility of doing more harm than good to the survey results. A measure of the relative bias of the independent estimate was available based on a comparison between the 1970 census estimate of the April 1970 housing inventory and the independent estimate of the April 1970 housing inventory which was projected over a 10-year period from the 1960 census using the same procedure that produced the October 1977 independent estimates.

1974-1977 lost units.—The 1974-1977 lost unit estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1974 housing inventory, as was described in the 1974 Current Housing Report, Series H-170, **Housing Characteristics for Selected Metropolitan Areas.** Since the 1974-1977 lost units existed, by definition, in the 1974 housing inventory, there was a 1974 housing inventory weight associated with each 1974-1977 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1974-1977 lost units.

1974 estimation procedure.—This report presents data on the housing characteristics of the 1974 housing inventory from the 1974 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1974.

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 to provide correct information on the part of respondents, mistakes in recording or coding the data, and other errors of collection, response, processing, coverage, and estimation for missing data. Nonsampling errors are not unique to sample surveys since they can, and do, occur in complete censuses as well.

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing, the 1977 AHS-SMSA sample, and the 1974 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 included space errors, definitional errors, and occupancy errors. The "content" errors measured the accuracy of the data collected for surveyed 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.—For the 1977 AHS-SMSA sample, a reinterview program was conducted to obtain a measurement of some of the components of the nonsampling error associated with the AHS estimates. A sample of households was revisited and answers to some of the questions on the AHS questionnaire were obtained again. The original interview and the reinterview were assumed to be two independent readings and, thus, were the basis for the measurement of the accuracy of the AHS data collected from interviewed households.

As part of the reinterview, a check was made at each of these households to determine if the following was done during the original interview:

1. The correct unit was visited.
2. The correct number of housing units were interviewed at that address.
3. The correct information on "Year Built" was obtained.
4. The correct information on "Tenure" was obtained.
5. The correct information on "Household Composition" was obtained.
6. The correct information on "Type of Housing Unit" was obtained.
7. The correct information on "Occupancy Status" was obtained.

This check was made for interviewer evaluation and control. That is, tolerance limits were derived to determine which interviewers passed or failed this reinterview check. The results of this study are available and are presented in the following Census Bureau memorandum, "Response Error in the Annual Housing Survey as Measured by Reinterviews—Year IV, 1977-1978 SMSA Sample."

For the A-2 reinterview program, measures of response inconsistency were taken on two subsamples—reconciled and unreconciled. After the question is answered in the reconciled subsample, the interviewer presented the original response and then asked the respondent to decide upon the best answer. Past experience has shown that the level of inconsistency is lower when the interviewer has the response available, and thus, it is believed that there is a downward bias in these estimates. In the unreconciled sample, the interviewer did not have access to previous responses. For the reconciled items, three-fourths showed low levels of response inconsistency while the remaining one-fourth were in the moderate range. For the unreconciled items, approximately one-half showed low levels of response disagreement, with the remaining one-half of the categories in the moderate and high inconsistency range. Although attitudinal items were excluded from the A-2 reinterview program, other reinterview programs have shown that, in general, attitudinal items such as items 102-104 (see facsimile of the 1977 AHS questionnaire, page App-34) have moderate to high levels of inconsistency. Moderate levels of inconsistency indicate that there is some problem with inconsistent reporting, and high levels indicate that improvements are needed in the method used to collect data or that the category concepts themselves are ambiguous.

A second measure of response error, the net difference rate, which is a measure of net error (or bias) in a particular category, was also provided by this reinterview program. None of the categories tested showed any significant evidence of bias.

The results of this study were based on sample data, so there is a sampling error associated with these estimates of nonsampling error. Therefore, the possibility of such errors should be taken into account when considering the results of this study.

A similar study was conducted for the 1974 AHS-SMSA sample. The results of which are presented in the following Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1974."

Coverage errors.—With respect to errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units, since, due to the relatively short time span involved, it is possible that construction of these units was not com-

pleted at the time the survey was conducted. In which case, they would not have been eligible for interview. In addition to these deficiencies, new construction in special places that do not require building permits, such as military bases, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit. Also, additonal coverage errors may exist in SMSA's where the complete Coverage Improvement Program was not applied (see the table on page App-45 of this appendix).

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all units located inside these ED's would be represented in the sample. However, it has been estimated that the 1974 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. Since these ED's were recanvassed for the 1977 survey, the number of missed housing units may be considerably less for 1977.

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 schedules, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by the standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

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

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

The figures presented in the following tables 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 1974 housing inventory can be found in the AHS Series H-170 reports for 1974.

Table I presents the standard errors applicable to estimates of characteristics of the 1977 housing inventory as well as estimates of characteristics of the 1974-1977 lost units. 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

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1977 Housing Inventory and for Estimated Number of 1974-1977 Los Units for the Anaheim-Santa Ana-Garden Grove, Calif., SMSA, for the Central City of the SMSA, and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]	Size of estimate	Standard error[1]
0	140	75,000. .	3,170
100.	140	100,000.	3,650
200.	170	150,000.	4,420
500.	260	200,000.	5,050
700.	310	250,000.	5,590
1,000 . . .	370	300,000.	6,060
2,500 . . .	590	400,000.	6,840
5,000 . . .	830	500,000.	7,460
10,000 . .	1,170	600,000.	7,980
25,000 . .	1,850	700,000.	8,390
50,000 . .	2,610		

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

tables A-1, B-1, and C-1 of part A of this report are 23,580 for the total SMSA, 12,090 for the central city of the SMSA, and 20,160 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 presents the standard errors of estimated percentages of the 1977 housing inventory as well as estimated percentages of the 1974-1977 lost units. 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 1977 there were 383,800 owner-occupied housing units in this SMSA. Interpolation in table I of this appendix shows that the standard error of an estimate of this size is approximately 6,710. The following interpolation procedure was used:

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

Size of estimate	Standard error
300,000.	6,060
383,800.	x
400,000.	6,840

The entry for x is determined as follows by vertically interpolating between 6,060 and 6,840:

$$383{,}800-300{,}000 = 83{,}800$$
$$400{,}000-300{,}000 = 100{,}000$$

$$6{,}060 + \frac{83{,}800}{100{,}000}(6{,}840-6{,}060)$$
$$= 6{,}710$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 377,090 to 390,510 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1977 owner-occupied housing units

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 Lost Units for the Anaheim-Santa Ana-Garden Grove, Calif., SMSA, for the Central City of the SMSA, and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100.	58.1	58.1	58.1	58.1	58.1	58.8
200.	40.9	40.9	40.9	40.9	40.9	41.6
500.	21.7	21.7	21.7	21.7	22.8	26.3
700.	16.5	16.5	16.5	16.5	19.3	22.2
1,000	12.2	12.2	12.2	12.2	16.1	18.6
2,500	5.2	5.2	5.2	7.1	10.2	11.8
5,000	2.7	2.7	3.6	5.0	7.2	8.3
10,000	1.4	1.4	2.6	3.5	5.1	5.9
25,000	0.6	0.7	1.6	2.2	3.2	3.7
50,000	0.3	0.5	1.1	1.6	2.3	2.6
75,000	0.2	0.4	0.9	1.3	1.9	2.1
100,000. . . .	0.14	0.4	0.8	1.1	1.6	1.9
150,000. . . .	0.09	0.3	0.7	0.9	1.3	1.5
200,000. . . .	0.07	0.3	0.6	0.8	1.1	1.3
250,000. . . .	0.06	0.2	0.5	0.7	1.0	1.2
300,000. . . .	0.05	0.2	0.5	0.6	0.9	1.1
400,000. . . .	0.03	0.2	0.4	0.6	0.8	0.9
500,000. . . .	0.03	0.2	0.4	0.5	0.7	0.8
600,000. . . .	0.02	0.15	0.3	0.5	0.7	0.8
700,000. . . .	0.02	0.14	0.3	0.4	0.6	0.7

[1] Standard error 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, for the central city, and for the balance (not in central city) estimates.

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 373,064 to 394,536 housing units with 90 percent confidence; and that the average estimate lies within the interval from 370,380 to 397,220 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 383,800 owner-occupied housing units, 85,300, or 22.2 percent, had two bedrooms. Interpolation in table II of the appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 22.2 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		
	10 or 90	22.2	25
300,000. . . .	0.6	a	0.9
383,800. . . .		p	
400,000. . . .	0.6	b	0.8

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

$$22.2-10.0 = 12.2$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{12.2}{15.0}(0.9-0.6) = 0.84$$

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

$$22.2-10.0 = 12.2$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{12.0}{15.0}(0.8-0.6) = 0.76$$

3. The entry for "p" was then determined by vertical interpolation between 0.84 and 0.76.

$$383{,}300-300{,}000 = 83{,}800$$
$$400{,}000-300{,}000 = 100{,}000$$

$$0.84 + \frac{83{,}800}{100{,}000}(0.76-0.84) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 21.4 to 23.0 percent; the 90-percent confidence interval is from 20.9 to 23.5 percent; and the 95-percent confidence interval is from 20.6 to 23.8 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. However, 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 1974 and 1977 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1974 and 1977 characteristics.

Illustration of the computation of the standard error of a difference.—Table A-1 of part A of this report shows that in 1977 there were 168,900 owner-occupied units with three bedrooms in this SMSA, Thus, the apparent difference, as shown by these data, between owner-occupied units with two bedrooms and owner-occupied units with three bedrooms is 83,600. Table I shows the standard error of 168,900 is approximately 4,660 and the standard error of 85,300 is approximately 3,370. Therefore, the standard error of the estimated difference of 83,600 is about:

$$5{,}750 = \sqrt{(4{,}660)^2 + (3{,}370)^2}$$

Consequently, the 68-percent confidence interval for the 83,600 difference is from 77,850 to 89,350 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of this difference lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 74,400 to 92,800 housing units, and the 95-percent confidence interval is from 72,100 to 95,100. Thus, we can conclude with 95 percent confidence that the number of 1977 owner-occupied 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, read off the confidence interval corresponding to the two points established in step 2.

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 383,800 housing units.

1. Interpolation in table II shows that the standard error of 50 percent on a base of 383,800 is approximately 0.9 percentage points.

2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.2 and 51.8.

3. From table A-1 of part A, it can be seen by cumulating the frequencies for the first two categories that 167,100 owner-occupied housing units, or 43.5 percent, had one or two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) and that an additional 67,600 owner-occupied housing units, or 17.6 percent, had three persons (i.e., 2.5 to 3.5 persons).

By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{48.2-43.5}{17.6}\right) = 2.8$$

Similarly, the upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{51.8-43.5}{17.6}\right) = 3.0$$

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

Boston 1977

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey . . . App-42
Designation of sample housing units for the 1977 survey . . . App-42
Selection of the 1974 sample . . . App-42
1977 sample reduction . . . App-43
1974-1977 additions to the housing inventory . . . App-44
Coverage improvement sample selection . . . App-44
Coverage improvement for deficiency 1 . . . App-44
Coverage improvement for deficiency 2 . . . App-44
Coverage improvement for deficiencies 3-6 . . . App-44
1970 Census of Population and Housing . . . App-45
ESTIMATION . . . App-45
1977 housing inventory . . . App-45
1974-1977 lost units . . . App-46
1974 estimation procedure . . . App-46
Ratio estimation procedure of the 1970 Census of Population and Housing . . . App-46
RELIABILITY OF THE ESTIMATES App-47
Nonsampling errors . . . App-47
1970 census . . . App-47
AHS-SMSA . . . App-47
Coverage errors . . . App-48
Rounding errors . . . App-48
Sampling errors for the AHS-SMSA sample . . . App-48
Illustration of the use of the standard error tables . . . App-50
Differences . . . App-51
Illustration of the computation of the standard error of a difference . . . App-51
Medians . . . App-51
Illustration of the computation of the 95-percent confidence interval of a median . . . App-52

SAMPLE DESIGN

Annual Housing Survey.—The estimates for each of the 20 SMSA's are based on data collected from the 1977 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 20 SMSA's, the data were collected for the 11-month period from April 1977 through February 1978 with one-eleventh of the sample units visited each month.

Each group of SMSA's is interviewed for the AHS on a rotating basis, and this group (A-2 SMSA's) was the first to be revisited. Nineteen of the 20 SMSA's were interviewed for the first time in 1974, while the Madison, Wis., SMSA was interviewed for the first time in 1975.

In each group of SMSA's, the largest SMSA from each of the 4 geographic regions is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All remaining SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the A-2 group (1977-1978) are: Boston, Mass.; Detroit, Mich.; Los Angeles-Long Beach, Calif.; and Washington, D.C.-Md.-Va.

The remaining SMSA's in the A-2 group are: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark; Minneapolis-St. Paul, Minn.; Newark, N. J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.

In this SMSA, 13,587 housing units were eligible for interview. Of these sample units, 717 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 units eligible for interview, 1,658 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 1977 survey.—The sample housing units designated for interview in the 1977 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 1974 survey which were not part of the 1977 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 1974 survey which were not part of the 1977 reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1977 AHS questionnaire, page App-20.)

3. All sample housing units that were selected from the list of new construction building permits issued since the 1974 survey. (This sample represented the housing units built in permit-issuing areas, since the 1974 survey.)

4. All sample housing units that were added to sample segments in the nonpermit universe since the 1974 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1974 survey.)

5. All sample housing units that were selected as part of the 1977 Coverage Improvement Program. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1974 sample.—The sample for the SMSA's which are 100-percent permit-issuing was selected from two sample frames—units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and units constructed in permit-

issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which are not 100-percent permit-issuing included a sample selected from a third frame—those units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). The following SMSA's are 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Los Angeles-Long Beach, Calif.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 14 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 four largest SMSA's, the overall sampling rate differed by central city and 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 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 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 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-$5,999. . .		
$6,000-$9,999. . .		
$10,000-$14,999 .		
$15,000 and over .		

Thus, for this SMSA, the occupied housing unit records from this 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 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 are 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 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 units in existence at the time of interview in these selected segments are in sample. Thus, units enumerated in the 1970 census as well as units built since the 1970 census are included.

1977 sample reduction.—In previous years, AHS sample units have been divided among 12 panels which were interviewed each month. Due to budget limitations, it became necessary to drop one of these panels from the 1977 survey. The designated sample size for 1977 was thus reduced by one-twelfth.

1974-1977 additions to the housing inventory.—In the permit-issuing universe, a sample of new construction building permits, issued since the 1974 survey, was selected to represent housing units built in permit-issuing areas since the 1974 survey. Sampling procedures were identical to those used in selecting the 1970-1974 new construction sample. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1974, to identify any units missed in the 1974 survey or any units added since the 1974 survey.

Coverage improvement sample selection.—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 census.
3. Units missed in the 1970 census.
4. 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.

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 A-2 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 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: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Orlando, Fla.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; and Wichita, Kans.

In the remaining 10 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 and the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 11,600 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, for each 1977-1978 SMSA, 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 no 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 units in multi-unit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multi-unit structures selected, all units were listed and matched to the 1970 census. Any missed units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply these two procedures in all 20 A-2 SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in three SMSA's. In these SMSA's, a string of 10 structures (instead of 8) that were eligible to be selected for the AHS, were listed. The full application of this technique was employed in 12 SMSA's, with 5 SMSA's not receiving any application of the first procedure. The second procedure was instituted in 13 SMSA's. The following table shows which of these two techniques were applied to the A-2 SMSA's.

The first procedure added an estimated 40,189 units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 5,371 units.

SMSA	Coverage improvement technique employed			
	First		Second	Neither
	Partial	Full		
Albany-Schenectady-Troy, N.Y	X		X	
Anaheim-Santa Ana-Garden Grove, Calif	X			
Boston, Mass		X	X	
Dallas, Tex				X
Detroit, Mich		X	X	
Fort Worth, Tex		X	X	
Los Angeles-Long Beach, Calif				X
Madison, Wis		X	X	
Memphis, Tenn.-Ark		X	X	
Minneapolis-St. Paul, Minn				X
Newark, N.J		X	X	
Orlando, Fla		X	X	
Phoenix, Ariz		X	X	
Pittsburgh, Pa				X
Saginaw, Mich		X	X	
Salt Lake City, Utah				X
Spokane, Wash	X			
Tacoma, Wash		X	X	
Washington, D.C.-Md.-Va		X	X	
Wichita, Kans		X	X	

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 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 1977 housing inventory) and estimates pertaining to characteristics of units removed from the housing inventory since 1974 (i.e., 1974-1977 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1977 housing inventory.—The AHS estimates of characteristics of the 1977 housing inventory employed a 1-stage ratio estimation procedure in 19 SMSA's, with a second-stage ratio estimation procedure being employed in 1 SMSA. Prior to the implementation of the ratio estimation procedure, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of 1 or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes from both the nonpermit universe and the coverage improvement universe, and one noninter-

view cell for other sample 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:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of 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 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 unit within the corresponding ratio estimation category.

The effect of this ratio estimation procedure was to somewhat reduce the variance due to the variation in sample size of strata used in the sample selection of the permit-issuing universe, thereby reducing the sampling error below what would have been obtained by simply weighting the results of the sample by the inverse of the probability of selection. Ordinarily, this would have been controlled by sampling within the strata during the sample selection process. However, prior to the AHS sample selection within each SMSA, units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. Thus, some variation in sample size was introduced during the AHS sample selection process.

For the Pittsburgh, Pa., SMSA, a second-stage ratio estimation procedure was also employed. This procedure involved the ratio estimation of the AHS weighted sample estimate of the October 1977 housing inventory to an independent estimate of the SMSA's October 1977 housing inventory. This ratio estimate factor equaled the following:

$$\frac{\text{SMSA independent estimate of October 1977 housing inventory}}{\text{AHS sample estimate of the housing inventory}}$$

The numerator of the ratio was derived by using the 1970 census estimate of total housing units for the SMSA in conjunction with the estimate of the change in the housing inventory since the census. The estimate of change was based on administrative records from utility companies. The denominator of that ratio was obtained from the weighted estimate of AHS sample units, using the existing weight after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for each sample unit and the resulting product was used as the final weight for tabulation purposes.

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

The second-stage ratio estimation procedure was not employed in all A-2 SMSA's, since it was felt that coverage improvement procedures adequately corrected known deficiencies in the sample where full coverage improvement procedures were employed. The independent estimates were only employed in SMSA's where full coverage improvements were not used and the estimated relative bias of the independent estimate was thought to be low enough (i.e., 2 percent or less over a 10-year period) to preclude the possibility of doing more harm than good to the survey results. A measure of the relative bias of the independent estimate was available based on a comparison between the 1970 census estimate of the April 1970 housing inventory and the independent estimate of the April 1970 housing inventory which was projected over a 10-year period from the 1960 census using the same procedure that produced the October 1977 independent estimates.

1974-1977 lost units.—The 1974-1977 lost unit estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1974 housing inventory, as was described in the 1974 Current Housing Report, Series H-170, **Housing Characteristics for Selected Metropolitan Areas.** Since the 1974-1977 lost units existed, by definition, in the 1974 housing inventory, there was a 1974 housing inventory weight associated with each 1974-1977 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1974-1977 lost units.

1974 estimation procedure.—This report presents data on the housing characteristics of the 1974 housing inventory from the 1974 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1974.

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 to provide correct information on the part of respondents, mistakes in recording or coding the data, and other errors of collection, response, processing, coverage, and estimation for missing data. Nonsampling errors are not unique to sample surveys since they can, and do, occur in complete censuses as well.

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing, the 1977 AHS-SMSA sample, and the 1974 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 included space errors, definitional errors, and occupancy errors. The "content" errors measured the accuracy of the data collected for surveyed 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.—For the 1977 AHS-SMSA sample, a reinterview program was conducted to obtain a measurement of some of the components of the nonsampling error associated with the AHS estimates. A sample of households was revisited and answers to some of the questions on the AHS questionnaire were obtained again. The original interview and the reinterview were assumed to be two independent readings and, thus, were the basis for the measurement of the accuracy of the AHS data collected from interviewed households.

As part of the reinterview, a check was made at each of these households to determine if the following was done during the original interview:

1. The correct unit was visited.
2. The correct number of housing units were interviewed at that address.
3. The correct information on "Year Built" was obtained.
4. The correct information on "Tenure" was obtained.
5. The correct information on "Household Composition" was obtained.
6. The correct information on "Type of Housing Unit" was obtained.
7. The correct information on "Occupancy Status" was obtained.

This check was made for interviewer evaluation and control. That is, tolerance limits were derived to determine which interviewers passed or failed this reinterview check. The results of this study are available and are presented in the following Census Bureau memorandum, "Response Error in the Annual Housing Survey as Measured by Reinterviews—Year IV, 1977-1978 SMSA Sample."

For the A-2 reinterview program, measures of response inconsistency were taken on two subsamples—reconciled and unreconciled. After the question is answered in the reconciled subsample, the interviewer presented the original response and then asked the respondent to decide upon the best answer. Past experience has shown that the level of inconsistency is lower when the interviewer has the response available, and thus, it is believed that there is a downward bias in these estimates. In the unreconciled sample, the interviewer did not have access to previous responses. For the reconciled items, three-fourths showed low levels of response inconsistency while the remaining one-fourth were in the moderate range. For the unreconciled items, approximately one-half showed low levels of response disagreement, with the remaining one-half of the categories in the moderate and high inconsistency range. Although attitudinal items were excluded from the A-2 reinterview program, other reinterview programs have shown that, in general, attitudinal items such as items 102-104 (see facsimile of the 1977 AHS questionnaire, page App-34) have moderate to high levels of inconsistency. Moderate levels of inconsistency indicate that there is some problem with inconsistent reporting, and high levels indicate that improvements are needed in the method used to collect data or that the category concepts themselves are ambiguous.

A second measure of response error, the net difference rate, which is a measure of net error (or bias) in a particular category, was also provided by this reinterview program. None of the categories tested showed any significant evidence of bias.

The results of this study were based on sample data, so there is a sampling error associated with these estimates of nonsampling error. Therefore, the possibility of such errors should be taken into account when considering the results of this study.

A similar study was conducted for the 1974 AHS-SMSA sample. The results of which are presented in the following Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1974."

Coverage errors.—With respect to errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units, since, due to the relatively short time span involved, it is possible that construction of these units was not completed at the time the survey was conducted. In which case, they would not have been eligible for interview. In addition to these deficiencies, new construction in special places that do not require building permits, such as military bases, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit. Also, additional coverage errors may exist in SMSA's where the complete Coverage Improvement Program was not applied (see the table on page App-45 of this appendix).

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all units located inside these ED's would be represented in the sample. However, it has been estimated that the 1974 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. Since these ED's were recanvassed for the 1977 survey, the number of missed housing units may be considerably less for 1977.

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 schedules, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by the standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.

2. Approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the average result of all possible samples.

3. Approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimate would include the average result of all possible samples.

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

The figures presented in the following tables 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 1974 housing inventory can be found in the AHS Series H-170 reports for 1974.

Table I presents the standard errors applicable to estimates of characteristics of the 1977 housing inventory as well as estimates of characteristics of the 1974-1977 lost units. 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 20,448 for the total SMSA, 7,712 for the central city of the SMSA, and 18,667 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

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1977 Housing Inventory and for Estimated Number of 1974-1977 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]		
	SMSA	In central city	Not in central city
0	90	40	120
100	90	60	120
200	130	90	160
500	210	140	250
700	250	170	290
1,000	300	200	350
2,500	470	320	550
5,000	660	450	770
10,000	930	630	1,090
25,000	1,470	980	1,720
50,000	2,060	1,330	2,400
75,000	2,500	1,570	2,900
100,000	2,870	1,730	3,310
150,000	3,460	1,910	3,950
200,000	3,940	1,940	4,430
250,000	4,330	1,820	4,810
300,000	4,670	—	5,110
400,000	5,210	—	5,510
500,000	5,620	—	5,700
600,000	5,910	—	5,680
700,000	6,110	—	5,470
800,000	6,230	—	5,030
900,000	6,270	—	—
1,000,000	6,240	—	—
1,110,000	6,120	—	—

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

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 Lost Units for the Boston, Mass., SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	46.6	46.6	46.6	46.6	46.6	46.7
200	30.4	30.4	30.4	30.4	30.4	33.0
500	14.9	14.9	14.9	14.9	18.1	20.9
700	11.1	11.1	11.1	11.1	15.3	17.7
1,000	8.0	8.0	8.0	8.9	12.8	14.8
2,500	3.4	3.4	4.1	5.6	8.1	9.3
5,000	1.7	1.7	2.9	4.0	5.7	6.6
10,000	0.9	0.9	2.0	2.8	4.0	4.7
25,000	0.3	0.6	1.3	1.8	2.6	3.0
50,000	0.2	0.4	0.9	1.3	1.8	2.1
75,000	0.12	0.3	0.7	1.0	1.5	1.7
100,000	0.09	0.3	0.6	0.9	1.3	1.5
150,000	0.06	0.2	0.5	0.7	1.0	1.2
200,000	0.04	0.2	0.5	0.6	0.9	1.0
250,000	0.03	0.2	0.4	0.6	0.8	0.9
300,000	0.03	0.2	0.4	0.5	0.7	0.9
400,000	0.02	0.15	0.3	0.4	0.6	0.7
500,000	0.02	0.13	0.3	0.4	0.6	0.7
600,000	0.01	0.12	0.3	0.4	0.5	0.6
700,000	0.01	0.11	0.2	0.3	0.5	0.6
800,000	0.01	0.10	0.2	0.3	0.5	0.5
900,000	0.01	0.10	0.2	0.3	0.4	0.5
1,000,000	0.01	0.09	0.2	0.3	0.4	0.5
1,100,000	0.01	0.09	0.2	0.3	0.4	0.4

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.3.

numerators of the percentages, particularly if the percentages are 50 percent or more.

Tables II through IV present the standard errors of estimated percentages of the 1977 housing inventory as well as estimated percentages of the 1974-1977 lost units. 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 1977 there were 494,700 owner-occupied housing units in this SMSA. Interpolation in table I of this appendix shows that the standard error of an estimate of this size is approximately 5,600. The following interpolation procedure was used:

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

Size of estimate	Standard error
400,000	5,210
494,700	x
500,000	5,620

The entry for x is determined as follows by vertically interpolating between 5,210 and 5,620:

494,700—400,000 = 94,700
500,000—400,000 = 100,000

$$5{,}210 + \frac{94{,}700}{100{,}000}(5{,}620-5{,}210) = 5{,}600$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 489,100 to 500,300 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1977 owner-occupied housing units lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 485,740 to 503,660 housing units with 90 percent confidence; and that the average estimate lies within the interval from 483,500 to 505,900 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 494,700 owner-occupied housing units, 114,300, or 23.1 percent, had two bedrooms. Interpolation in table II of the appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 23.1 percent is approximately 0.6 percentage points. The following interpolation procedure was used:

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

Base of percentage	Estimated percentage		
	10 or 90	23.1	25 or 75
400,000 ...	0.4	a	0.6
494,700 ...		p	
500,000 ...	0.4	b	0.6

1. The entry for cell "a" is determined by horizontal interpolation between 0.4 and 0.6

23.1—10.0 = 13.1
25.0—10.0 = 15.0

$$0.4 + \frac{13.1}{15.0}(0.6-0.4) = 0.6$$

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

23.0—10.0 = 13.1
25.0—10.0 = 15.0

$$0.4 + \frac{13.1}{15.0}(0.6-0.4) = 0.6$$

TABLE III. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 Lost Housing Units for the Central City of the Boston, Mass., SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	29.1	29.1	29.1	29.1	29.1	32.0
200	17.0	17.0	17.0	17.0	19.6	22.6
500	7.6	7.6	7.6	8.6	12.4	14.3
700	5.5	5.5	5.5	7.3	10.5	12.1
1,000	3.9	3.9	4.4	6.1	8.8	10.1
2,500	1.6	1.6	2.8	3.8	5.5	6.4
5,000	0.8	0.9	2.0	2.7	3.9	4.5
10,000	0.4	0.6	1.4	1.9	2.8	3.2
25,000	0.2	0.4	0.9	1.2	1.8	2.0
50,000	0.08	0.3	0.6	0.9	1.2	1.4
75,000	0.05	0.2	0.5	0.7	1.0	1.2
100,000 ...	0.04	0.2	0.4	0.6	0.9	1.0
150,000 ...	0.03	0.2	0.4	0.5	0.7	0.8
200,000 ...	0.02	0.14	0.3	0.4	0.6	0.7
250,000 ...	0.02	0.13	0.3	0.4	0.6	0.6

[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.

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

$$494{,}700-400{,}000 = 94{,}700$$
$$500{,}000-400{,}000 = 100{,}000$$

$$0.6 + \frac{94{,}700}{100{,}000}(0.6-0.6) = 0.6$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 22.5 to 23.7 percent; the 90-percent confidence interval is from 22.1 to 24.1 percent; and the 95-percent interval is from 21.9 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. However, 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 1974 and 1977 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1974 and 1977 characteristics.

Illustration of the computation of the standard error of a difference.—Table A-1 of part A of this report shows that in 1977 there were 223,400 owner-occupied units with three bedrooms in this SMSA, Thus, the apparent difference, as shown by these data, between owner-occupied units with two bedrooms and owner-occupied units with three bedrooms is 109,100. Table I shows the standard error of 223,400 is approximately 4,120 and the standard error of 114,300 is approximately 3,040. Therefore, the standard error of the estimated difference of 109,100 is about:

$$5{,}120 = \sqrt{(4{,}120)^2 + (3{,}040)^2}$$

Consequently, the 68-percent confidence interval for the 109,100 difference is from 103,980 to 114,220 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of this difference lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 100,908 to 117,292 housing units, and the 95-percent confidence interval is from 98,860 to 119,340. Thus, we can conclude with 95 percent confidence that the number of 1977 owner-occupied 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

TABLE IV. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 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]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	54.6	54.6	54.6	54.6	54.6	54.9
200	37.6	37.6	37.6	37.6	37.6	38.8
500	19.4	19.4	19.4	19.4	21.3	24.5
700	14.7	14.7	14.7	14.7	18.0	20.7
1,000	10.7	10.7	10.7	10.7	15.0	17.4
2,500	4.6	4.6	4.8	6.6	9.5	11.0
5,000	2.4	2.4	3.4	4.7	6.7	7.8
10,000	1.2	1.2	2.4	3.3	4.8	5.5
25,000	0.5	0.7	1.5	2.1	3.0	3.5
50,000	0.2	0.5	1.1	1.5	2.1	2.5
75,000	0.2	0.4	0.9	1.2	1.7	2.0
100,000	0.12	0.3	0.8	1.0	1.5	1.7
150,000	0.08	0.3	0.6	0.9	1.2	1.4
200,000	0.06	0.2	0.5	0.7	1.1	1.2
250,000	0.05	0.2	0.5	0.7	1.0	1.1
300,000	0.04	0.2	0.4	0.6	0.9	1.0
400,000	0.03	0.2	0.4	0.5	0.8	0.9
500,000	0.02	0.15	0.3	0.5	0.7	0.8
600,000	0.02	0.14	0.3	0.4	0.6	0.7
700,000	0.02	0.13	0.3	0.4	0.6	0.7
800,000	0.02	0.12	0.3	0.4	0.5	0.6

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

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, read off the confidence interval corresponding to the two points established in step 2.

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

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

Illustration of the computation of the 95-percent confidence interval of a median.—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 3.1. The base of the distribution from which this median was determined is 494,700 housing units.

1. Interpolation in table II shows that the standard error of 50 percent on a base of 494,700 is approximately 0.7 percentage points.

2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.6 and 51.4.

3. From table A-1 of part A, it can be seen by cumulating the frequencies for the first two categories that 195,400 owner-occupied housing units, or 39.5 percent, had one or two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) and that an additional 87,200 owner-occupied housing units, or 17.6 percent, had three persons (i.e., 2.5 to 3.5 persons).

By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{48.6-39.5}{17.6}\right) = 3.0$$

Similarly, the uppper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{51.4-39.5}{17.6}\right) = 3.2$$

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

Dallas 1977

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1977 survey App-42
Selection of the 1974 sample App-42
1977 sample reduction App-43
1974-1977 additions to the housing inventory App-43
Coverage improvement sample selection App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1970 Census of Population and Housing App-45
ESTIMATION App-45
1977 housing inventory App-45
1974-1977 lost units App-46
1974 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES App-46
Nonsampling errors App-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-50
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-51

SAMPLE DESIGN

Annual Housing Survey.—The estimates for each of the 20 SMSA's are based on data collected from the 1977 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 20 SMSA's, the data were collected for the 11-month period from April 1977 through February 1978 with one-eleventh of the sample units visited each month.

Each group of SMSA's is interviewed for the AHS on a rotating basis, and this group (A-2 SMSA's) was the first to be revisited. Nineteen of the 20 SMSA's were interviewed for the first time in 1974, while the Madison, Wis., SMSA was interviewed for the first time in 1975.

In each group of SMSA's, the largest SMSA from each of the 4 geographic regions is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All remaining SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the A-2 group (1977-1978) are: Boston, Mass.; Detroit, Mich.; Los Angeles-Long Beach, Calif.; and Washington, D.C.-Md.-Va.

The remaining SMSA's in the A-2 group are: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.

In this SMSA, 4,691 housing units were eligible for interview. Of these sample units, 183 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 units eligible for interview, 475 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 1977 survey.—The sample housing units designated for interview in the 1977 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 1974 survey which were not part of the 1977 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 1974 survey which were not part of the 1977 reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1977 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from the list of new construction building permits issued since the 1974 survey. (This sample represented the housing units built in permit-issuing areas, since the 1974 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1974 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1974 survey.)
5. All sample housing units that were selected as part of the 1977 Coverage Improvement Program. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1974 sample.—The sample for the SMSA's which are 100-percent permit-issuing was selected from two sample frames—units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which

are not 100-percent permit-issuing included a sample selected from a third frame—those units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). The following SMSA's are 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Los Angeles-Long Beach, Calif.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 14 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 four largest SMSA's, the overall sampling rate differed by central city and 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 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 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 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-$5,999. . .		
$6,000-$9,999. . .		
$10,000-$14,999 .		
$15,000 and over .		

Thus, for this SMSA, the occupied housing unit records from this 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 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 are 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 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 units in existence at the time of interview in these selected segments are in sample. Thus, units enumerated in the 1970 census as well as units built since the 1970 census are included.

1977 sample reduction.—In previous years, AHS sample units have been divided among 12 panels which were interviewed each month. Due to budget limitations, it became necessary to drop one of these panels from the 1977 survey. The designated sample size for 1977 was thus reduced by one-twelfth.

1974-1977 additions to the housing inventory.—In the permit-issuing universe, a

sample of new construction building permits, issued since the 1974 survey, was selected to represent housing units built in permit-issuing areas since the 1974 survey. Sampling procedures were identical to those used in selecting the 1970-1974 new construction sample. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1974, to identify any units missed in the 1974 survey or any units added since the 1974 survey.

Coverage improvement sample selection.—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 census.
3. Units missed in the 1970 census.
4. 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.

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 A-2 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 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: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Orlando, Fla.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; and Wichita, Kans.

In the remaining 10 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 and the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 8,288 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, for each 1977-1978 SMSA, 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 2,805 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 no 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 units in multi-unit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multi-unit structures selected, all units were listed and matched to the 1970 census. Any missed units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply these two procedures in all 20 A-2 SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in three SMSA's. In these SMSA's, a string of 10 structures (instead of 8) that were eligible to be selected for the AHS, were listed. The full application of this technique was employed in 12 SMSA's, with 5 SMSA's not receiving any application of the first procedure. The second procedure was instituted in 13 SMSA's. The following table shows which of these two techniques were applied to the A-2 SMSA's.

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 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 1977 housing inventory) and estimates pertaining to characteristics of units removed from the housing inventory since 1974 (i.e., 1974-1977 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1977 housing inventory.—The AHS estimates of characteristics of the 1977 housing inventory employed a 1-stage ratio estimation procedure in 19 SMSA's, with a second-stage ratio estimation procedure being employed in 1 SMSA. Prior to the implementation of the ratio estimation procedure, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of 1 or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample 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 estima-

SMSA	Coverage improvement technique employed			
	First		Second	Neither
	Partial	Full		
Albany-Schenectady-Troy, N.Y	X		X	
Anaheim-Santa Ana-Garden Grove, Calif	X			
Boston, Mass		X	X	
Dallas, Tex				X
Detroit, Mich		X	X	
Fort Worth, Tex		X	X	
Los Angeles-Long Beach, Calif				X
Madison, Wis		X	X	
Memphis, Tenn.-Ark		X	X	
Minneapolis-St. Paul, Minn				X
Newark, N.J		X	X	
Orlando, Fla		X	X	
Phoenix, Ariz		X	X	
Pittsburgh, Pa				X
Saginaw, Mich		X	X	
Salt Lake City, Utah				X
Spokane, Wash	X			
Tacoma, Wash		X	X	
Washington, D.C.-Md.-Va		X	X	
Wichita, Kans		X	X	

tion factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of 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 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 unit within the corresponding ratio estimation category.

The effect of this ratio estimation procedure was to somewhat reduce the variance due to the variation in sample size of strata used in the sample selection of the permit-issuing universe, thereby reducing the sampling error below what would have been obtained by simply weighting the results of the sample by the inverse of the probability of selection. Ordinarily, this would have been controlled by sampling within the strata during the sample selection process. However, prior to the AHS sample selection within each SMSA, units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. Thus, some variation in sample size was introduced during the AHS sample selection process.

For the Pittsburgh, Pa., SMSA, a second-stage ratio estimation procedure was also employed. This procedure involved the ratio estimation of the AHS weighted sample estimate of the October 1977 housing inventory to an independent estimate of the SMSA's October 1977 housing inventory. This ratio estimate factor equaled the following:

$$\frac{\text{SMSA independent estimate of October 1977 housing inventory}}{\text{AHS sample estimate of the housing inventory}}$$

The numerator of the ratio was derived by using the 1970 census estimate of total housing units for the SMSA in conjunction with the estimate of the change in the housing inventory since the census. The estimate of change was based on administrative records from utility companies. The denominator of that ratio was obtained from the weighted estimate of AHS sample units, using the existing weight after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for each sample unit and the resulting product was used as the final weight for tabulation purposes.

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

The second-stage ratio estimation procedure was not employed in all A-2 SMSA's, since it was felt that coverage improvement procedures adequately corrected known deficiencies in the sample where full coverage improvement procedures were employed. The independent estimates were only employed in SMSA's where full coverage improvements were not used and the estimated relative bias of the independent estimate was thought to be low enough (i.e., 2 percent or less over a 10-year period) to preclude the possibility of doing more harm than good to the survey results. A measure of the relative bias of the independent estimate was available based on a comparison between the 1970 census estimate of the April 1970 housing inventory and the independent estimate of the April 1970 housing inventory which was projected over a 10-year period from the 1960 census using the same procedure that produced the October 1977 independent estimates.

1974-1977 lost units.—The 1974-1977 lost unit estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1974 housing inventory, as was described in the 1974 Current Housing Report, Series H-170, **Housing Characteristics for Selected Metropolitan Areas.** Since the 1974-1977 lost units existed, by definition, in the 1974 housing inventory, there was a 1974 housing inventory weight associated with each 1974-1977 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1974-1977 lost units.

1974 estimation procedure.—This report presents data on the housing characteristics of the 1974 housing inventory from the 1974 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1974.

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 to provide correct information on the part of respondents, mistakes in recording or coding the data, and other errors of collection, response, processing, coverage, and estimation for missing data. Nonsampling errors are not unique to sample surveys since they can, and do, occur in complete censuses as well.

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing, the 1977 AHS-SMSA sample, and the 1974 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 included space errors, definitional errors, and occupancy errors. The "content" errors measured the accuracy of the data collected for surveyed 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.—For the 1977 AHS-SMSA sample, a reinterview program was conducted to obtain a measurement of some of the components of the nonsampling error associated with the AHS estimates. A sample of households was revisited and answers to some of the questions on the AHS questionnaire were obtained again. The original interview and the reinterview were assumed to be two independent readings and, thus, were the basis for the measurement of the accuracy of the AHS data collected from interviewed households.

As part of the reinterview, a check was made at each of these households to determine if the following was done during the original interview:

1. The correct unit was visited.
2. The correct number of housing units were interviewed at that address.
3. The correct information on "Year Built" was obtained.
4. The correct information on "Tenure" was obtained.
5. The correct information on "Household Composition" was obtained.
6. The correct information on "Type of Housing Unit" was obtained.
7. The correct information on "Occupancy Status" was obtained.

This check was made for interviewer evaluation and control. That is, tolerance limits were derived to determine which interviewers passed or failed this reinterview check. The results of this study are available and are presented in the following Census Bureau memorandum, "Response Error in the Annual Housing Survey as Measured by Reinterviews—Year IV, 1977-1978 SMSA Sample."

For the A-2 reinterview program, measures of response inconsistency were taken on two subsamples—reconciled and unreconciled. After the question is answered in the reconciled subsample, the interviewer presented the original response and then asked the respondent to decide upon the best answer. Past experience has shown that the level of inconsistency is lower when the interviewer has the response available, and thus, it is believed that there is a downward bias in these estimates. In the unreconciled sample, the interviewer did not have access to previous responses. For the reconciled items, three-fourths showed low levels of response inconsistency while the remaining one-fourth were in the moderate range. For the unreconciled items, approximately one-half showed low levels of response disagreement, with the remaining one-half of the categories in the moderate and high inconsistency range. Although attitudinal items were excluded from the A-2 reinterview program, other reinterview programs have shown that, in general, attitudinal items such as items 102-104 (see facsimile of the 1977 AHS questionnaire, page App-34) have moderate to high levels of inconsistency. Moderate levels of inconsistency indicate that there is some problem with inconsistent reporting, and high levels indicate that improvements are needed in the method used to collect data or that the category concepts themselves are ambiguous.

A second measure of response error, the net difference rate, which is a measure of net error (or bias) in a particular category, was also provided by this reinterview program. None of the categories tested showed any significant evidence of bias.

The results of this study were based on sample data, so there is a sampling error associated with these estimates of nonsampling error. Therefore, the possibility of such errors should be taken into account when considering the results of this study.

A similar study was conducted for the 1974 AHS-SMSA sample. The results of which are presented in the following Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1974."

Coverage errors.—With respect to errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing

units, since, due to the relatively short time span involved, it is possible that construction of these units was not completed at the time the survey was conducted. In which case, they would not have been eligible for interview. In addition to these deficiencies, new construction in special places that do not require building permits, such as military bases, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit. Also, additional coverage errors may exist in SMSA's where the complete Coverage Improvement Program was not applied (see the table on page App-45 of this appendix).

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all units located inside these ED's would be represented in the sample. However, it has been estimated that the 1974 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. Since these ED's were recanvassed for the 1977 survey, the number of missed housing units may be considerably less for 1977.

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 schedules, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by the standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.

2. Approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the average result of all possible samples.

3. Approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimate would include the average result of all possible samples.

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

The figures presented in the following tables 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 1974 housing inventory can be found in the AHS Series H-170 reports for 1974.

Table I presents the standard errors applicable to estimates of characteristics of the 1977 housing inventory as well as estimates of characteristics of the 1974-1977 lost units. 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 31,610 for the total SMSA, 18,200 for the central city of the SMSA, and 25,750 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 presents the standard errors of estimated percentages of the 1977 housing inventory as well as estimated percentages of the 1974-1977 lost units. 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 1977 there were 372,700 owner-occupied housing units in this SMSA. Interpolation in table I of this appendix shows that the standard error of an estimate of this size is approximately 7,900. 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	7,040
372,700	x
400,000	8,220

The entry for x is determined as follows by vertically interpolating between 7,040 and 8,220:

372,700—300,000 = 72,700
400,000—300,000 = 100,000

$$7{,}040+\frac{72{,}700}{100{,}000}(8{,}220-7{,}040)$$
$$=7{,}900$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 364,800 to 380,600 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1977 owner-occupied housing units lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 360,060 to 385,340 housing units with 90 percent confidence; and that the average estimate lies within the interval from 356,900 to 388,500 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 372,700 owner-occupied housing units, 89,700, or 24.1 percent, had two bedrooms. Interpolation in table II of the appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 24.1 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	24.1	25
300,000	0.7	a	1.0
372,700		p	
400,000	0.6	b	0.8

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

24.1—10.0 = 14.1
25.0—10.0 = 15.0

$$0.7+\frac{14.1}{15.0}(1.0-0.7)=0.98$$

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

24.1—10.0 = 14.1
25.0—10.0 = 15.0

$$0.6+\frac{14.1}{15.0}(0.8-0.6)=0.79$$

3. The entry for "p" was then determined by vertical interpolation between 0.98 and 0.79.

372,700—300,000 = 72,700
400,000—300,000 = 100,000

$$0.98+\frac{72{,}700}{100{,}000}(0.79-0.98)=0.8$$

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1977 Housing Inventory and for Estimated Number of 1974-1977 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]		
	SMSA	In central city	Not in central city
0	150	160	140
100	150	160	140
200	180	180	170
500	280	280	270
700	330	330	320
1,000 ...	390	390	380
2,500 ...	620	620	600
5,000 ...	880	880	850
10,000 ..	1,240	1,240	1,210
25,000 ..	1,970	1,950	1,940
50,000 ..	2,790	2,720	2,820
75,000 ..	3,430	3,280	3,540
100,000 .	3,970	3,730	4,180
150,000 .	4,890	4,420	5,350
200,000 .	5,680	4,930	6,440
250,000 .	6,390	5,320	7,480
300,000 .	7,040	5,600	8,480
400,000 .	8,220	5,910	10,430
500,000 .	9,290	—	—
600,000 .	10,280	—	—
700,000 .	11,220	—	—

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.1 for the total SMSA, for the central city, and for the balance (not in central city) estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 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)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	60.6	60.6	60.6	60.6	60.6	62.0
200	43.5	43.5	43.5	43.5	43.5	43.9
500	23.5	23.5	23.5	23.5	24.0	27.7
700	18.0	18.0	18.0	18.0	20.3	23.4
1,000	13.3	13.3	13.3	13.3	17.0	19.6
2,500	5.8	5.8	5.8	7.4	10.7	12.4
5,000	3.0	3.0	3.8	5.3	7.6	8.8
10,000	1.5	1.5	2.7	3.7	5.4	6.2
25,000	0.6	0.8	1.7	2.4	3.4	3.9
50,000	0.3	0.6	1.2	1.7	2.4	2.8
75,000	0.2	0.5	1.0	1.4	2.0	2.3
100,000	0.15	0.4	0.9	1.2	1.7	2.0
150,000	0.10	0.3	0.7	1.0	1.4	1.6
200,000	0.08	0.3	0.6	0.8	1.2	1.4
250,000	0.06	0.2	0.5	0.7	1.1	1.2
300,000	0.05	0.2	0.5	0.7	1.0	1.1
400,000	0.04	0.2	0.4	0.6	0.8	1.0
500,000	0.03	0.2	0.4	0.5	0.8	0.9
600,000	0.03	0.2	0.3	0.5	0.7	0.8
700,000	0.02	0.15	0.3	0.4	0.6	0.7

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.1 for the total SMSA, for the central city, and for the balance (not in central city) estimates.

Consequently, the 68-percent confidence interval, as shown by these data, is from 23.3 to 24.9 percent; the 90-percent confidence interval is from 22.8 to 25.4 percent; and the 95-percent confidence interval is from 22.5 to 25.7 percent.

Differences.—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. However, 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 1974 and 1977 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1974 and 1977 characteristics.

Illustration of the computation of the standard error of a difference.—Table A-1 of part A of this report shows that in 1977 there were 203,300 owner-occupied units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied units with two bedrooms and owner-occupied units with three bedrooms is 113,600. Table I shows the standard error of 203,300 is approximately 5,730, and the standard error of 89,700 is approximately 3,750. Therefore, the standard error of the estimated difference of 113,600 is about:

$$6{,}850 = \sqrt{(5{,}730)^2 + (3{,}750)^2}$$

Consequently, the 68-percent confidence interval for the 113,600 difference is from 106,750 to 120,450 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 102,640 to 124,560 housing units, and the 95-percent confidence interval is from 99,900 to 127,300. Thus, we can conclude with 95 percent confidence that the number of 1977 owner-occupied 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, read off the confidence interval corresponding to the two points established in step 2.

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 372,700 housing units.

1. Interpolation in table II shows that the standard error of 50 percent on a base of 372,700 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 table A-1 of part A, it can be seen by cumulating the frequencies for the first two categories that 166,600 owner-occupied housing units, or 44.7 percent, had one or two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) and that an additional 74,500 owner-occupied housing units, or 20.0 percent, had three persons (i.e., 2.5 to 3.5 persons).

By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{48.0-44.7}{20.0} = 2.7$$

Similarly, the upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{52.0-44.7}{20.0} = 2.9$$

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

Detroit 1977

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1977 survey App-42
Selection of the 1974 sample ... App-42
1977 sample reduction App-43
1974-1977 additions to the housing inventory App-44
Coverage improvement sample selection App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1970 Census of Population and Housing App-45
ESTIMATION App-45
1977 housing inventory App-45
1974-1977 lost units App-46
1974 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES App-46
Nonsampling errors App-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-50
Differences App-51
Illustration of the computation of the standard error of a difference App-51
Medians App-52
Illustration of the computation of the 95-percent confidence interval of a median ... App-52

SAMPLE DESIGN

Annual Housing Survey.—The estimates for each of the 20 SMSA's are based on data collected from the 1977 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 20 SMSA's, the data were collected for the 11-month period from April 1977 through February 1978 with one-eleventh of the sample units visited each month.

Each group of SMSA's is interviewed for the AHS on a rotating basis, and this group (A-2 SMSA's) was the first to be revisited. Nineteen of the 20 SMSA's were interviewed for the first time in 1974, while the Madison, Wis., SMSA was interviewed for the first time in 1975.

In each group of SMSA's, the largest SMSA from each of the 4 geographic regions is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All remaining SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the A-2 group (1977-1978) are: Boston, Mass.; Detroit, Mich.; Los Angeles-Long Beach, Calif.; and Washington, D.C.-Md.-Va.

The remaining SMSA's in the A-2 group are: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.

In this SMSA, 13,394 housing units were eligible for interview. Of these sample units, 673 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 units eligible for interview, 892 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 1977 survey.—The sample housing units designated for interview in the 1977 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 1974 survey which were not part of the 1977 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 1974 survey which were not part of the 1977 reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1977 AHS questionnaire, page App-20.)

3. All sample housing units that were selected from the list of new construction building permits issued since the 1974 survey. (This sample represented the housing units built in permit-issuing areas since the 1974 survey.)

4. All sample housing units that were added to sample segments in the nonpermit universe since the 1974 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1974 survey.)

5. All sample housing units that were selected as part of the 1977 Coverage Improvement Program. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1974 sample.—The sample for the SMSA's which are 100-percent permit-issuing was selected from two sample frames—units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and units constructed in

permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which are not 100-percent permit-issuing included a sample selected from a third frame—those units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). The following SMSA's are 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Los Angeles-Long Beach, Calif.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 14 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 four largest SMSA's, the overall sampling rate differed by central city and 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 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 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 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-$5,999. . .		
$6,000-$9,999. . .		
$10,000-$14,999 .		
$15,000 and over .		

Thus, for this SMSA, the occupied housing unit records from this 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 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 are 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 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 units in existence at the time of interview in these selected segments are in sample. Thus, units enumerated in the 1970 census as well as units built since the 1970 census are included.

1977 sample reduction.—In previous years, AHS sample units have been divided among 12 panels which were interviewed each month. Due to budget limitations, it became necessary to drop one of these panels from the 1977 survey. The designated sample size for 1977 was thus reduced by one-twelfth.

1974-1977 additions to the housing inventory.—In the permit-issuing universe, a sample of new construction building permits, issued since the 1974 survey, was selected to represent housing units built in permit-issuing areas since the 1974 survey. Sampling procedures were identical to those used in selecting the 1970-1974 new construction sample. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1974, to identify any units missed in the 1974 survey or any units added since the 1974 survey.

Coverage improvement sample selection.—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 census.
3. Units missed in the 1970 census.
4. 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.

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 A-2 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 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: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Orlando, Fla.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; and Wichita, Kans.

In the remaining 10 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 and the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 13,249 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, for each 1977-1978 SMSA, 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 6,245 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 no 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 units in multi-unit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multi-unit structures selected, all units were listed and matched to the 1970 census. Any missed units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply these two procedures in all 20 A-2 SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in three SMSA's. In these SMSA's, a string of 10 structures (instead of 8) that were eligible to be selected for the AHS, were listed. The full application of this technique was employed in 12 SMSA's, with 5 SMSA's not receiving any application of the first procedure. The second procedure was instituted in 13 SMSA's. The following table shows which of these two techniques were applied to the A-2 SMSA's.

The first procedure added an estimated 9,407 units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 1,770 units.

SMSA	Coverage improvement technique employed			
	First		Second	Neither
	Partial	Full		
Albany-Schenectady-Troy, N.Y	X		X	
Anaheim-Santa Ana-Garden Grove, Calif	X			
Boston, Mass		X	X	
Dallas, Tex				X
Detroit, Mich		X	X	
Fort Worth, Tex		X	X	
Los Angeles-Long Beach, Calif				X
Madison, Wis		X	X	
Memphis, Tenn.-Ark		X	X	
Minneapolis-St. Paul, Minn				X
Newark, N.J		X	X	
Orlando, Fla		X	X	
Phoenix, Ariz		X	X	
Pittsburgh, Pa				X
Saginaw, Mich		X	X	
Salt Lake City, Utah				X
Spokane, Wash	X			
Tacoma, Wash		X	X	
Washington, D.C.-Md.-Va		X	X	
Wichita, Kans		X	X	

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 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 1977 housing inventory) and estimates pertaining to characteristics of units removed from the housing inventory since 1974 (i.e., 1974-1977 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1977 housing inventory.—The AHS estimates of characteristics of the 1977 housing inventory employed a 1-stage ratio estimation procedure in 19 SMSA's, with a second-stage ratio estimation procedure being employed in 1 SMSA. Prior to the implementation of the ratio estimation procedure, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of 1 or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample 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 hous-

ing 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:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of 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 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 unit within the corresponding ratio estimation category.

The effect of this ratio estimation procedure was to somewhat reduce the variance due to the variation in sample size of strata used in the sample selection of the permit-issuing universe, thereby reducing the sampling error below what would have been obtained by simply weighting the results of the sample by the inverse of the probability of selection. Ordinarily, this would have been controlled by sampling within the strata during the sample selection process. However, prior to the AHS sample selection within each SMSA, units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. Thus, some variation in sample size was introduced during the AHS sample selection process.

For the Pittsburgh, Pa., SMSA, a second-stage ratio estimation procedure was also employed. This procedure involved the ratio estimation of the AHS weighted sample estimate of the October 1977 housing inventory to an independent estimate of the SMSA's October 1977 housing inventory. This ratio estimate factor equaled the following:

$$\frac{\text{SMSA independent estimate of October 1977 housing inventory}}{\text{AHS sample estimate of the housing inventory}}$$

The numerator of the ratio was derived by using the 1970 census estimate of total housing units for the SMSA in conjunction with the estimate of the change in the housing inventory since the census. The estimate of change was based on administrative records from utility companies. The denominator of that ratio was obtained from the weighted estimate of AHS sample units, using the existing weight after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for each sample unit and the resulting product was used as the final weight for tabulation purposes.

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

The second-stage ratio estimation procedure was not employed in all A-2 SMSA's, since it was felt that coverage improvement procedures adequately corrected known deficiencies in the sample where full coverage improvement procedures were employed. The independent estimates were only employed in SMSA's where full coverage improvements were not used and the estimated relative bias of the independent estimate was thought to be low enough (i.e., 2 percent or less over a 10-year period) to preclude the possibility of doing more harm than good to the survey results. A measure of the relative bias of the independent estimate was available based on a comparison between the 1970 census estimate of the April 1970 housing inventory and the independent estimate of the April 1970 housing inventory which was projected over a 10-year period from the 1960 census using the same procedure that produced the October 1977 independent estimates.

1974-1977 lost units.—The 1974-1977 lost unit estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1974 housing inventory, as was described in the 1974 Current Housing Report, Series H-170, **Housing Characteristics for Selected Metropolitan Areas.** Since the 1974-1977 lost units existed, by definition, in the 1974 housing inventory, there was a 1974 housing inventory weight associated with each 1974-1977 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1974-1977 lost units.

1974 estimation procedure.—This report presents data on the housing characteristics of the 1974 housing inventory from the 1974 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1974.

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 non-

sampling 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 to provide correct information on the part of respondents, mistakes in recording or coding the data, and other errors of collection, response, processing, coverage, and estimation for missing data. Nonsampling errors are not unique to sample surveys since they can, and do, occur in complete censuses as well.

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing, the 1977 AHS-SMSA sample, and the 1974 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 included space errors, definitional errors, and occupancy errors. The "content" errors measured the accuracy of the data collected for surveyed 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.—For the 1977 AHS-SMSA sample, a reinterview program was conducted to obtain a measurement of some of the components of the nonsampling error associated with the AHS estimates. A sample of households was revisited and answers to some of the questions on the AHS questionnaire were obtained again. The original interview and the reinterview were assumed to be two independent readings and, thus, were the basis for the measurement of the accuracy of the AHS data collected from interviewed households.

As part of the reinterview, a check was made at each of these households to determine if the following was done during the original interview:

1. The correct unit was visited.
2. The correct number of housing units were interviewed at that address.
3. The correct information on "Year Built" was obtained.
4. The correct information on "Tenure" was obtained.
5. The correct information on "Household Composition" was obtained.
6. The correct information on "Type of Housing Unit" was obtained.
7. The correct information on "Occupancy Status" was obtained.

This check was made for interviewer evaluation and control. That is, tolerance limits were derived to determine which interviewers passed or failed this reinterview check. The results of this study are available and are presented in the following Census Bureau memorandum, "Response Error in the Annual Housing Survey as Measured by Reinterviews—Year IV, 1977-1978 SMSA Sample."

For the A-2 reinterview program, measures of response inconsistency were taken on two subsamples—reconciled and unreconciled. After the question is answered in the reconciled subsample, the interviewer presented the original response and then asked the respondent to decide upon the best answer. Past experience has shown that the level of inconsistency is lower when the interviewer has the response available, and thus, it is believed that there is a downward bias in these estimates. In the unreconciled sample, the interviewer did not have access to previous responses. For the reconciled items, three-fourths showed low levels of response inconsistency while the remaining one-fourth were in the moderate range. For the unreconciled items, approximately one-half showed low levels of response disagreement, with the remaining one-half of the categories in the moderate and high inconsistency range. Although attitudinal items were excluded from the A-2 reinterview program, other reinterview programs have shown that, in general, attitudinal items such as items 102-104 (see facsimile of the 1977 AHS questionnaire, page App-34) have moderate to high levels of inconsistency. Moderate levels of inconsistency indicate that there is some problem with inconsistent reporting, and high levels indicate that improvements are needed in the method used to collect data or that the category concepts themselves are ambiguous.

A second measure of response error, the net difference rate, which is a measure of net error (or bias) in a particular category, was also provided by this reinterview program. None of the categories tested showed any significant evidence of bias.

The results of this study were based on sample data, so there is a sampling error associated with these estimates of nonsampling error. Therefore, the possibility of such errors should be taken into account when considering the results of this study.

A similar study was conducted for the 1974 AHS-SMSA sample. The results of which are presented in the following Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1974."

Coverage errors.—With respect to errors of coverage and estimation for missing data, the AHS new construction sample

had deficiencies in the representation of conventional new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units, since, due to the relatively short time span involved, it is possible that construction of these units was not completed at the time the survey was conducted. In which case, they would not have been eligible for interview. In addition to these deficiencies, new construction in special places that do not require building permits, such as military bases, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit. Also, additional coverage errors may exist in SMSA's where the complete Coverage Improvement Program was not applied (see the table on page App-45 of this appendix).

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all units located inside these ED's would be represented in the sample. However, it has been estimated that the 1974 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. Since these ED's were recanvassed for the 1977 survey, the number of missed housing units may be considerably less for 1977.

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 schedules, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by the standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.

2. Approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the average result of all possible samples.

3. Approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimate would include the average result of all possible samples.

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

The figures presented in the following tables 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 1974 housing inventory can be found in the AHS Series H-170 reports for 1974.

Table I presents the standard errors applicable to estimates of characteristics of the 1977 housing inventory as well as estimates of characteristics of the 1974-1977 lost units. 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 33,360 for the total SMSA, 12,660 for the central city of the SMSA, and 30,660 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, particu-

larly if the percentages are 50 percent or more.

Tables II through IV present the standard errors of estimated percentages of the 1977 housing inventory as well as estimated percentages of the 1974-1977 lost units. 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

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1977 Housing Inventory and for Estimated Number of 1974-1977 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]		
	SMSA	In central city	Not in central city
0	110	80	150
100	110	90	150
200	150	130	170
500	240	200	270
700	280	240	320
1,000	330	290	380
2,500	530	460	600
5,000	750	650	850
10,000	1,060	910	1,200
25,000	1,670	1,420	1,900
50,000	2,360	1,970	2,680
75,000	2,880	2,360	3,270
100,000	3,330	2,670	3,760
150,000	4,060	3,120	4,580
200,000	4,670	3,420	5,260
250,000	5,200	3,610	5,840
300,000	5,680	3,710	6,350
400,000	6,500	3,650	7,240
500,000	7,220	3,210	7,980
600,000	7,850	—	8,620
700,000	8,410	—	9,180
800,000	8,920	—	9,670
900,000	9,380	—	10,100
1,000,000	9,810	—	10,480
1,100,000	10,200	—	—
1,200,000	10,570	—	—
1,300,000	10,900	—	—
1,400,000	11,210	—	—
1,500,000	11,510	—	—

[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, 1.6 for the central city, and 1.3 for the balance (Not in central city).

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 Lost Units for the Detroit, Mich., SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	52.9	52.9	52.9	52.9	52.9	53.0
200	35.9	35.9	35.9	35.9	35.9	37.4
500	18.3	18.3	18.3	18.3	20.5	23.7
700	13.8	13.8	13.8	13.8	17.3	20.0
1,000	10.1	10.1	10.1	10.1	14.5	16.7
2,500	4.3	4.3	4.6	6.4	9.2	10.6
5,000	2.2	2.2	3.3	4.5	6.5	7.5
10,000	1.1	1.1	2.3	3.2	4.6	5.3
25,000	.4	.7	1.5	2.0	2.9	3.3
50,000	.2	.5	1.0	1.4	2.1	2.4
75,000	.1	.4	.8	1.2	1.7	1.9
100,000	.11	.3	.7	1.0	1.5	1.7
150,000	.07	.3	.6	.8	1.2	1.4
200,000	.06	.2	.5	.7	1.0	1.2
250,000	.04	.2	.5	.6	.9	1.1
300,000	.04	.2	.4	.6	.8	1.0
400,000	.03	.2	.4	.5	.7	.8
500,000	.02	.1	.3	.4	.6	.7
600,000	.02	.14	.3	.4	.6	.7
700,000	.02	.13	.3	.4	.5	.6
800,000	.01	.12	.3	.4	.5	.6
900,000	.01	.11	.2	.3	.5	.6
1,000,000	.01	.11	.2	.3	.5	.5
1,100,000	.01	.10	.2	.3	.4	.5
1,200,000	.01	.10	.2	.3	.4	.5
1,300,000	.01	.09	.2	.3	.4	.5
1,400,000	.01	.09	.2	.3	.4	.4
1,500,000	.01	.09	.2	.3	.4	.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 1.4.

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 1977 there were 979,000 owner-occupied housing units in this SMSA. Interpolation in table I of this appendix shows that the standard error of an estimate of this size is approximately 9,720. The following interpolation procedure was used:

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

Size of estimate	Standard error
900,000	9,380
979,000	x
100,000	9,810

The entry for x is determined as follows by vertically interpolating between 9,380 and 9,810:

979,000–900,000 = 79,000
1,000,000–900,000 = 100,000

$$9{,}380 + \frac{79{,}000}{100{,}000}(9{,}810-9{,}380)$$
$$= 9{,}720$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 969,280 to 988,720 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1977 owner-occupied housing units lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 963,448 to 994,552 housing units with 90 percent confidence; and that the average estimate lies within the interval from 959,560 to 998,440 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 979,000 owner-occupied housing units, 241,200, or 24.6 percent, had two bedrooms. Interpolation in table II of the appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 24.6 percent is approximately 0.5 percentage points. The following interpolation procedure was used:

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

Base of percentage	Estimated percentage		
	10 or 90	24.6	25 or 75
900,000 ...	0.3	a	0.5
979,000 ...		p	
1,000,000 ..	0.3	b	0.5

TABLE III. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 Lost Housing Units for the Central City of the Detroit, Mich., SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	45.6	45.6	45.6	45.6	45.6	45.8
200	29.5	29.5	29.5	29.5	29.5	32.4
500	14.4	14.4	14.4	14.4	17.7	20.5
700	10.7	10.7	10.7	10.7	15.0	17.3
1,000	7.7	7.7	7.7	8.7	12.5	14.5
2,500	3.2	3.2	4.0	5.5	7.9	9.2
5,000	1.6	1.6	2.8	3.9	5.6	6.5
10,000	.8	.9	2.0	2.7	4.0	4.6
25,000	.3	.6	1.3	1.7	2.5	2.9
50,000	.2	.4	.9	1.2	1.8	2.0
75,000	.11	.3	.7	1.0	1.4	1.7
100,000 ...	.08	.3	.6	.9	1.3	1.4
150,000 ...	.06	.2	.5	.7	1.0	1.2
200,000 ...	.04	.2	.4	.6	.9	1.0
250,000 ...	.03	.2	.4	.5	.8	.9
300,000 ...	.03	.2	.4	.5	.7	.8
400,000 ...	.02	.14	.3	.4	.6	.7
500,000 ...	.02	.13	.3	.4	.6	.6

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.6.

1. The entry for cell "a" is determined by horizontal interpolation between 0.3 and 0.5

$$24.6-10.0 = 14.6$$
$$25.0-10.0 = 15.0$$
$$0.3 + \frac{14.6}{15.0}(0.5-0.3) = 0.5$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.3 and 0.5.

$$24.6-10 = 14.6$$
$$25.0-10 = 15.0$$
$$0.3 + \frac{14.6}{15.0}(0.5-0.3) = 0.5$$

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

$$979{,}000-900{,}000 = 79{,}000$$
$$1{,}000{,}000-900{,}000 = 100{,}000$$
$$0.5 + \frac{79{,}000}{100{,}000}(0.5-0.5) = 0.5$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 24.1 to 25.1 percent; the 90-percent confidence interval is from 23.8 to 25.4 percent; and the 95-percent interval is from 23.6 to 25.6 percent.

TABLE IV. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 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]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	59.2	59.2	59.2	59.2	59.2	60.2
200	42.1	42.1	42.1	42.1	42.1	42.6
500	22.5	22.5	22.5	22.5	23.3	26.9
700	17.2	17.2	17.2	17.2	19.7	22.8
1,000	12.7	12.7	12.7	12.7	16.5	19.1
2,500	5.5	5.5	5.5	7.2	10.4	12.0
5,000	2.8	2.8	3.7	5.1	7.4	8.5
10,000	1.4	1.4	2.6	3.6	5.2	6.0
25,000	.6	.8	1.7	2.3	3.3	3.8
50,000	.3	.5	1.2	1.6	2.3	2.7
75,000	.2	.4	1.0	1.3	1.9	2.2
100,000	.14	.4	.8	1.1	1.6	1.9
150,000	.10	.3	.7	.9	1.3	1.6
200,000	.07	.3	.6	.8	1.2	1.3
250,000	.06	.2	.5	.7	1.0	1.2
300,000	.05	.2	.5	.7	1.0	1.1
400,000	.04	.2	.4	.6	.8	1.0
500,000	.03	.2	.4	.5	.7	.9
600,000	.02	.2	.3	.5	.7	.8
700,000	.02	.14	.3	.4	.6	.7
800,000	.02	.13	.3	.4	.6	.7
900,000	.02	.13	.3	.4	.5	.6
1,000,000	.01	.12	.3	.4	.5	.6

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.3.

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. However, 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 1974 and 1977 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1974 and 1977 characteristics.

Illustration of the computation of the standard error of a difference.—Table A-1 of part A of this report shows that in 1977 there were 543,000 owner-occupied units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied units with two bedrooms and owner-occupied units with three bedrooms is 301,800. Table I shows the standard error of 543,000 is approximately 7,490, and the standard error of 241,200 is approximately 5,110. Therefore, the standard error of the estimated difference of 301,800 is about:

$$9{,}070 = \sqrt{(7{,}490)^2 + (5{,}110)^2}$$

Consequently, the 68-percent confidence interval for the 301,800 difference is from 292,730 to 310,870 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of this difference lies within a range

computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 287,288 to 316,312 housing units, and the 95-percent confidence interval is from 283,660 to 319,940. Thus, we can conclude with 95 percent confidence that the number of 1977 owner-occupied 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, read off the confidence interval corresponding to the two points established in step 2.

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 979,000 housing units.

1. Interpolation in table II shows that the standard error of 50 percent on a base of 979,000 is approximately 0.5 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 49.0 and 51.0.

3. From table A-1 of part A, it can be seen by cumulating the frequencies for the first two categories that 397,700 owner-occupied housing units, or 40.6 percent, had one or two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) and that an additional 181,200 owner-occupied housing units, or 18.5 percent, had three persons (i.e., 2.5 to 3.5 persons).

By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{49.0-40.6}{18.5}\right) = 3.0$$

Similarly, the upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{51.0-40.6}{18.5}\right) = 3.1$$

Thus, the 95-percent confidence interval ranges from 3.0 to 3.1 persons. Although it appears that this confidence interval has the sample estimate as the lower limit, it actually is a reflection of the rounding error associated with this median (see the paragraph on rounding errors in the nonsampling errors section of this appendix).

Ft. Worth 1977

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1977 survey App-42
Selection of the 1974 sample App-42
1977 sample reduction App-43
1974-1977 additions to the housing inventory App-43
Coverage improvement sample selection App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1970 Census of Population and Housing App-45
ESTIMATION App-45
1977 housing inventory App-45
1974-1977 lost units App-46
1974 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES App-46
Nonsampling errors App-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-50
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-51

SAMPLE DESIGN

Annual Housing Survey.—The estimates for each of the 20 SMSA's are based on data collected from the 1977 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 20 SMSA's, the data were collected for the 11-month period from April 1977 through February 1978 with one-eleventh of the sample units visited each month.

Each group of SMSA's is interviewed for the AHS on a rotating basis, and this group (A-2 SMSA's) was the first to be revisited. Nineteen of the 20 SMSA's were interviewed for the first time in 1974, while the Madison, Wis., SMSA was interviewed for the first time in 1975.

In each group of SMSA's, the largest SMSA from each of the 4 geographic regions is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All remaining SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the A-2 group (1977-1978) are: Boston, Mass.; Detroit, Mich.; Los Angeles-Long Beach, Calif.; and Washington, D.C.-Md.-Va.

The remaining SMSA's in the A-2 group are: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.

In this SMSA, 4,961 housing units were eligible for interview. Of these sample 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 units eligible for interview, 482 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 1977 survey.—The sample housing units designated for interview in the 1977 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 1974 survey which were not part of the 1977 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 1974 survey which were not part of the 1977 reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1977 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from the list of new construction building permits issued since the 1974 survey. (This sample represented the housing units built in permit-issuing areas, since the 1974 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1974 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1974 survey.)
5. All sample housing units that were selected as part of the 1977 Coverage Improvement Program. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1974 sample.—The sample for the SMSA's which are 100-percent permit-issuing was selected from two sample frames—units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which are not 100-percent permit-issuing in-

cluded a sample selected from a third frame—those units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). The following SMSA's are 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Los Angeles-Long Beach, Calif.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 14 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 four largest SMSA's, the overall sampling rate differed by central city and 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 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 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 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-$5,999. . .		
$6,000-$9,999. . .		
$10,000-$14,999 .		
$15,000 and over .		

Thus, for this SMSA, the occupied housing unit records from this 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 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 are 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 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 units in existence at the time of interview in these selected segments are in sample. Thus, units enumerated in the 1970 census as well as units built since the 1970 census are included.

1977 sample reduction.—In previous years, AHS sample units have been divided among 12 panels which were interviewed each month. Due to budget limitations, it became necessary to drop one of these panels from the 1977 survey. The designated sample size for 1977 was thus reduced by one-twelfth.

1974-1977 additions to the housing inventory.—In the permit-issuing universe, a sample of new construction building permits, issued since the 1974 survey, was

selected to represent housing units built in permit-issuing areas since the 1974 survey. Sampling procedures were identical to those used in selecting the 1970-1974 new construction sample. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1974, to identify any units missed in the 1974 survey or any units added since the 1974 survey.

Coverage improvement sample selection.—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 census.
3. Units missed in the 1970 census.
4. 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.

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 A-2 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 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: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Orlando, Fla.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; and Wichita, Kans.

In the remaining 10 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 and the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 5,212 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, for each 1977-1978 SMSA, 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 1,862 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 no 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 units in multi-unit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multi-unit structures

selected, all units were listed and matched to the 1970 census. Any missed units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply these two procedures in all 20 A-2 SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in three SMSA's. In these SMSA's, a string of 10 structures (instead of 8) that were eligible to be selected for the AHS, were listed. The full application of this technique was employed in 12 SMSA's, with 5 SMSA's not receiving any application of the first procedure. The second procedure was instituted in 13 SMSA's. The following table shows which of these two techniques were applied to the A-2 SMSA's.

The first procedure added an estimated 10,654 units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 183 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 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 1977 housing inventory) and estimates pertaining to characteristics of units removed from the housing inventory since 1974 (i.e., 1974-1977 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1977 housing inventory.—The AHS estimates of characteristics of the 1977 housing inventory employed a 1-stage ratio estimation procedure in 19 SMSA's, with a second-stage ratio estimation procedure being employed in 1 SMSA. Prior to the implementation of the ratio estimation procedure, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of 1 or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample 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

SMSA	Coverage improvement technique employed			
	First		Second	Neither
	Partial	Full		
Albany-Schenectady-Troy, N.Y	X		X	
Anaheim-Santa Ana-Garden Grove, Calif	X			
Boston, Mass		X	X	
Dallas, Tex				X
Detroit, Mich		X	X	
Fort Worth, Tex		X	X	
Los Angeles-Long Beach, Calif				X
Madison, Wis		X	X	
Memphis, Tenn.-Ark		X	X	
Minneapolis-St. Paul, Minn				X
Newark, N.J		X	X	
Orlando, Fla		X	X	
Phoenix, Ariz		X	X	
Pittsburgh, Pa				X
Saginaw, Mich		X	X	
Salt Lake City, Utah				X
Spokane, Wash	X			
Tacoma, Wash		X	X	
Washington, D.C.-Md.-Va		X	X	
Wichita, Kans		X	X	

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:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of 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 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 unit within the corresponding ratio estimation category.

The effect of this ratio estimation procedure was to somewhat reduce the variance due to the variation in sample size of strata used in the sample selection of the permit-issuing universe, thereby reducing the sampling error below what would have been obtained by simply weighting the results of the sample by the inverse of the probability of selection. Ordinarily, this would have been controlled by sampling within the strata during the sample selection process. However, prior to the AHS sample selection within each SMSA, units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. Thus, some variation in sample size was introduced during the AHS sample selection process.

For the Pittsburgh, Pa., SMSA, a second-stage ratio estimation procedure was also employed. This procedure involved the ratio estimation of the AHS weighted sample estimate of the October 1977 housing inventory to an independent estimate of the SMSA's October 1977 housing inventory. This ratio estimate factor equaled the following:

$$\frac{\text{SMSA independent estimate of October 1977 housing inventory}}{\text{AHS sample estimate of the housing inventory}}$$

The numerator of the ratio was derived by using the 1970 census estimate of total housing units for the SMSA in conjunction with the estimate of the change in the housing inventory since the census. The estimate of change was based on administrative records from utility companies. The denominator of that ratio was obtained from the weighted estimate of AHS sample units, using the existing weight after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for each sample unit and the resulting product was used as the final weight for tabulation purposes.

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

The second-stage ratio estimation procedure was not employed in all A-2 SMSA's, since it was felt that coverage improvement procedures adequately corrected known deficiencies in the sample where full coverage improvement procedures were employed. The independent estimates were only employed in SMSA's where full coverage improvements were not used and the estimated relative bias of the independent estimate was thought to be low enough (i.e., 2 percent or less over a 10-year period) to preclude the possibility of doing more harm than good to the survey results. A measure of the relative bias of the independent estimate was available based on a comparison between the 1970 census estimate of the April 1970 housing inventory and the independent estimate of the April 1970 housing inventory which was projected over a 10-year period from the 1960 census using the same procedure that produced the October 1977 independent estimates.

1974-1977 lost units:—The 1974-1977 lost unit estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1974 housing inventory, as was described in the 1974 Current Housing Report, Series H-170, **Housing Characteristics for Selected Metropolitan Areas.** Since the 1974-1977 lost units existed, by definition, in the 1974 housing inventory, there was a 1974 housing inventory weight associated with each 1974-1977 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1974-1977 lost units.

1974 estimation procedure.—This report presents data on the housing characteristics of the 1974 housing inventory from the 1974 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1974.

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 to provide correct information on the part of respondents, mistakes in recording or coding the data, and other errors of collection, response, processing, coverage, and estimation for missing data. Nonsampling errors are not unique to sample surveys since they can, and do, occur in complete censuses as well.

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing, the 1977 AHS-SMSA sample, and the 1974 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 included space errors, definitional errors, and occupancy errors. The "content" errors measured the accuracy of the data collected for surveyed 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.—For the 1977 AHS-SMSA sample, a reinterview program was conducted to obtain a measurement of some of the components of the nonsampling error associated with the AHS estimates. A sample of households was revisited and answers to some of the questions on the AHS questionnaire were obtained again. The original interview and the reinterview were assumed to be two independent readings and thus, were the basis for the measurement of the accuracy of the AHS data collected from interviewed households.

As part of the reinterview, a check was made at each of these households to determine if the following was done during the original interview:

1. The correct unit was visited.
2. The correct number of housing units were interviewed at that address.
3. The correct information on "Year Built" was obtained.
4. The correct information on "Tenure" was obtained.
5. The correct information on "Household Composition" was obtained.
6. The correct information on "Type of Housing Unit" was obtained.
7. The correct information on "Occupancy Status" was obtained.

This check was made for interviewer evaluation and control. That is, tolerance limits were derived to determine which interviewers passed or failed this reinterview check. The results of this study are available and are presented in the following Census Bureau memorandum, "Response Error in the Annual Housing Survey as Measured by Reinterviews—Year IV, 1977-1978 SMSA Sample."

For the A-2 reinterview program, measures of response inconsistency were taken on two subsamples—reconciled and unreconciled. After the question is answered in the reconciled subsample, the interviewer presented the original response and then asked the respondent to decide upon the best answer. Past experience has shown that the level of inconsistency is lower when the interviewer has the response available, and thus, it is believed that there is a downward bias in these estimates. In the unreconciled sample, the interviewer did not have access to previous responses. For the reconciled items, three-fourths showed low levels of response inconsistency while the remaining one-fourth were in the moderate range. For the unreconciled items, approximately one-half showed low levels of response disagreement, with the remaining one-half of the categories in the moderate and high inconsistency range. Although attitudinal items were excluded from the A-2 reinterview program, other reinterview programs have shown that, in general, attitudinal items such as items 102-104 (see facsimile of the 1977 AHS questionnaire, page App-34) have moderate to high levels of inconsistency. Moderate levels of inconsistency indicate that there is some problem with inconsistent reporting, and high levels indicate that improvements are needed in the method used to collect data or that the category concepts themselves are ambiguous.

A second measure of response error, the net difference rate, which is a measure of net error (or bias) in a particular category, was also provided by this reinterview program. None of the categories tested showed any significant evidence of bias.

The results of this study were based on sample data, so there is a sampling error associated with these estimates of nonsampling error. Therefore, the possibility of such errors should be taken into account when considering the results of this study.

A similar study was conducted for the 1974 AHS-SMSA sample. The results of which are presented in the following Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1974."

Coverage errors.—With respect to errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing

units, since, due to the relatively short time span involved, it is possible that construction of these units was not completed at the time the survey was conducted. In which case, they would not have been eligible for interview. In addition to these deficiencies, new construction in special places that do not require building permits, such as military bases, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit. Also, additional coverage errors may exist in SMSA's where the complete Coverage Improvement Program was not applied (see the table on page App-45 of this appendix).

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all units located inside these ED's would be represented in the sample. However, it has been estimated that the 1974 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. Since these ED's were recanvassed for the 1977 survey, the number of missed housing units may be considerably less for 1977.

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 schedules, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by the standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

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

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

The figures presented in the following tables 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 1974 housing inventory can be found in the AHS Series H-170 reports for 1974.

Table I presents the standard errors applicable to estimates of characteristics of the 1977 housing inventory as well as estimates of characteristics of the 1974-1977 lost units. 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 16,620 for the total SMSA, 6,990 for the central city of the SMSA, and 14,960 for the balance of the SMSA.

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

Table II presents the standard errors of estimated percentages of the 1977 housing inventory as well as estimated percentages of the 1974-1977 lost units. 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.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1977 Housing Inventory and for Estimated Number of 1974-1977 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]		
	SMSA	Central city	Not in central city
0	70	60	70
100.	80	80	90
200.	120	110	120
500.	180	170	190
700.	220	210	230
1,000 . . .	260	250	270
2,500 . . .	410	390	430
5,000 . . .	580	550	600
10,000 . .	820	770	860
25,000 . .	1,290	1,180	1,370
50,000 . .	1,820	1,570	1,980
75,000 . .	2,230	1,800	2,470
100,000. .	2,560	1,930	2,910
150,000. .	3,130	1,950	3,680
200,000. .	3,600	1,630	4,390
250,000. .	4,010	—	—
300,000. .	4,370	—	—
400,000. .	5,010	—	—

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2 for the total SMSA, 1.3 for the central city, and 1.1 for the balance (not in central city) estimates.

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 1977 there were 206,400 owner-occupied housing units in this SMSA. Interpolation in table I of this appendix shows that the standard error of an estimate of this size is approximately 3,650. The following interpolation procedure was used:

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

Size of estimate	Standard error
200,000.	3,600
206,400.	x
250,000.	4,010

The entry for x is determined as follows by vertically interpolating between 3,600 and 4,010:

$$206{,}400-200{,}000 = 6{,}400$$
$$250{,}000-200{,}000 = 50{,}000$$

$$3{,}600+\frac{6{,}400}{50{,}000}(4{,}010-3{,}600) = 3{,}650$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 202,750 to 210,050 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1977 owner-occupied housing units lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 200,560 to 212,240 housing units with 90 percent confidence; and that the average estimate lies within the interval from 199,100 to 213,700 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 206,400 owner-occupied housing units, 55,500, or 26.9 percent, had two bedrooms. Interpolation in table II of the appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 26.9 percent is approximately 0.8 percentage points. The following interpolation procedure was used:

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

Base of percentage	Estimated percentage		
	25	26.9	50
200,000. . . .	0.8	a	1.0
206,400. . . .		p	
250,000. . . .	0.7	b	0.9

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

$$26.9-25.0 = 1.9$$
$$50.0-25.0 = 25.0$$

$$0.8+\frac{1.9}{25.0}(1.0-0.8) = 0.81$$

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

$$26.9-25.0 = 1.9$$
$$50.0-25.0 = 25.0$$

$$0.7+\frac{1.9}{25.0}(0.9-0.7) = 0.71$$

3. The entry for "p" was then determined by vertical interpolation between 0.81 and 0.71.

$$206{,}400-200{,}000 = 6{,}400$$
$$250{,}000-200{,}000 = 50{,}000$$

$$0.81+\frac{6{,}400}{50{,}000}(0.71-0.81) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 26.1 to 27.8 percent; the 90-percent

confidence interval is from 25.6 to 28.2 percent; and the 95-percent confidence interval is from 25.3 to 28.5 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. However, 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 1974 and 1977 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1974 and 1977 characteristics.

Illustration of the computation of the standard error of a difference.—Table A-1 of part A of this report shows that in 1977 there were 116,600 owner-occupied units with three bedrooms in this SMSA, Thus, the apparent difference, as shown by these data, between owner-occupied units with two bedrooms and owner-occupied units with three bedrooms is 61,100. Table I shows the standard error of 116,600 is approximately 2,750 and the standard error of 55,000 is approximately 1,910. Therefore, the standard error of the estimated difference of 61,100 is about:

$$3{,}350 = \sqrt{(2{,}750)^2 + (1{,}910)^2}$$

Consequently, the 68-percent confidence interval for the 61,100 difference is from 57,750 to 64,450 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 55,740 to 66,460 housing units, and the 95-percent confidence interval is from 54,400 to 67,800. Thus, we can conclude with 95 percent confidence that the number of 1977 owner-occupied 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.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 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)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	42.1	42.1	42.1	42.1	42.1	42.6
200	26.6	26.6	26.6	26.6	26.6	30.1
500	12.7	12.7	12.7	12.7	16.5	19.1
700	9.4	9.4	9.4	9.7	13.9	16.1
1,000	6.8	6.8	6.8	8.1	11.7	13.5
2,500	2.8	2.8	3.7	5.1	7.4	8.5
5,000	1.4	1.4	2.6	3.6	5.2	6.0
10,000	0.7	0.8	1.9	2.6	3.7	4.3
25,000	0.3	0.5	1.2	1.6	2.3	2.7
50,000	0.15	0.4	0.8	1.1	1.7	1.9
75,000	0.10	0.3	0.7	0.9	1.3	1.6
100,000	0.07	0.3	0.6	0.8	1.2	1.3
150,000	0.05	0.2	0.5	0.7	1.0	1.1
200,000	0.04	0.2	0.4	0.6	0.8	1.0
250,000	0.03	0.2	0.4	0.5	0.7	0.9
300,000	0.02	0.15	0.3	0.5	0.7	0.8
400,000	0.02	0.13	0.3	0.4	0.6	0.7

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2 for the total SMSA, 1.3 for the central city, and 1.1 for the balance (not in central city) estimates.

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, read off the confidence interval corresponding to the two points established in step 2.

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 206,400 housing units.

1. Interpolation in table II shows that the standard error of 50 percent on a base of 206,400 is approximately 1.0 percentage points.

2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.0 and 52.0.

3. From table A-1 of part A, it can be seen by cumulating the frequencies for the first two categories that 90,400 owner-occupied housing units, or 43.8 percent, had one or two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) and that an additional 40,300 owner-occupied housing units, or 19.5 percent, had three persons (i.e., 2.5 to 3.5 persons).

By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{48.0-43.8}{19.5}\right) = 2.7$$

Similarly, the upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{52.0-43.8}{19.5}\right) = 2.9$$

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

Los Angeles-Long Beach 1977

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1977 survey App-42
Selection of the 1974 sample App-42
1977 sample reduction App-43
1974-1977 additions to the housing inventory App-44
Coverage improvement sample selection App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1970 Census of Population and Housing App-45
ESTIMATION App-45
1977 housing inventory App-45
1974-1977 lost units App-46
1974 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES App-46
Nonsampling errors App-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-50
Illustration of the computation of the standard error of a difference App-50
Medians App-51
Illustration of the computation of the 95-percent confidence interval of a median ... App-51

SAMPLE DESIGN

Annual Housing Survey.—The estimates for each of the 20 SMSA's are based on data collected from the 1977 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 20 SMSA's, the data were collected for the 11-month period from April 1977 through February 1978 with one-eleventh of the sample units visited each month.

Each group of SMSA's is interviewed for the AHS on a rotating basis, and this group (A-2 SMSA's) was the first to be revisited. Nineteen of the 20 SMSA's were interviewed for the first time in 1974, while the Madison, Wis., SMSA was interviewed for the first time in 1975.

In each group of SMSA's, the largest SMSA from each of the 4 geographic regions is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All remaining SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the A-2 group (1977-1978) are: Boston, Mass.; Detroit, Mich.; Los Angeles-Long Beach, Calif.; and Washington, D.C.-Md.-Va.

The remaining SMSA's in the A-2 group are: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.

In this SMSA, 13,831 housing units were eligible for interview. Of these sample units, 776 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 units eligible for interview, 707 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 1977 survey.—The sample housing units designated for interview in the 1977 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 1974 survey which were not part of the 1977 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 1974 survey which were not part of the 1977 reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1977 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from the list of new construction building permits issued since the 1974 survey. (This sample represented the housing units built in permit-issuing areas, since the 1974 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1974 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1974 survey.)
5. All sample housing units that were selected as part of the 1977 Coverage Improvement Program. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1974 sample.—The sample for the SMSA's which are 100-percent permit-issuing was selected from two sample frames—units enumerated in the 1970 Census of Population and Housing in areas under the jurisdic-

tion of permit-issuing offices (the permit-issuing universe) and units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which are not 100-percent permit-issuing included a sample selected from a third frame—those units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). The following SMSA's are 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Los Angeles-Long Beach, Calif.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 14 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 four largest SMSA's, the overall sampling rate differed by central city and 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 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 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 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-$5,999. . .		
$6,000-$9,999. . .		
$10,000-$14,999 .		
$15,000 and over .		

Thus, for this SMSA, the occupied housing unit records from this 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 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 are 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 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 units in existence at the time of interview in these selected segments are in sample. Thus, units enumerated in the 1970 census as well as units built since the 1970 census are included.

1977 sample reduction.—In previous years, AHS sample units have been divided among 12 panels which were interviewed each month. Due to budget limitations, it became necessary to drop one of these panels from the 1977 survey. The designated sample size for 1977 was thus reduced by one-twelfth.

1974-1977 additions to the housing inventory.—In the permit-issuing universe, a sample of new construction building permits, issued since the 1974 survey, was selected to represent housing units built in permit-issuing areas since the 1974 survey. Sampling procedures were identical to those used in selecting the 1970-1974 new construction sample. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1974, to identify any units missed in the 1974 survey or any units added since the 1974 survey.

Coverage improvement sample selection.—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 census.
3. Units missed in the 1970 census.
4. 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.

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 A-2 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 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: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Orlando, Fla.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; and Wichita, Kans.

In the remaining 10 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 and the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 14,014 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, for each 1977-1978 SMSA, 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 10,747 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 no 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 units in multi-unit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multi-unit structures selected, all units were listed and matched to the 1970 census. Any missed units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply these two procedures in all 20 A-2 SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in three SMSA's. In these SMSA's, a string of 10 structures (instead of 8) that were eligible to be selected for the AHS, were listed. The full application of this technique was employed in 12 SMSA's, with 5 SMSA's not receiving any application of the first procedure. The second procedure was instituted in 13 SMSA's. The following table shows which of these two techniques were applied to the A-2 SMSA's:

SMSA	Coverage improvement technique employed			
	First		Second	Neither
	Partial	Full		
Albany-Schenectady-Troy, N.Y	X		X	
Anaheim-Santa Ana-Garden Grove, Calif	X			
Boston, Mass		X	X	
Dallas, Tex				X
Detroit, Mich		X	X	
Fort Worth, Tex		X	X	
Los Angeles-Long Beach, Calif				X
Madison, Wis		X	X	
Memphis, Tenn.-Ark		X	X	
Minneapolis-St. Paul, Minn				X
Newark, N.J		X	X	
Orlando, Fla		X	X	
Phoenix, Ariz		X	X	
Pittsburgh, Pa				X
Saginaw, Mich		X	X	
Salt Lake City, Utah				X
Spokane, Wash	X			
Tacoma, Wash		X	X	
Washington, D.C.-Md.-Va		X	X	
Wichita, Kans		X	X	

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 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 1977 housing inventory) and estimates pertaining to characteristics of units removed from the housing inventory since 1974 (i.e., 1974-1977 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1977 housing inventory.—The AHS estimates of characteristics of the 1977 housing inventory employed a 1-stage ratio estimation procedure in 19 SMSA's, with a second-stage ratio estimation procedure being employed in 1 SMSA. Prior to the implementation of the ratio estimation procedure, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of 1 or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample 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:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of 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 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 unit within the corresponding ratio estimation category.

The effect of this ratio estimation procedure was to somewhat reduce the variance due to the variation in sample size of strata used in the sample selection of the permit-issuing universe, thereby reducing the sampling error below what would have been obtained by simply weighting the results of the sample by the inverse of the probability of selection. Ordinarily, this would have been controlled by sampling within the strata during the sample selection process. However, prior to the AHS sample selection within each SMSA, units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. Thus, some variation in sample size was introduced during the AHS sample selection process.

For the Pittsburgh, Pa., SMSA, a second-stage ratio estimation procedure was also employed. This procedure involved the ratio estimation of the AHS weighted sample estimate of the October 1977 housing inventory to an independent estimate of the SMSA's October 1977 housing inventory. This ratio estimate factor equaled the following:

$$\frac{\text{SMSA independent estimate of October 1977 housing inventory}}{\text{AHS sample estimate of the housing inventory}}$$

The numerator of the ratio was derived by using the 1970 census estimate of total housing units for the SMSA in conjunction with the estimate of the change in the housing inventory since the census. The estimate of change was based on administrative records from utility companies. The denominator of that ratio was obtained from the weighted estimate of AHS sample units, using the existing weight after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for each sample unit and the resulting product was used as the final weight for tabulation purposes.

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

The second-stage ratio estimation procedure was not employed in all A-2 SMSA's, since it was felt that coverage improvement procedures adequately corrected known deficiencies in the sample where full coverage improvement procedures were employed. The independent estimates were only employed in SMSA's where full coverage improvements were not used and the estimated relative bias of the independent estimate was thought to be low enough (i.e., 2 percent or less over a 10-year period) to preclude the possibility of doing more harm than good to the survey results. A measure of the relative bias of the independent estimate was available based on a comparison between the 1970 census estimate of the April 1970 housing inventory and the independent estimate of the April 1970 housing inventory which was projected over a 10-year period from the 1960 census using the same procedure that produced the October 1977 independent estimates.

1974-1977 lost units.—The 1974-1977 lost unit estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1974 housing inventory, as was described in the 1974 Current Housing Report, Series H-170, **Housing Characteristics for Selected Metropolitan Areas.** Since the 1974-1977 lost units existed, by definition, in the 1974 housing inventory, there was a 1974 housing inventory weight associated with each 1974-1977 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1974-1977 lost units.

1974 estimation procedure.—This report presents data on the housing characteristics of the 1974 housing inventory from the 1974 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1974.

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 to provide correct information on the part of respondents, mistakes in recording or coding the data, and other errors of collection, response, processing, coverage, and estimation for missing data. Nonsampling errors are not unique to sample surveys since they can, and do, occur in complete censuses as well.

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing, the 1977 AHS-SMSA sample, and the 1974 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 included space errors, definitional errors, and occupancy errors. The "content" errors measured the accuracy of the data collected for surveyed 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.—For the 1977 AHS-SMSA sample, a reinterview program was conducted to obtain a measurement of some of the components of the nonsampling error associated with the AHS estimates. A sample of households was revisited and answers to some of the questions on the AHS questionnaire were obtained again. The original interview and the reinterview were assumed to be two independent readings and, thus, were the basis for the measurement of the accuracy of the AHS data collected from interviewed households.

As part of the reinterview, a check was made at each of these households to determine if the following was done during the original interview:

1. The correct unit was visited.
2. The correct number of housing units were interviewed at that address.
3. The correct information on "Year Built" was obtained.
4. The correct information on "Tenure" was obtained.
5. The correct information on "Household Composition" was obtained.
6. The correct information on "Type of Housing Unit" was obtained.
7. The correct information on "Occupancy Status" was obtained.

This check was made for interviewer evaluation and control. That is, tolerance limits were derived to determine which interviewers passed or failed this reinterview check. The results of this study are available and are presented in the following Census Bureau memorandum, "Response Error in the Annual Housing Survey as Measured by Reinterviews—Year IV, 1977-1978 SMSA Sample."

For the A-2 reinterview program, measures of response inconsistency were taken on two subsamples—reconciled and unreconciled. After the question is answered in the reconciled subsample, the interviewer presented the original response and then asked the respondent to decide upon the best answer. Past experience has shown that the level of inconsistency is lower when the interviewer has the response available, and thus, it is believed that there is a downward bias in these estimates. In the unreconciled sample, the interviewer did not have access to previous responses. For the reconciled items, three-fourths showed low levels of response inconsistency while the remaining one-fourth were in the moderate range. For the unreconciled items, approximately one-half showed low levels of response disagreement, with the remaining one-half of the categories in the moderate and high inconsistency range. Although attitudinal items were excluded from the A-2 reinterview program, other reinterview programs have shown that, in general, attitudinal items such as items 102-104 (see facsimile of the 1977 AHS questionnaire, page App-34) have moderate to high levels of inconsistency. Moderate levels of inconsistency indicate that there is some problem with inconsistent reporting, and high levels indicate that improvements are needed in the method used to collect data or that the category concepts themselves are ambiguous.

A second measure of response error, the net difference rate, which is a measure of net error (or bias) in a particular category, was also provided by this reinterview program. None of the categories tested showed any significant evidence of bias.

The results of this study were based on sample data, so there is a sampling error associated with these estimates of nonsampling error. Therefore, the possibility of such errors should be taken into account when considering the results of this study.

A similar study was conducted for the 1974 AHS-SMSA sample. The results of which are presented in the following Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1974."

Coverage errors.—With respect to errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units, since, due to the relatively short time span involved, it is possible that construction of these units was not completed at the time the survey was con-

ducted. In which case, they would not have been eligible for interview. In addition to these deficiencies, new construction in special places that do not require building permits, such as military bases, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit. Also, additional coverage errors may exist in SMSA's where the complete Coverage Improvement Program was not applied (see the table on page App-45 of this appendix).

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all units located inside these ED's would be represented in the sample. However, it has been estimated that the 1974 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. Since these ED's were recanvassed for the 1977 survey, the number of missed housing units may be considerably less for 1977.

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 schedules, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by the standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.

2. Approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the average result of all possible samples.

3. Approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimate would include the average result of all possible samples.

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

The figures presented in the following tables 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 1974 housing inventory can be found in the AHS Series H-170 reports for 1974.

Table I presents the standard errors applicable to estimates of characteristics of the 1977 housing inventory as well as estimates of characteristics of the 1974-1977 lost units. 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 50,533 for the total SMSA, 33,468 for the central cities of the SMSA, and 37,851 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 presents the standard errors of estimated percentages of the 1977 housing inventory as well as estimated percentages of the 1974-1977 lost units. 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

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1977 Housing Inventory and for Estimated Number of 1974-1977 Lost Units for the Los Angeles-Long Beach, 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]		
	SMSA	In central cities	Not in central cities
0	210	190	220
100	210	190	220
200	210	200	220
500	330	310	330
700	380	370	390
1,000	460	440	470
2,500	730	690	740
5,000	1,030	980	1,050
10,000 . . .	1,450	1,380	1,480
25,000 . . .	2,290	2,170	2,330
50,000 . . .	3,230	3,050	3,280
75,000 . . .	3,940	3,710	3,990
100,000 . .	4,540	4,250	4,580
150,000 . .	5,520	5,130	5,530
200,000 . .	6,330	5,830	6,300
250,000 . .	7,020	6,410	6,940
300,000 . .	7,640	6,910	7,490
400,000 . .	8,690	7,710	8,380
500,000 . .	9,570	8,300	9,060
600,000 . .	10,320	8,740	9,570
700,000 . .	10,970	9,030	9,950
800,000 . .	11,530	9,200	10,200
900,000 . .	12,020	9,260	10,330
1,000,000 .	12,440	9,190	10,360
1,100,000 .	12,810	9,020	10,270
1,200,000 .	13,120	8,710	10,070
1,300,000 .	13,390	8,270	9,750
1,400,000 .	13,600	—	9,300
1,500,000 .	13,770	—	8,690
1,750,000 .	14,020	—	—
2,000,000 .	14,010	—	—
2,250,000 .	13,750	—	—
3,000,000 .	11,150	—	—

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.3 for the total SMSA, 1.4 for the central cities, and 1.2 for the balance (not in central cities) estimates.

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 1977 there were 1,266,600 owner-occupied housing units in this SMSA. Interpolation in table I of this appendix shows that the standard error of an estimate of this size is approximately 13,300. The following interpolation procedure was used:

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

Size of estimate	Standard error
1,200,000	13,120
1,266,600	x
1,300,000	13,390

The entry for x is determined as follows by vertically interpolating between 13,120 and 13,390:

1,266,600−1,200,000 = 66,600
1,300,000−1,200,000 = 100,000

$$13{,}120+\frac{66{,}600}{100{,}000}(13{,}390-13{,}120)$$
$$=13{,}300$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 1,253,300 to 1,279,900 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1977 owner-occupied housing units lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 1,245,320 to 1,287,880 housing units with 90 percent confidence; and that the average estimate lies within the interval from 1,240,000 to 1,293,200 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 1,266,600 owner-occupied housing units, 440,700, or 34.8 percent, had two bedrooms. Interpolation in table II of the appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 34.8 percent is approximately 0.6 percentage points. The following interpolation procedure was used:

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

Base of percentage	Estimated percentage		
	25	34.8	50
1,200,000 . .	0.6	a	0.7
1,266,600 . .		p	
1,300,000 . .	0.6	b	0.7

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

34.8−25.0 = 9.8
50.0−25.0 = 25.0

$$0.6+\frac{9.8}{25.0}(0.7-0.6)=0.6$$

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

$$34.8-25.0 = 9.8$$
$$50.0-25.0 = 25.0$$

$$0.6 + \frac{9.8}{25.0}(0.7-0.6) = 0.6$$

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

$$1{,}266{,}600-1{,}200{,}000 = 66{,}600$$
$$13{,}00{,}000-1{,}200{,}000 = 100{,}000$$

$$0.6 + \frac{66{,}600}{100{,}000}(0.6-0.6) = 0.6$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 34.2 to 35.4 percent; the 90-percent confidence interval is from 33.8 to 35.8 percent; and the 95-percent confidence interval is from 33.6 to 36.0 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. However, 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 1974 and 1977 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1974 and 1977 characteristics.

Illustration of the computation of the standard error of a difference.—Table A-1 of part A of this report shows that in 1977 there were 565,300 owner-occupied units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied units with two bedrooms and owner-occupied units with three bedrooms is 124,600. Table I shows the standard error of 440,700 is approximately 9,050, and the standard error of 565,300 is approximately 10,060. Therefore, the standard error of the estimated difference of 124,600 is about:

$$13{,}530 = \sqrt{(10{,}060)^2 + (9{,}050)^2}$$

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 Lost Units for the Los Angeles-Long Beach, 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)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	68.8	68.8	68.8	68.8	68.8	74.3
200	52.5	52.5	52.5	52.5	52.5	52.5
500	30.6	30.6	30.6	30.6	30.6	33.2
700	24.0	24.0	24.0	24.0	24.3	28.1
1,000	18.1	18.1	18.1	18.1	20.4	23.5
2,500	8.1	8.1	8.1	8.9	12.9	14.9
5,000	4.2	4.2	4.6	6.3	9.1	10.5
10,000	2.2	2.2	3.2	4.5	6.4	7.4
25,000	0.9	0.9	2.0	2.8	4.1	4.7
50,000	0.4	0.7	1.4	2.0	2.9	3.3
75,000	0.3	0.5	1.2	1.6	2.4	2.7
100,000	0.2	0.5	1.0	1.4	2.0	2.4
150,000	0.15	0.4	0.8	1.2	1.7	1.9
200,000	0.11	0.3	0.7	1.0	1.4	1.7
250,000	0.09	0.3	0.6	0.9	1.3	1.5
300,000	0.07	0.3	0.6	0.8	1.2	1.4
400,000	0.06	0.2	0.5	0.7	1.0	1.2
500,000	0.04	0.2	0.5	0.6	0.9	1.1
600,000	0.04	0.2	0.4	0.6	0.8	1.0
700,000	0.03	0.2	0.4	0.5	0.8	0.9
800,000	0.03	0.17	0.4	0.5	0.7	0.8
900,000	0.02	0.16	0.3	0.5	0.7	0.8
1,000,000	0.02	0.15	0.3	0.4	0.6	0.7
1,100,000	0.02	0.14	0.3	0.4	0.6	0.7
1,200,000	0.02	0.13	0.3	0.4	0.6	0.7
1,300,000	0.02	0.13	0.3	0.4	0.6	0.7
1,400,000	0.02	0.12	0.3	0.4	0.5	0.6
1,500,000	0.01	0.12	0.3	0.4	0.5	0.6
1,750,000	0.01	0.11	0.2	0.3	0.5	0.6
2,000,000	0.01	0.10	0.2	0.3	0.5	0.5
2,250,000	0.01	0.10	0.2	0.3	0.4	0.5
2,500,000	0.01	0.09	0.2	0.3	0.4	0.5
3,000,000	0.01	0.09	0.2	0.3	0.4	0.4

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.3 for the total SMSA, 1.4 for the central cities, and 1.2 for the balance (not in central cities) estimates.

Consequently, the 68-percent confidence interval for the 124,600 difference is from 111,070 to 138,130 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of this difference lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 102,952 to 146,248 housing units, and the 95-percent confidence interval is from 97,540 to 151,660. Thus, we can conclude with 95 percent confidence that the number of 1977 owner-occupied 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, read off the confidence interval corresponding to the two points established in step 2.

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 1,266,600 housing units.

1. Interpolation in table II shows that the standard error of 50 percent on a base of 1,266,600 is approximately 0.7 percentage points.

2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.6 and 51.4.

3. From table A-1 of part A, it can be seen by cumulating the frequencies for the first two categories that 570,100 owner-occupied housing units, or 45.0 percent, had one or two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) and that an additional 248,300 owner-occupied housing units, or 19.6 percent, had three persons (i.e., 2.5 to 3.5 persons).

By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{48.6-45.0}{19.6} = 2.7$$

Similarly, the upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{51.4-45.0}{19.6} = 2.8$$

Thus, the 95-percent confidence interval ranges from 2.7 to 2.8 persons. Although it appears that this confidence interval has the sample estimate as the upper limit, it actually is a reflection of the rounding error associated with this median (see the paragraph on rounding errors in the nonsampling errors section of this appendix).

Madison 1977

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1977 survey App-42
Selection of the original sample ... App-42
1977 sample reduction App-43
1975-1977 additions to the housing inventory App-43
Coverage improvement sample selection App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1970 Census of Population and Housing App-45
ESTIMATION App-45
1977 housing inventory App-45
1975-1977 lost units App-46
1975 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES App-46
Nonsampling errors App-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-50
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

SAMPLE DESIGN

Annual Housing Survey.—The estimates for each of the 20 SMSA's are based on data collected from the 1977 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 20 SMSA's, the data were collected for the 11-month period from April 1977 through February 1978 with one-eleventh of the sample units visited each month.

Each group of SMSA's is interviewed for the AHS on a rotating basis, and this group (A-2 SMSA's) was the first to be revisited. Nineteen of the 20 SMSA's were interviewed for the first time in 1974, while the Madison, Wis., SMSA was interviewed for the first time in 1975.

In each group of SMSA's, the largest SMSA from each of the 4 geographic regions is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All remaining SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the A-2 group (1977-1978) are: Boston, Mass.; Detroit, Mich.; Los Angeles-Long Beach, Calif.; and Washington, D.C.-Md.-Va.

The remaining SMSA's in the A-2 group are: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.

In this SMSA, 4,839 housing units were eligible for interview. Of these sample units, 154 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 units eligible for interview, 419 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 1977 survey.—The sample housing units designated for interview in the 1977 Madison, Wis. survey consisted of the following categories which are described in detail in the succeeding sections.

1. All sample housing units that were interviewed in the 1975 survey which were not part of the 1977 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 1975 survey which were not part of the 1977 reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1977 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from the list of new construction building permits issued since the 1975 survey. (This sample represented the housing units built in permit-issuing areas, since the 1975 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1975 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1975 survey.)
5. All sample housing units that were selected as part of the 1977 Coverage Improvement Program. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the original sample.—The sample for the SMSA's which are 100-percent permit-issuing was selected from two sample frames—units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which are

not 100-percent permit-issuing included a sample selected from a third frame—those units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). The following SMSA's are 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Los Angeles-Long Beach, Calif.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 14 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 four largest SMSA's, the overall sampling rate differed by central city and 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 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 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 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-$5,999. . .		
$6,000-$9,999. . .		
$10,000-$14,999 .		
$15,000 and over .		

Thus, for this SMSA, the occupied housing unit records from this 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 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 are 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 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 units in existence at the time of interview in these selected segments are in sample. Thus, units enumerated in the 1970 census as well as units built since the 1970 census are included.

1977 sample reduction.—In previous years, AHS sample units have been divided among 12 panels which were interviewed each month. Due to budget limitations, it became necessary to drop one of these panels from the 1977 survey. The designated sample size for 1977 was thus reduced by one-twelfth.

1975-1977 additions to the housing inventory.—In the permit-issuing universe, a sample of new construction building permits, issued since the 1975 survey, was

selected to represent housing units built in permit-issuing areas since the 1975 survey. Sampling procedures were identical to those used in selecting the 1970-1975 new construction sample. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1975, to identify any units missed in the 1975 survey or any units added since the 1975 survey.

Coverage improvement sample selection.—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 census.
3. Units missed in the 1970 census.
4. 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.

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 A-2 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 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: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Orlando, Fla.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; and Wichita, Kans.

In the remaining 10 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 and the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 974 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, for each 1977-1978 SMSA, 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 169 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 no 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 units in multi-unit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multi-unit structures

selected, all units were listed and matched to the 1970 census. Any missed units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply these two procedures in all 20 A-2 SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in three SMSA's. In these SMSA's, a string of 10 structures (instead of 8) that were eligible to be selected for the AHS, were listed. The full application of this technique was employed in 12 SMSA's, with 5 SMSA's not receiving any application of the first procedure. The second procedure was instituted in 13 SMSA's. The following table shows which of these two techniques were applied to the A-2 SMSA's.

The first procedure added an estimated 1,602 units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 44 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 AHS sample produced two types of estimates for this SMSA: Estimates pertaining to characteristics of the housing inventory at the time of the interview (i.e., the 1977 housing inventory) and estimates pertaining to characteristics of units removed from the housing inventory since 1975 (i.e., 1975-1977 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1977 housing inventory.—The AHS estimates of characteristics of the 1977 housing inventory employed a 1-stage ratio estimation procedure in 19 SMSA's, with a second-stage ratio estimation procedure being employed in 1 SMSA. Prior to the implementation of the ratio estimation procedure, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of 1 or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample 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 estima-

SMSA	Coverage improvement technique employed			
	First		Second	Neither
	Partial	Full		
Albany-Schenectady-Troy, N.Y	X		X	
Anaheim-Santa Ana-Garden Grove, Calif	X			
Boston, Mass		X	X	
Dallas, Tex				X
Detroit, Mich		X	X	
Fort Worth, Tex		X	X	
Los Angeles-Long Beach, Calif				X
Madison, Wis		X	X	
Memphis, Tenn.-Ark		X	X	
Minneapolis-St. Paul, Minn				X
Newark, N.J		X	X	
Orlando, Fla		X	X	
Phoenix, Ariz		X	X	
Pittsburgh, Pa				X
Saginaw, Mich		X	X	
Salt Lake City, Utah				X
Spokane, Wash	X			
Tacoma, Wash		X	X	
Washington, D.C.-Md.-Va		X	X	
Wichita, Kans		X	X	

tion factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of 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 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 unit within the corresponding ratio estimation category.

The effect of this ratio estimation procedure was to somewhat reduce the variance due to the variation in sample size of strata used in the sample selection of the permit-issuing universe, thereby reducing the sampling error below what would have been obtained by simply weighting the results of the sample by the inverse of the probability of selection. Ordinarily, this would have been controlled by sampling within the strata during the sample selection process. However, prior to the AHS sample selection within each SMSA, units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. Thus, some variation in sample size was introduced during the AHS sample selection process.

For the Pittsburgh, Pa., SMSA, a second-stage ratio estimation procedure was also employed. This procedure involved the ratio estimation of the AHS weighted sample estimate of the October 1977 housing inventory to an independent estimate of the SMSA's October 1977 housing inventory. This ratio estimate factor equaled the following:

$$\frac{\text{SMSA independent estimate of October 1977 housing inventory}}{\text{AHS sample estimate of the housing inventory}}$$

The numerator of the ratio was derived by using the 1970 census estimate of total housing units for the SMSA in conjunction with the estimate of the change in the housing inventory since the census. The estimate of change was based on administrative records from utility companies. The denominator of that ratio was obtained from the weighted estimate of AHS sample units, using the existing weight after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for each sample unit and the resulting product was used as the final weight for tabulation purposes.

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

The second-stage ratio estimation procedure was not employed in all A-2 SMSA's, since it was felt that coverage improvement procedures adequately corrected known deficiencies in the sample where full coverage improvement procedures were employed. The independent estimates were only employed in SMSA's where full coverage improvements were not used and the estimated relative bias of the independent estimate was thought to be low enough (i.e., 2 percent or less over a 10-year period) to preclude the possibility of doing more harm than good to the survey results. A measure of the relative bias of the independent estimate was available based on a comparison between the 1970 census estimate of the April 1970 housing inventory and the independent estimate of the April 1970 housing inventory which was projected over a 10-year period from the 1960 census using the same procedure that produced the October 1977 independent estimates.

1975-1977 lost units.—The 1975-1977 Madison, Wis. lost unit estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1975 housing inventory, as was described in the 1975 Current Housing Report, Series H-170, **Housing Characteristics for Selected Metropolitan Areas.** Since the 1975-1977 lost units existed, by definition, in the 1975 housing inventory, there was a 1975 housing inventory weight associated with each 1975-1977 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1975-1977 lost units.

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

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

RELIABILITY OF THE ESTIMATES

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

Volume I, **Housing Characteristics for States, Cities, and Counties**, Part 1.

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

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing, the 1977 AHS-SMSA sample, and the 1975 AHS-SMSA sample.

1970 census.—A number of studies were conducted to measure two types of general errors associated with 1970 census estimates—"coverage" and "content" errors. The "coverage" errors determined how completely housing units were counted in the census and included space errors, definitional errors, and occupancy errors. The "content" errors measured the accuracy of the data collected for surveyed 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.—For the 1977 AHS-SMSA sample, a reinterview program was conducted to obtain a measurement of some of the components of the nonsampling error associated with the AHS estimates. A sample of households was revisited and answers to some of the questions on the AHS questionnaire were obtained again. The original interview and the reinterview were assumed to be two independent readings and, thus, were the basis for the measurement of the accuracy of the AHS data collected from interviewed households.

As part of the reinterview, a check was made at each of these households to determine if the following was done during the original interview:

1. The correct unit was visited.
2. The correct number of housing units were interviewed at that address.
3. The correct information on "Year Built" was obtained.
4. The correct information on "Tenure" was obtained.
5. The correct information on "Household Composition" was obtained.
6. The correct information on "Type of Housing Unit" was obtained.
7. The correct information on "Occupancy Status" was obtained.

This check was made for interviewer evaluation and control. That is, tolerance limits were derived to determine which interviewers passed or failed this reinterview check. The results of this study are available and are presented in the following Census Bureau memorandum, "Response Error in the Annual Housing Survey as Measured by Reinterviews—Year IV, 1977-1978 SMSA Sample."

For the A-2 reinterview program, measures of response inconsistency were taken on two subsamples—reconciled and unreconciled. After the question is answered in the reconciled subsample, the interviewer presented the original response and then asked the respondent to decide upon the best answer. Past experience has shown that the level of inconsistency is lower when the interviewer has the response available, and thus, it is believed that there is a downward bias in these estimates. In the unreconciled sample, the interviewer did not have access to previous responses. For the reconciled items, three-fourths showed low levels of response inconsistency while the remaining one-fourth were in the moderate range. For the unreconciled items, approximately one-half showed low levels of response disagreement, with the remaining one-half of the categories in the moderate and high inconsistency range. Although attitudinal items were excluded from the A-2 reinterview program, other reinterview programs have shown that, in general, attitudinal items such as items 102-104 (see facsimile of the 1977 AHS questionnaire, page App-34) have moderate to high levels of inconsistency. Moderate levels of inconsistency indicate that there is some problem with inconsistent reporting, and high levels indicate that improvements are needed in the method used to collect data or that the category concepts themselves are ambiguous.

A second measure of response error, the net difference rate, which is a measure of net error (or bias) in a particular category, was also provided by this reinterview program. None of the categories tested showed any significant evidence of bias.

The results of this study were based on sample data, so there is sampling error associated with these estimates of nonsampling error. Therefore, the possibility of such errors should be taken into account when considering the results of this study.

A similar study was conducted for the 1975 AHS-SMSA sample. The results of which are presented in the following Census Bureau memorandum, "Response Error in the Annual Housing Survey Data—Year II SMSA Sample."

Coverage errors.—With respect to errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units, since, due to the relatively short time span involved, it is possible that construction of these units was not com-

pleted at the time the survey was conducted. In which case, they would not have been eligible for interview. In addition to these deficiencies, new construction in special places that do not require building permits, such as military bases, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit. Also, additional coverage errors may exist in SMSA's where the complete Coverage Improvement Program was not applied (see the table on page App-45 of this appendix).

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all units located inside these ED's would be represented in the sample. However, it has been estimated that the 1975 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. Since these ED's were recanvassed for the 1977 survey, the number of missed housing units may be considerably less for 1977.

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 schedules, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by the standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

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

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

The figures presented in the following tables are approximations to the standard errors of various estimates shown in this report for this SMSA. In order to derive standard errors that would be applicable to a wide variety of items and also could be prepared at a moderate cost, a number of approximations were required. As a result, the tables of standard errors provide an indication of the order of magnitude of the standard errors rather than precise standard errors for any specific item. Standard errors applicable to estimates of characteristics of the 1975 housing inventory can be found in the AHS Series H-170 reports for 1975.

Table I presents the standard errors applicable to estimates of characteristics of the 1977 housing inventory as well as estimates of characteristics of the 1975-1977 lost units. 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 4,230 for the total SMSA, 2,440 for the central city of the SMSA, and 3,380 for the balance of the SMSA.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1977 Housing Inventory and for Estimated Number of 1975-1977 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 . .	350
100.	50	10,000. .	490
200.	70	25,000. .	760
500.	110	50,000. .	1,010
700.	130	75,000. .	1,170
1,000 . . .	160	100,000.	1,260
2,500 . . .	250	150,000.	1,290

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

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

Table II presents the standard errors of estimated percentages of the 1977 housing inventory as well as estimated percentages of the 1975-1977 lost units. 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 1977 there were 61,300 owner-occupied housing units in this SMSA. Interpolation in table I of this appendix shows that the standard error of an estimate of this size is approximately 1,080. The following interpolation procedure was used:

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1975-1977 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)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100.	20.1	20.1	20.1	20.1	21.7	25.1
200.	11.2	11.2	11.2	11.2	15.4	17.7
500.	4.8	4.8	4.9	6.7	9.7	11.2
700.	3.5	3.5	4.1	5.7	8.2	9.5
1,000	2.5	2.5	3.5	4.8	6.9	7.9
2,500	1.0	1.0	2.2	3.0	4.3	5.0
5,000.	0.5	0.7	1.5	2.1	3.1	3.5
10,000	0.3	0.5	1.1	1.5	2.2	2.5
25,000	0.10	0.3	0.7	1.0	1.4	1.6
50,000	0.05	0.2	0.5	0.7	1.0	1.1
75,000	0.03	0.2	0.4	0.5	0.8	0.9
100,000. . . .	0.03	0.2	0.3	0.5	0.7	0.8
150,000. . . .	0.02	0.13	0.3	0.4	0.6	0.6

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

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	1,010
61,300	x
75,000	1,170

The entry for x is determined as follows by vertically interpolating between 1,010 and 1,170:

$$61{,}300-50{,}000 = 11{,}300$$
$$75{,}000-50{,}000 = 25{,}000$$

$$1{,}010+\frac{11{,}300}{25{,}000}(1{,}170-1{,}010)$$
$$= 1{,}080$$

Consequently, the 68 percent confidence interval, as shown by these data, is from 60,220 to 62,380 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1977 owner-occupied housing units lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 59,572 to 63,028 housing units with 90 percent confidence; and that the average estimate lies within the interval from 59,140 to 63,460 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 61,300 owner-occupied housing units, 11,400, or 18.6 percent, had two bedrooms. Interpolation in table II of the appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 18.6 percent is approximately 0.8 percentage points. The following interpolation procedure was used:

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

Base of percentage	Estimated percentage		
	10 or 90	18.6	25 or 75
50,000	0.7	a	1.0
61,300		p	
75,000	0.5	b	0.8

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

$$18.6-10.0 = 8.6$$
$$25.0-10.0 = 15.0$$
$$0.7 + \frac{8.6}{15.0}(1.0-0.7) = 0.87$$

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

$$18.6-10.0 = 8.6$$
$$25.0-10.0 = 15.0$$
$$0.5 + \frac{8.6}{15.0}(0.8-0.5) = 0.67$$

3. The entry for "p" was then determined by vertical interpolation between 0.87 and 0.67.

$$61,300-50,000 = 11,300$$
$$75,000-50,000 = 25,000$$
$$0.87 + \frac{11,300}{25,000}(0.67-0.87) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 17.8 to 19.4 percent; the 90-percent confidence interval is from 17.3 to 19.9 percent; and the 95-percent confidence interval is from 17.0 to 20.2 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. However, if there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1975 and 1977 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1975 and 1977 characteristics.

Illustration of the computation of the standard error of a difference.—Table A-1 of part A of this report shows that in 1977 there were 33,800 owner-occupied units with three bedrooms in this SMSA, Thus, the apparent difference, as shown by these data, between owner-occupied units with two bedrooms and owner-occupied units with three bedrooms is 22,400. Table I shows the standard error of 33,800 is approximately 850 and the standard error of 11,400 is approximately 520. Therefore, the standard error of the estimated difference of 22,400 is about:

$$1,000 = \sqrt{(850)^2 + (520)^2}$$

Consequently, the 68-percent confidence interval for the 22,400 difference is from 21,400 to 23,400 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 20,800 to 24,000 housing units, and the 95-percent confidence interval is from 20,400 to 24,400. Thus, we can conclude with 95 percent confidence that the number of 1977 owner-occupied 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, read off the confidence interval corresponding to the two points established in step 2.

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 61,300 housing units.

1. Interpolation in table II shows that the standard error of 50 percent on a base of 61,300 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 de-

termined in step 1. This yields percentage limits of 48.0 and 52.0.

3. From table A-1 of part A, it can be seen by cumulating the frequencies for the first two categories that 25,800 owner-occupied housing units, or 42.1 percent, had one or two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) and that an additional 10,500 owner-occupied housing units, or 17.1 percent, had three persons (i.e., 2.5 to 3.5 persons).

By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{48.0-42.1}{17.1}\right) = 2.8$$

Similarly, the upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{52.0-42.1}{17.1}\right) = 3.1$$

Thus, the 95-percent confidence interval ranges from 2.8 to 3.1 persons.

Memphis 1977

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1977 survey App-42
Selection of the 1974 sample App-42
1977 sample reduction App-43
1974-1977 additions to the housing inventory App-44
Coverage improvement sample selection App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1970 Census of Population and Housing App-45
ESTIMATION App-45
1977 housing inventory App-45
1974-1977 lost units App-46
1974 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES App-46
Nonsampling errors App-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-50
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-51

SAMPLE DESIGN

Annual Housing Survey.—The estimates for each of the 20 SMSA's are based on data collected from the 1977 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 20 SMSA's, the data were collected for the 11-month period from April 1977 through February 1978 with one-eleventh of the sample units visited each month.

Each group of SMSA's is interviewed for the AHS on a rotating basis, and this group (A-2 SMSA's) was the first to be revisited. Nineteen of the 20 SMSA's were interviewed for the first time in 1974, while the Madison, Wis., SMSA was interviewed for the first time in 1975.

In each group of SMSA's, the largest SMSA from each of the 4 geographic regions is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All remaining SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the A-2 group (1977-1978) are: Boston, Mass.; Detroit, Mich.; Los Angeles-Long Beach, Calif.; and Washington, D.C.-Md.-Va.

The remaining SMSA's in the A-2 group are: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.

In this SMSA, 4,544 housing units were eligible for interview. Of these sample units, 230 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 units eligible for interview, 392 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 1977 survey.—The sample housing units designated for interview in the 1977 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 1974 survey which were not part of the 1977 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 1974 survey which were not part of the 1977 reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1977 AHS questionnaire, page App-20.)

3. All sample housing units that were selected from the list of new construction building permits issued since the 1974 survey. (This sample represented the housing units built in permit-issuing areas, since the 1974 survey.)

4. All sample housing units that were added to sample segments in the nonpermit universe since the 1974 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1974 survey.)

5. All sample housing units that were selected as part of the 1977 Coverage Improvement Program. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1974 sample.—The sample for the SMSA's which are 100-percent permit-issuing was selected from two sample frames—units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and units constructed in per-

mit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which are not 100-percent permit-issuing included a sample selected from a third frame—those units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). The following SMSA's are 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Los Angeles-Long Beach, Calif.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 14 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 four largest SMSA's, the overall sampling rate differed by central city and 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 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 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 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-$5,999. . .		
$6,000-$9,999. . .		
$10,000-$14,999 .		
$15,000 and over .		

Thus, for this SMSA, the occupied housing unit records from this 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 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 are 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 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 units in existence at the time of interview in these selected segments are in sample. Thus, units enumerated in the 1970 census as well as units built since the 1970 census are included.

1977 sample reduction.—In previous years, AHS sample units have been divided among 12 panels which were interviewed each month. Due to budget limitations, it became necessary to drop one of these panels from the 1977 survey. The designated sample size for 1977 was thus reduced by one-twelfth.

1974-1977 additions to the housing inventory.—In the permit-issuing universe, a sample of new construction building permits, issued since the 1974 survey, was selected to represent housing units built in permit-issuing areas since the 1974 survey. Sampling procedures were identical to those used in selecting the 1970-1974 new construction sample. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1974, to identify any units missed in the 1974 survey or any units added since the 1974 survey.

Coverage improvement sample selection.—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 census.
3. Units missed in the 1970 census.
4. 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.

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 A-2 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 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: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Orlando, Fla.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; and Wichita, Kans.

In the remaining 10 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 and the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 3,671 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, for each 1977-1978 SMSA, 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 712 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 no 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 units in multi-unit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multi-unit structures selected, all units were listed and matched to the 1970 census. Any missed units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply these two procedures in all 20 A-2 SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in three SMSA's. In these SMSA's, a string of 10 structures (instead of 8) that were eligible to be selected for the AHS, were listed. The full application of this technique was employed in 12 SMSA's, with 5 SMSA's not receiving any application of the first procedure. The second procedure was instituted in 13 SMSA's. The following table shows which of these two techniques were applied to the A-2 SMSA's.

The first procedure added an estimated 9,058 units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 1,026 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.

SMSA	Coverage improvement technique employed			
	First		Second	Neither
	Partial	Full		
Albany-Schenectady-Troy, N.Y	X		X	
Anaheim-Santa Ana-Garden Grove, Calif	X			
Boston, Mass		X	X	
Dallas, Tex				X
Detroit, Mich		X	X	
Fort Worth, Tex		X	X	
Los Angeles-Long Beach, Calif				X
Madison, Wis		X	X	
Memphis, Tenn.-Ark		X	X	
Minneapolis-St. Paul, Minn				X
Newark, N.J		X	X	
Orlando, Fla		X	X	
Phoenix, Ariz		X	X	
Pittsburgh, Pa				X
Saginaw, Mich		X	X	
Salt Lake City, Utah				X
Spokane, Wash	X			
Tacoma, Wash		X	X	
Washington, D.C.-Md.-Va		X	X	
Wichita, Kans		X	X	

ESTIMATION

The 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 1977 housing inventory) and estimates pertaining to characteristics of units removed from the housing inventory since 1974 (i.e., 1974-1977 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1977 housing inventory.—The AHS estimates of characteristics of the 1977 housing inventory employed a 1-stage ratio estimation procedure in 19 SMSA's, with a second-stage ratio estimation procedure being employed in 1 SMSA. Prior to the implementation of the ratio estimation procedure, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of 1 or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample 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:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of 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 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 unit within the corresponding ratio estimation category.

The effect of this ratio estimation procedure was to somewhat reduce the variance due to the variation in sample size of strata used in the sample selection of the permit-issuing universe, thereby reducing the sampling error below what would have been obtained by simply weighting the results of the sample by the inverse of the probability of selection. Ordinarily, this would have been controlled by sampling within the strata during the sample selection process. However, prior to the AHS sample selection within each SMSA, units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. Thus, some variation in sample size was introduced during the AHS sample selection process.

For the Pittsburgh, Pa., SMSA, a second-stage ratio estimation procedure was also employed. This procedure involved the ratio estimation of the AHS weighted sample estimate of the October 1977 housing inventory to an independent estimate of the SMSA's October 1977 housing inventory. This ratio estimate factor equaled the following:

$$\frac{\text{SMSA independent estimate of October 1977 housing inventory}}{\text{AHS sample estimate of the housing inventory}}$$

The numerator of the ratio was derived by using the 1970 census estimate of total housing units for the SMSA in conjunction with the estimate of the change in the housing inventory since the census. The estimate of change was based on administrative records from utility companies. The denominator of that ratio was obtained from the weighted estimate of AHS sample units, using the existing weight after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for each sample unit and the resulting product was used as the final weight for tabulation purposes.

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

The second-stage ratio estimation procedure was not employed in all A-2 SMSA's, since it was felt that coverage improvement procedures adequately corrected known deficiencies in the sample where full coverage improvement procedures were employed. The independent estimates were only employed in SMSA's where full coverage improvements were not used and the estimated relative bias of the independent estimate was thought to be low enough (i.e., 2 percent or less over a 10-year period) to preclude the possibility of doing more harm than good to the survey results. A measure of the relative bias of the independent estimate was available based on a comparison between the 1970 census estimate of the April 1970 housing inventory and the independent estimate of the April 1970 housing inventory which was projected over a 10-year period from the 1960 census using the same procedure that produced the October 1977 independent estimates.

1974-1977 lost units.—The 1974-1977 lost unit estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1974 housing inventory, as was described in the 1974 Current Housing Report, Series H-170, **Housing Characteristics for Selected Metropolitan Areas.** Since the 1974-1977 lost units existed, by definition, in the 1974 housing inventory, there was a 1974 housing inventory weight associated with each 1974-1977 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1974-1977 lost units.

1974 estimation procedure.—This report presents data on the housing characteristics of the 1974 housing inventory from the 1974 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1974.

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 to provide correct information on the part of respondents, mistakes in recording or coding the data, and other errors of collection, response, processing, coverage, and estimation for missing data. Nonsampling errors are not unique to sample surveys since they can, and do, occur in complete censuses as well.

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing, the 1977 AHS-SMSA sample, and the 1974 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 included space errors, definitional errors, and occupancy errors. The "content" errors measured the accuracy of the data collected for surveyed 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.—For the 1977 AHS-SMSA sample, a reinterview program was conducted to obtain a measurement of some of the components of the nonsampling error associated with the AHS estimates. A sample of households was revisited and answers to some of the questions on the AHS questionnaire were obtained again. The original interview and the reinterview were assumed to be two independent readings and, thus, were the basis for the measurement of the accuracy of the AHS data collected from interviewed households.

As part of the reinterview, a check was made at each of these households to determine if the following was done during the original interview:

1. The correct unit was visited.
2. The correct number of housing units were interviewed at that address.
3. The correct information on "Year Built" was obtained.
4. The correct information on "Tenure" was obtained.
5. The correct information on "Household Composition" was obtained.
6. The correct information on "Type of Housing Unit" was obtained.
7. The correct information on "Occupancy Status" was obtained.

This check was made for interviewer evaluation and control. That is, tolerance limits were derived to determine which interviewers passed or failed this reinterview check. The results of this study are available and are presented in the following Census Bureau memorandum, "Response Error in the Annual Housing Survey as Measured by Reinterviews—Year IV, 1977-1978 SMSA Sample."

For the A-2 reinterview program, measures of response inconsistency were taken on two subsamples—reconciled and unreconciled. After the question is answered in the reconciled subsample, the interviewer presented the original response and then asked the respondent to decide upon the best answer. Past experience has shown that the level of inconsistency is lower when the interviewer has the response available, and thus, it is believed that there is a downward bias in these estimates. In the unreconciled sample, the interviewer did not have access to previous responses. For the reconciled items, three-fourths showed low levels of response inconsistency while the remaining one-fourth were in the moderate range. For the unreconciled items, approximately one-half showed low levels of response disagreement, with the remaining one-half of the categories in the moderate and high inconsistency range. Although attitudinal items were excluded from the A-2 reinterview program, other reinterview programs have shown that, in general, attitudinal items such as items 102-104 (see facsimile of the 1977 AHS questionnaire, page App-34) have moderate to high levels of inconsistency. Moderate levels of inconsistency indicate that there is some problem with inconsistent reporting, and high levels indicate that improvements are needed in the method used to collect data or that the category concepts themselves are ambiguous.

A second measure of response error, the net difference rate, which is a measure of net error (or bias) in a particular category, was also provided by this reinterview program. None of the categories tested showed any significant evidence of bias.

The results of this study were based on sample data, so there is a sampling error associated with these estimates of nonsampling error. Therefore, the possibility of such errors should be taken into account when considering the results of this study.

A similar study was conducted for the 1974 AHS-SMSA sample. The results of which are presented in the following Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1974."

Coverage errors.—With respect to errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of

conventional new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units, since, due to the relatively short time span involved, it is possible that construction of these units was not completed at the time the survey was conducted. In which case, they would not have been eligible for interview. In addition to these deficiencies, new construction in special places that do not require building permits, such as military bases, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit. Also, additional coverage errors may exist in SMSA's where the complete Coverage Improvement Program was not applied (see the table on page App-45 of this appendix).

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all units located inside these ED's would be represented in the sample. However, it has been estimated that the 1974 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. Since these ED's were recanvassed for the 1977 survey, the number of missed housing units may be considerably less for 1977.

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 schedules, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by the standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.

2. Approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the average result of all possible samples.

3. Approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimate would include the average result of all possible samples.

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

The figures presented in the following tables 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 1974 housing inventory can be found in the AHS Series H-170 reports for 1974.

Table I presents the standard errors applicable to estimates of characteristics of the 1977 housing inventory as well as estimates of characteristics of the 1974-1977 lost units. Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 13,470 for the total SMSA, 8,600 for the central city of the SMSA, and 10,280 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 presents the standard errors of estimated percentages of the 1977 hous-

ing inventory as well as estimated percentages of the 1974-1977 lost units. Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1977 Housing Inventory and for Estimated Number of 1974-1977 Lost Units for the Memphis, Tenn.-Ark., SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city
0	80	70	80
100	90	80	90
200	120	120	130
500	190	180	200
700	230	220	240
1,000	270	260	280
2,500	430	410	450
5,000	610	580	650
10,000	860	810	950
25,000	1,350	1,250	1,620
50,000	1,870	1,690	2,550
75,000	2,250	1,970	3,410
100,000	2,550	2,160	—
150,000	3,000	2,330	—
200,000	3,320	2,270	—
250,000	3,540	1,960	—
300,000	3,680	1,960	—

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2 for the total SMSA, 1.4 for the central city, and 1.2 for the balance (not in central city) estimates.

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.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 Lost Units for the Memphis, Tenn.-Ark., SMSA, for the Central City of the SMSA, and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	44.5	44.5	44.5	44.5	44.5	44.8
200	28.6	28.6	28.6	28.6	28.6	31.7
500	13.8	13.8	13.8	13.8	17.3	20.0
700	10.3	10.3	10.3	10.3	14.7	16.9
1,000	7.4	7.4	7.4	8.5	12.3	14.2
2,500	3.1	3.1	3.9	5.4	7.8	9.0
5,000	1.6	1.6	2.8	3.8	5.5	6.3
10,000	0.8	0.9	2.0	2.7	3.9	4.5
25,000	0.3	0.6	1.2	1.7	2.5	2.8
50,000	0.2	0.4	0.9	1.2	1.7	2.0
75,000	0.11	0.2	0.7	1.0	1.4	1.6
100,000	0.08	0.3	0.6	0.8	1.2	1.4
150,000	0.05	0.2	0.5	0.7	1.0	1.2
200,000	0.04	0.2	0.4	0.6	0.9	1.0
250,000	0.03	0.2	0.4	0.5	0.8	0.9
300,000	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 for the total SMSA, 1.4 for the central city, and 1.2 for the balance (not in central city) estimates.

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 1977 there were 155,400 owner-occupied housing units in this SMSA. Interpolation in table I of this appendix shows that the standard error of an estimate of this size is approximately 3,030. 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	3,000
155,400	x
200,000	3,320

The entry for x is determined as follows by vertically interpolating between 3,000 and 3,320:

$$155{,}400-150{,}000 = 5{,}400$$
$$200{,}000-150{,}000 = 50{,}000$$

$$3{,}000 + \frac{5{,}400}{50{,}000}(3{,}320-3{,}000) = 3{,}030$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 152,370 to 158,430 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1977 owner-occupied housing units lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 150,552 to 160,248 housing units with 90 percent confidence; and that the average estimate lies within the interval from 149,340 to 161,460 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 155,400 owner-occupied housing units, 42,300, or 27.2 percent, had two bedrooms. Interpolation in table II of the appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 27.2 percent is approximately 1.0 percentage points. The following interpolation procedure was used:

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

Base of percentage	Estimated percentage		
	25 or 75	27.2	50
150,000	1.0	a	1.2
155,400		p	
200,000	0.9	b	1.0

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

$$27.2-25.0 = 2.2$$
$$50.0-25.0 = 25.0$$
$$1.0 + \frac{2.2}{25.0}(1.2-1.0) = 1.0$$

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

$$27.2-25.0 = 2.2$$
$$50.0-25.0 = 25.0$$
$$0.9 + \frac{2.2}{25.0}(1.0-0.9) = 0.9$$

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

$$155{,}400-150{,}000 = 5{,}400$$
$$200{,}000-150{,}000 = 50{,}000$$
$$1.0 + \frac{5{,}400}{50{,}000}(0.9-1.0) = 1.0$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 26.2 to 28.2 percent; the 90-percent confidence interval is from 25.6 to 28.8 percent; and the 95-percent interval is from 25.2 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. However, 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 1974 and 1977 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1974 and 1977 characteristics.

Illustration of the computation of the standard error of a difference.—Table A-1 of part A of this report shows that in 1977 there were 84,700 owner-occupied units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied units with two bedrooms and owner-occupied units with three bedrooms is 42,400. Table I shows the standard error of 84,700 is approximately 2,370 and the standard error of 42,300 is approximately 1,710. Therefore, the standard error of the estimated difference of 42,400 is about:

$$2{,}920 = \sqrt{(2{,}370)^2 + (1{,}710)^2}$$

Consequently, the 68-percent confidence interval for the difference is from 39,480 to 45,320 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of this difference lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 37,728 to 47,072 housing units, and the 95-percent confidence interval is from 36,560 to 48,240. Thus, we can conclude with 95 percent confidence that the number of 1977 owner-occupied 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, read off the confidence interval corresponding to the two points established in step 2.

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 155,400 housing units.

1. Interpolation in table II shows that the standard error of 50 percent on a base of 155,400 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 table A-1 of part A, it can be seen by cumulating the frequencies for the first two categories that 67,700 owner-occupied housing units, or 43.6 percent, had one or two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) and that an additional 31,300 owner-occupied housing units, or 20.1 percent had three persons (i.e., 2.5 to 3.5 persons).

By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{47.6-43.6}{20.1}\right) = 2.7$$

Similarly, the upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{52.4-43.6}{20.1}\right) = 2.9$$

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

Minneapolis-St Paul 1977

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1977 survey App-42
Selection of the 1974 sample App-42
1977 sample reduction App-43
1974-1977 additions to the housing inventory App-43
Coverage improvement sample selection App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1970 Census of Population and Housing App-45
ESTIMATION App-45
1977 housing inventory App-45
1974-1977 lost units App-46
1974 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES App-46
Nonsampling errors App-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-50
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-51

SAMPLE DESIGN

Annual Housing Survey.—The estimates for each of the 20 SMSA's are based on data collected from the 1977 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 20 SMSA's, the data were collected for the 11-month period from April 1977 through February 1978 with one-eleventh of the sample units visited each month.

Each group of SMSA's is interviewed for the AHS on a rotating basis, and this group (A-2 SMSA's) was the first to be revisited. Nineteen of the 20 SMSA's were interviewed for the first time in 1974, while the Madison, Wis., SMSA was interviewed for the first time in 1975.

In each group of SMSA's, the largest SMSA from each of the 4 geographic regions is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All remaining SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the A-2 group (1977-1978) are: Boston, Mass.; Detroit, Mich.; Los Angeles-Long Beach, Calif.; and Washington, D.C.-Md.-Va.

The remaining SMSA's in the A-2 group are: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.

In this SMSA, 4,605 housing units were eligible for interview. Of these sample units, 139 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 units eligible for interview, 222 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 1977 survey.—The sample housing units designated for interview in the 1977 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 1974 survey which were not part of the 1977 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 1974 survey which were not part of the 1977 reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1977 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from the list of new construction building permits issued since the 1974 survey. (This sample represented the housing units built in permit-issuing areas, since the 1974 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1974 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1974 survey.)
5. All sample housing units that were selected as part of the 1977 Coverage Improvement Program. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1974 sample.—The sample for the SMSA's which are 100-percent permit-issuing was selected from two sample frames—units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which are

not 100-percent permit-issuing included a sample selected from a third frame—those units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). The following SMSA's are 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Los Angeles-Long Beach, Calif.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 14 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 four largest SMSA's, the overall sampling rate differed by central city and 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 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 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 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-$5,999. . .		
$6,000-$9,999. . .		
$10,000-$14,999 .		
$15,000 and over .		

Thus, for this SMSA, the occupied housing unit records from this 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 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 are 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 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 units in existence at the time of interview in these selected segments are in sample. Thus, units enumerated in the 1970 census as well as units built since the 1970 census are included.

1977 sample reduction.—In previous years, AHS sample units have been divided among 12 panels which were interviewed each month. Due to budget limitations, it became necessary to drop one of these panels from the 1977 survey. The designated sample size for 1977 was thus reduced by one-twelfth.

1974-1977 additions to the housing inventory.—In the permit-issuing universe, a sample of new construction building permits, issued since the 1974 survey, was selected to represent housing units built

in permit-issuing areas since the 1974 survey. Sampling procedures were identical to those used in selecting the 1970-1974 new construction sample. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1974, to identify any units missed in the 1974 survey or any units added since the 1974 survey.

Coverage improvement sample selection.—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 census.
3. Units missed in the 1970 census.
4. 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.

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 A-2 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 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: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Orlando, Fla.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; and Wichita, Kans.

In the remaining 10 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 and the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 6,812 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, for each 1977-1978 SMSA, 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 5,657 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 no 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 units in multi-unit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multi-unit structures selected, all units were listed and matched

to the 1970 census. Any missed units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply these two procedures in all 20 A-2 SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in three SMSA's. In these SMSA's, a string of 10 structures (instead of 8) that were eligible to be selected for the AHS, were listed. The full application of this technique was employed in 12 SMSA's, with 5 SMSA's not receiving any application of the first procedure. The second procedure was instituted in 13 SMSA's. The following table shows which of these two techniques were applied to the A-2 SMSA's.

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 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 1977 housing inventory) and estimates pertaining to characteristics of units removed from the housing inventory since 1974 (i.e., 1974-1977 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1977 housing inventory.—The AHS estimates of characteristics of the 1977 housing inventory employed a 1-stage ratio estimation procedure in 19 SMSA's, with a second-stage ratio estimation procedure being employed in 1 SMSA. Prior to the implementation of the ratio estimation procedure, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of 1 or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample 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:

SMSA	Coverage improvement technique employed			
	First		Second	Neither
	Partial	Full		
Albany-Schenectady-Troy, N.Y	X		X	
Anaheim-Santa Ana-Garden Grove, Calif	X			
Boston, Mass		X	X	
Dallas, Tex				X
Detroit, Mich		X	X	
Fort Worth, Tex		X	X	
Los Angeles-Long Beach, Calif				X
Madison, Wis		X	X	
Memphis, Tenn.-Ark		X	X	
Minneapolis-St. Paul, Minn				X
Newark, N.J		X	X	
Orlando, Fla		X	X	
Phoenix, Ariz		X	X	
Pittsburgh, Pa				X
Saginaw, Mich		X	X	
Salt Lake City, Utah				X
Spokane, Wash	X			
Tacoma, Wash		X	X	
Washington, D.C.-Md.-Va		X	X	
Wichita, Kans		X	X	

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of 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 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 unit within the corresponding ratio estimation category.

The effect of this ratio estimation procedure was to somewhat reduce the variance due to the variation in sample size of strata used in the sample selection of the permit-issuing universe, thereby reducing the sampling error below what would have been obtained by simply weighting the results of the sample by the inverse of the probability of selection. Ordinarily, this would have been controlled by sampling within the strata during the sample selection process. However, prior to the AHS sample selection within each SMSA, units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. Thus, some variation in sample size was introduced during the AHS sample selection process.

For the Pittsburgh, Pa., SMSA, a second-stage ratio estimation procedure was also employed. This procedure involved the ratio estimation of the AHS weighted sample estimate of the October 1977 housing inventory to an independent estimate of the SMSA's October 1977 housing inventory. This ratio estimate factor equaled the following:

$$\frac{\text{SMSA independent estimate of October 1977 housing inventory}}{\text{AHS sample estimate of the housing inventory}}$$

The numerator of the ratio was derived by using the 1970 census estimate of total housing units for the SMSA in conjunction with the estimate of the change in the housing inventory since the census. The estimate of change was based on administrative records from utility companies. The denominator of that ratio was obtained from the weighted estimate of AHS sample units, using the existing weight after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for each sample unit and the resulting product was used as the final weight for tabulation purposes.

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

The second-stage ratio estimation procedure was not employed in all A-2 SMSA's, since it was felt that coverage improvement procedures adequately corrected known deficiencies in the sample where full coverage improvement procedures were employed. The independent estimates were only employed in SMSA's where full coverage improvements were not used and the estimated relative bias of the independent estimate was thought to be low enough (i.e., 2 percent or less over a 10-year period) to preclude the possibility of doing more harm than good to the survey results. A measure of the relative bias of the independent estimate was available based on a comparison between the 1970 census estimate of the April 1970 housing inventory and the independent estimate of the April 1970 housing inventory which was projected over a 10-year period from the 1960 census using the same procedure that produced the October 1977 independent estimates.

1974-1977 lost units.—The 1974-1977 lost unit estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1974 housing inventory, as was described in the 1974 Current Housing Report, Series H-170, **Housing Characteristics for Selected Metropolitan Areas.** Since the 1974-1977 lost units existed, by definition, in the 1974 housing inventory, there was a 1974 housing inventory weight associated with each 1974-1977 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1974-1977 lost units.

1974 estimation procedure.—This report presents data on the housing characteristics of the 1974 housing inventory from the 1974 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1974.

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 to provide correct information on the part of respondents, mistakes in recording or coding the data, and other errors of collection, response, processing, coverage, and estimation for missing data. Nonsampling errors are not unique to sample surveys since they can, and do, occur in complete censuses as well.

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing, the 1977 AHS-SMSA sample, and the 1974 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 included space errors, definitional errors, and occupancy errors. The "content" errors measured the accuracy of the data collected for surveyed 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.—For the 1977 AHS-SMSA sample, a reinterview program was conducted to obtain a measurement of some of the components of the nonsampling error associated with the AHS estimates. A sample of households was revisited and answers to some of the questions on the AHS questionnaire were obtained again. The original interview and the reinterview were assumed to be two independent readings and, thus, were the basis for the measurement of the accuracy of the AHS data collected from interviewed households.

As part of the reinterview, a check was made at each of these households to determine if the following was done during the original interview:

1. The correct unit was visited.
2. The correct number of housing units were interviewed at that address.
3. The correct information on "Year Built" was obtained.
4. The correct information on "Tenure" was obtained.
5. The correct information on "Household Composition" was obtained.
6. The correct information on "Type of Housing Unit" was obtained.
7. The correct information on "Occupancy Status" was obtained.

This check was made for interviewer evaluation and control. That is, tolerance limits were derived to determine which interviewers passed or failed this reinterview check. The results of this study are available and are presented in the following Census Bureau memorandum, "Response Error in the Annual Housing Survey as Measured by Reinterviews—Year IV, 1977-1978 SMSA Sample."

For the A-2 reinterview program, measures of response inconsistency were taken on two subsamples—reconciled and unreconciled. After the question is answered in the reconciled subsample, the interviewer presented the original response and then asked the respondent to decide upon the best answer. Past experience has shown that the level of inconsistency is lower when the interviewer has the response available, and thus, it is believed that there is a downward bias in these estimates. In the unreconciled sample, the interviewer did not have access to previous responses. For the reconciled items, three-fourths showed low levels of response inconsistency while the remaining one-fourth were in the moderate range. For the unreconciled items, approximately one-half showed low levels of response disagreement, with the remaining one-half of the categories in the moderate and high inconsistency range. Although attitudinal items were excluded from the A-2 reinterview program, other reinterview programs have shown that, in general, attitudinal items such as items 102-104 (see facsimile of the 1977 AHS questionnaire, page App-34) have moderate to high levels of inconsistency. Moderate levels of inconsistency indicate that there is some problem with inconsistent reporting, and high levels indicate that improvements are needed in the method used to collect data or that the category concepts themselves are ambiguous.

A second measure of response error, the net difference rate, which is a measure of net error (or bias) in a particular category, was also provided by this reinterview program. None of the categories tested showed any significant evidence of bias.

The results of this study were based on sample data, so there is a sampling error associated with these estimates of nonsampling error. Therefore, the possibility of such errors should be taken into account when considering the results of this study.

A similar study was conducted for the 1974 AHS-SMSA sample. The results of which are presented in the following Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1974."

Coverage errors.—With respect to errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units, since, due to the relatively short time span involved, it is possible that construction of these units was not completed at the time the survey was conducted. In which case, they would not

have been eligible for interview. In addition to these deficiencies, new construction in special places that do not require building permits, such as military bases, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit. Also, additional coverage errors may exist in SMSA's where the complete Coverage Improvement Program was not applied (see the table on page App-45 of this appendix).

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all units located inside these ED's would be represented in the sample. However, it has been estimated that the 1974 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. Since these ED's were recanvassed for the 1977 survey, the number of missed housing units may be considerably less for 1977.

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 schedules, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by the standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

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

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

The figures presented in the following tables 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 1974 housing inventory can be found in the AHS Series H-170 reports for 1974.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1977 Housing Inventory and for Estimated Number of 1974-1977 Lost Units for the Minneapolis-St. Paul, Minn., SMSA, for the Central Cities of the SMSA and for the Balance (Not in Central Cities) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]		
	SMSA	In central cities	Not in central cities
0	150	140	160
100.	150	140	160
200.	170	170	180
500.	270	260	290
700.	320	310	340
1,000 . . .	380	370	410
2,500 . . .	610	590	640
5,000 . . .	860	830	900
10,000 . .	1,210	1,160	1,270
25,000 . .	1,900	1,810	2,000
50,000 . .	2,660	2,500	2,780
75,000 . .	3,230	2,990	3,350
100,000. .	3,690	3,350	3,800
150,000. .	4,420	3,860	4,490
200,000. .	4,990	4,160	4,980
250,000. .	5,450	4,300	5,340
300,000. .	5,820	4,280	5,580
400,000. .	6,370	–	5,780
500,000. .	6,710	–	5,630
600,000. .	6,860	–	–
700,000. .	6,840	–	–

[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, 1.7 for the central cities, and 1.2 for the balance (not in central cities) estimates.

Table I presents the standard errors applicable to estimates of characteristics of the 1977 housing inventory as well as estimates of characteristics of the 1974-1977 lost units. Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 23,233 for the total SMSA, 10,842 for the central city of the SMSA, and 20,478 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 presents the standard errors of estimated percentages of the 1977 housing inventory as well as estimated percentages of the 1974-1977 lost units. 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

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 Lost Units for the Minneapolis-St. Paul, Minn., SMSA, for the Central Cities of the SMSA, and for the Balance (Not in Central Cities) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	62.2	62.2	62.2	62.2	62.2	64.1
200	45.1	45.1	45.1	45.1	45.1	45.3
500	24.7	24.7	24.7	24.7	24.8	28.7
700	19.0	19.0	19.0	19.0	21.0	24.2
1,000	14.1	14.1	14.1	14.1	17.6	20.3
2,500	6.2	6.2	6.2	7.7	11.1	12.8
5,000	3.2	3.2	4.0	5.4	7.9	9.1
10,000	1.6	1.6	2.8	3.8	5.6	6.4
25,000	0.7	0.8	1.8	2.4	3.5	4.1
50,000	0.3	0.6	1.2	1.7	2.5	2.9
75,000	0.2	0.5	1.0	1.4	2.0	2.3
100,000	0.16	0.4	0.9	1.2	1.8	2.0
150,000	0.11	0.3	0.7	1.0	1.4	1.7
200,000	0.08	0.3	0.6	0.9	1.2	1.4
250,000	0.07	0.3	0.6	0.8	1.1	1.3
300,000	0.05	0.2	0.5	0.7	1.0	1.2
400,000	0.04	0.2	0.4	0.6	0.9	1.0
500,000	0.03	0.2	0.4	0.5	0.8	0.9
600,000	0.03	0.2	0.4	0.5	0.7	0.8
700,000	0.02	0.15	0.3	0.5	0.7	0.8

[1]Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.3 for the total SMSA, 1.7 for the central cities, and 1.2 for the balance (not in central cities) estimates.

Illustration of the use of the standard error tables.—Table A-1 of part A of this report shows that in 1977 there were 430,900 owner-occupied housing units in this SMSA. Interpolation in table I of this appendix shows that the standard error of an estimate of this size is approximately 6,480. The following interpolation procedure was used:

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

Size of estimate	Standard error
400,000	6,370
430,900	x
500,000	6,710

The entry for x is determined as follows by vertically interpolating between 6,370 and 6,710.

$$430{,}900-400{,}000 = 30{,}900$$
$$500{,}000-400{,}000 = 100{,}000$$

$$6{,}370 + \frac{30{,}900}{100{,}000}(6{,}710-6{,}370)$$
$$= 6{,}480$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 424,420 to 437,380 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1977 owner-occupied housing units lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 420,532 to 441,268 housing units with 90 percent confidence; and that the average estimate lies within the interval from 417,940 to 443,860 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 430,900 owner-occupied housing units, 116,800, or 27.1 percent, had two bedrooms. Interpolation in table II of the appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 27.1 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	27.1	50
400,000. . . .	0.9	a	1.0
430,900. . . .		p	
500,000. . . .	0.8	b	0.9

1. The entry for cell "a" is determined by horizontal interpolation between 0.9 and 1.0.

$$27.1-25.0 = 2.1$$
$$50.0-25.0 = 25.0$$

$$0.9 + \frac{2.1}{25.0}(1.0-0.9) = 0.9$$

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

$$27.1-25.0 = 2.1$$
$$50.0-25.0 = 25.0$$

$$0.8 + \frac{2.1}{25.0}(0.9-0.8) = 0.8$$

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

$$430{,}900-400{,}000 = 30{,}900$$
$$500{,}000-400{,}000 = 100{,}000$$

$$0.9 + \frac{30{,}900}{100{,}000}(0.8-0.9) = 0.9$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 26.2 to 28.0 percent; the 90-percent confidence interval is from 25.7 to 28.5 percent; and the 95-percent confidence interval is from 25.3 to 28.9 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. However, 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 1974 and 1977 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1974 and 1977 characteristics.

Illustration of the computation of the standard error of a difference.—Table A-1 of part A of this report shows that in 1977 there were 198,500 owner-occupied units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied units with two bedrooms and owner-occupied units with three bedrooms is 81,700. Table I shows the standard error of 198,500 is approximately 4,970 and the standard error of 116,800 is approximately 3,940. Table I shows the standard error of 198,500 is approximately 4,970. Therefore, the standard error of the estimated difference of 81,700 is about:

$$6{,}340 = \sqrt{(4{,}970)^2 + (3{,}940)^2}$$

Consequently, the 68-percent confidence interval for the 81,700 difference is from 75,360 to 88,040 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 71,556 to 91,844 housing units, and the 95-percent confidence interval is from 69,020 to 94,380. Thus, we can conclude with 95 percent confidence that the number of 1977 owner-occupied 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, read off the confidence interval corresponding to the two points established in step 2.

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

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

Illustration of the computation of the 95-percent confidence interval of a median.—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 3.1. The base of the distribution from which this median was determined is 430,900 housing units.

1. Interpolation in table II shows that the standard error of 50 percent on a base of 430,900 is approximately 1.0 percentage points.

2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.0 and 52.0.

3. From table A-1 of part A, it can be seen by cumulating the frequencies for the first two categories that 167,100 owner-occupied housing units, or 38.8 percent, had one or two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) and that an additional 79,100 owner-occupied housing units, or 18.4 percent, had three persons (i.e., 2.5 to 3.5 persons).

By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{48.0-38.8}{18.4}\right) = 3.0$$

Similarly, the upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{52.0-38.8}{18.4}\right) = 3.2$$

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

Newark 1977

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1977 survey App-42
Selection of the 1974 sample App-42
1977 sample reduction App-43
1974-1977 additions to the housing inventory App-44
Coverage improvement sample selection App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1970 Census of Population and Housing App-45
ESTIMATION App-45
1977 housing inventory App-45
1974-1977 lost units App-46
1974 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES App-47
Nonsampling errors App-47
1970 census App-47
AHS-SMSA App-47
Coverage errors App-48
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-50
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-51

SAMPLE DESIGN

Annual Housing Survey.—The estimates for each of the 20 SMSA's are based on data collected from the 1977 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 20 SMSA's, the data were collected for the 11-month period from April 1977 through February 1978 with one-eleventh of the sample units visited each month.

Each group of SMSA's is interviewed for the AHS on a rotating basis, and this group (A-2 SMSA's) was the first to be revisited. Nineteen of the 20 SMSA's were interviewed for the first time in 1974, while the Madison, Wis., SMSA was interviewed for the first time in 1975.

In each group of SMSA's, the largest SMSA from each of the 4 geographic regions is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All remaining SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the A-2 group (1977-1978) are: Boston, Mass.; Detroit, Mich.; Los Angeles-Long Beach, Calif.; and Washington, D.C.-Md.-Va.

The remaining SMSA's in the A-2 group are: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.

In this SMSA, 4,563 housing units were eligible for interview. Of these sample units, 242 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 units eligible for interview, 256 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 1977 survey.—The sample housing units designated for interview in the 1977 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 1974 survey which were not part of the 1977 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 1974 survey which were not part of the 1977 reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1977 AHS questionnaire, page App-20.)

3. All sample housing units that were selected from the list of new construction building permits issued since the 1974 survey. (This sample represented the housing units built in permit-issuing areas, since the 1974 survey.)

4. All sample housing units that were added to sample segments in the nonpermit universe since the 1974 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1974 survey.)

5. All sample housing units that were selected as part of the 1977 Coverage Improvement Program. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1974 sample.—The sample for the SMSA's which are 100-percent permit-issuing was selected from two sample frames—units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and units constructed in

permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which are not 100-percent permit-issuing included a sample selected from a third frame—those units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). The following SMSA's are 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Los Angeles-Long Beach, Calif.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 14 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 four largest SMSA's, the overall sampling rate differed by central city and 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 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 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 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-$5,999. . .		
$6,000-$9,999. . .		
$10,000-$14,999 .		
$15,000 and over .		

Thus, for this SMSA, the occupied housing unit records from this 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 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 are 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 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 units in existence at the time of interview in these selected segments are in sample. Thus, units enumerated in the 1970 census as well as units built since the 1970 census are included.

1977 sample reduction.—In previous years, AHS sample units have been divided among 12 panels which were interviewed each month. Due to budget limitations, it became necessary to drop one of these panels from the 1977 survey. The

designated sample size for 1977 was thus reduced by one-twelfth.

1974-1977 additions to the housing inventory.—In the permit-issuing universe, a sample of new construction building permits, issued since the 1974 survey, was selected to represent housing units built in permit-issuing areas since the 1974 survey. Sampling procedures were identical to those used in selecting the 1970-1974 new construction sample. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1974, to identify any units missed in the 1974 survey or any units added since the 1974 survey.

Coverage improvement sample selection.—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 census.

3. Units missed in the 1970 census.

4. 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.

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 A-2 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 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: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Orlando, Fla.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; and Wichita, Kans.

In the remaining 10 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 and the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 3,308 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, for each 1977-1978 SMSA, 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 no 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 units in multi-unit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multi-unit structures selected, all units were listed and matched to the 1970 census. Any missed units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply these two procedures in all 20 A-2 SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in three SMSA's. In these SMSA's, a string of 10 structures (instead of 8) that were eligible to be selected for the AHS, were listed. The full application of this technique was employed in 12 SMSA's, with 5 SMSA's not receiving any application of the first procedure. The second procedure was instituted in 13 SMSA's. The following table shows which of these two techniques were applied to the A-2 SMSA's.

SMSA	Coverage improvement technique employed			
	First		Second	Neither
	Partial	Full		
Albany-Schenectady-Troy, N.Y	X		X	
Anaheim-Santa Ana-Garden Grove, Calif	X			
Boston, Mass		X	X	
Dallas, Tex				X
Detroit, Mich		X	X	
Fort Worth, Tex		X	X	
Los Angeles-Long Beach, Calif				X
Madison, Wis		X	X	
Memphis, Tenn.-Ark		X	X	
Minneapolis-St. Paul, Minn				X
Newark, N.J		X	X	
Orlando, Fla		X	X	
Phoenix, Ariz		X	X	
Pittsburgh, Pa				X
Saginaw, Mich		X	X	
Salt Lake City, Utah				X
Spokane, Wash	X			
Tacoma, Wash		X	X	
Washington, D.C.-Md.-Va		X	X	
Wichita, Kans		X	X	

The first procedure added an estimated 13,437 units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 5,269 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 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 1977 housing inventory) and estimates pertaining to characteristics of units removed from the housing inventory since 1974 (i.e., 1974-1977 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1977 housing inventory.—The AHS estimates of characteristics of the 1977 housing inventory employed a 1-stage ratio estimation procedure in 19 SMSA's, with a second-stage ratio estimation procedure being employed in 1 SMSA. Prior to the implementation of the ratio estimation procedure, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of 1 or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes from both the

nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample 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:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of 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 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 unit within the corresponding ratio estimation category.

The effect of this ratio estimation procedure was to somewhat reduce the variance due to the variation in sample size of strata used in the sample selection of the permit-issuing universe, thereby reducing the sampling error below what would have been obtained by simply weighting the results of the sample by the inverse of the probability of selection. Ordinarily, this would have been controlled by sampling within the strata during the sample selection process. However, prior to the AHS sample selection within each SMSA, units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. Thus, some variation in sample size was introduced during the AHS sample selection process.

For the Pittsburgh, Pa., SMSA, a second-stage ratio estimation procedure was also employed. This procedure involved the ratio estimation of the AHS weighted sample estimate of the October 1977 housing inventory to an independent estimate of the SMSA's October 1977 housing inventory. This ratio estimate factor equaled the following:

$$\frac{\text{SMSA independent estimate of October 1977 housing inventory}}{\text{AHS sample estimate of the housing inventory}}$$

The numerator of the ratio was derived by using the 1970 census estimate of total housing units for the SMSA in conjunction with the estimate of the change in the housing inventory since the census. The estimate of change was based on administrative records from utility companies. The denominator of that ratio was obtained from the weighted estimate of AHS sample units, using the existing weight after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for each sample unit and the resulting product was used as the final weight for tabulation purposes.

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

The second-stage ratio estimation procedure was not employed in all A-2 SMSA's, since it was felt that coverage improvement procedures adequately corrected known deficiencies in the sample where full coverage improvement procedures were employed. The independent estimates were only employed in SMSA's where full coverage improvements were not used and the estimated relative bias of the independent estimate was thought to be low enough (i.e., 2 percent or less over a 10-year period) to preclude the possibility of doing more harm than good to the survey results. A measure of the relative bias of the independent estimate was available based on a comparison between the 1970 census estimate of the April 1970 housing inventory and the independent estimate of the April 1970 housing inventory which was projected over a 10-year period from the 1960 census using the same procedure that produced the October 1977 independent estimates.

1974-1977 lost units.—The 1974-1977 lost unit estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1974 housing inventory, as was described in the 1974 Current Housing Report, Series H-170, **Housing Characteristics for Selected Metropolitan Areas.** Since the 1974-1977 lost units existed, by definition, in the 1974 housing inventory, there was a 1974 housing inventory weight associated with each 1974-1977 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1974-1977 lost units.

1974 estimation procedure.—This report presents data on the housing characteristics of the 1974 housing inventory from the 1974 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1974.

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 to provide correct information on the part of respondents, mistakes in recording or coding the data, and other errors of collection, response, processing, coverage, and estimation for missing data. Nonsampling errors are not unique to sample surveys since they can, and do, occur in complete censuses as well.

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing, the 1977 AHS-SMSA sample, and the 1974 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 included space errors, definitional errors, and occupancy errors. The "content" errors measured the accuracy of the data collected for surveyed 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.—For the 1977 AHS-SMSA sample, a reinterview program was conducted to obtain a measurement of some of the components of the nonsampling error associated with the AHS estimates. A sample of households was revisited and answers to some of the questions on the AHS questionnaire were obtained again. The original interview and the reinterview were assumed to be two independent readings and, thus, were the basis for the measurement of the accuracy of the AHS data collected from interviewed households.

As part of the reinterview, a check was made at each of these households to determine if the following was done during the original interview:

1. The correct unit was visited.
2. The correct number of housing units were interviewed at that address.
3. The correct information on "Year Built" was obtained.
4. The correct information on "Tenure" was obtained.
5. The correct information on "Household Composition" was obtained.
6. The correct information on "Type of Housing Unit" was obtained.
7. The correct information on "Occupancy Status" was obtained.

This check was made for interviewer evaluation and control. That is, tolerance limits were derived to determine which interviewers passed or failed this reinterview check. The results of this study are available and are presented in the following Census Bureau memorandum, "Response Error in the Annual Housing Survey as Measured by Reinterviews—Year IV, 1977-1978 SMSA Sample."

For the A-2 reinterview program, measures of response inconsistency were taken on two subsamples—reconciled and unreconciled. After the question is answered in the reconciled subsample, the interviewer presented the original response and then asked the respondent to decide upon the best answer. Past experience has shown that the level of inconsistency is lower when the interviewer has the response available, and thus, it is believed that there is a downward bias in these estimates. In the unreconciled sample, the interviewer did not have access to previous responses. For the reconciled items, three-fourths showed low levels of response inconsistency while the remaining one-fourth were in the moderate range. For the unreconciled items, approximately one-half showed low levels of response disagreement, with the remaining one-half of the categories in the moderate and high inconsistency range. Although attitudinal items were excluded from the A-2 reinterview program, other reinterview programs have shown that, in general, attitudinal items such as items 102-104 (see facsimile of the 1977 AHS questionnaire, page App-34) have moderate to high levels of inconsistency. Moderate levels of inconsistency indicate that there is some problem with inconsistent reporting, and high levels indicate that improvements are needed in the method used to collect data or that the category concepts themselves are ambiguous.

A second measure of response error, the net difference rate, which is a measure of net error (or bias) in a particular category, was also provided by this reinterview program. None of the categories tested showed any significant evidence of bias.

The results of this study were based on sample data, so there is a sampling error associated with these estimates of nonsampling error. Therefore, the possibility of such errors should be taken into account when considering the results of this study.

A similar study was conducted for the 1974 AHS-SMSA sample. The results of which are presented in the following

Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1974."

Coverage errors.—With respect to errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units, since, due to the relatively short time span involved, it is possible that construction of these units was not completed at the time the survey was conducted. In which case, they would not have been eligible for interview. In addition to these deficiencies, new construction in special places that do not require building permits, such as military bases, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit. Also, additional coverage errors may exist in SMSA's where the complete Coverage Improvement Program was not applied (see the table on page App-45 of this appendix).

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all units located inside these ED's would be represented in the sample. However, it has been estimated that the 1974 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. Since these ED's were recanvassed for the 1977 survey, the number of missed housing units may be considerably less for 1977.

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 schedules, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by the standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

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

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

The figures presented in the following tables 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 1974 housing inventory can be found in the AHS Series H-170 reports for 1974.

Table I presents the standard errors applicable to estimates of characteristics of the 1977 housing inventory as well as estimates of characteristics of the 1974-1977 lost units. 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 16,844 for the total SMSA, 9,527 for the central city of the SMSA, and 13,843 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

TABLE I. Standard Errors for Estimated Numbers of Housing Units in the 1977 Housing Inventory and for Estimated Number of 1974-1977 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]
0	150	75,000 . .	3,200
100	150	100,000 .	3,630
200	170	150,000 .	4,270
500	270	200,000 .	4,730
700	330	250,000 .	5,060
1,000 . .	390	300,000 .	5,270
2,500 . .	610	400,000 .	5,400
5,000 . .	870	500,000 .	5,160
10,000 .	1,220	600,000 .	4,500
25,000 .	1,910	700,000 .	3,140
50,000 .	2,660		

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.4.

which the percentage is based. Estimated percentages are relatively more reliable than the corresponding estimates of the numerators of the percentages, particularly if the percentages are 50 percent or more.

Table II presents the standard errors of estimated percentages of the 1977 housing inventory as well as estimated percentages of the 1974-1977 lost units. 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 1977 there were 329,700 owner-occupied housing units in this SMSA. Interpolation in table I of this appendix shows that the standard error of an estimate of this size is approximately 5,310. 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	5,270
329,700	x
400,000	5,400

The entry for x is determined as follows by vertically interpolating between 5,270 and 5,400:

$$329{,}700-300{,}000 = 29{,}700$$
$$400{,}000-300{,}000 = 100{,}000$$

$$5{,}270 + \frac{29{,}700}{100{,}000}(5{,}400-5{,}270) = 5{,}310$$

Consequently, the 68 percent confidence interval, as shown by these data, is from 324,390 to 335,010 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1977 owner-occupied housing units lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 321,204 to 338,196 housing units with 90 percent confidence; and that the average estimate lies within the interval from 319,080 to 340,320 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 329,700 owner-occupied housing units, 71,000, or 21.5 percent, had two bedrooms. Interpolation in table II of the appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 21.5 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	21.5	25 or 75
300,000	0.7	a	1.0
329,700		p	
400,000	0.6	b	0.8

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

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

$$0.7 + \frac{11.5}{15.0}(1.0-0.7) = 0.9$$

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

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

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

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

$$329{,}700-300{,}000 = 29{,}700$$
$$400{,}000-300{,}000 = 100{,}000$$

$$0.9 + \frac{29{,}700}{100{,}000}(0.8-0.9) = 0.9$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 20.6 to 22.4 percent; the 90-percent confidence interval is from 20.1 to 22.9 percent; and the 95-percent interval is from 19.7 to 23.3 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. However, 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 1974 and 1977 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1974 and 1977 characteristics.

Illustration of the computation of the standard error of a difference.—Table A-1 of part A of this report shows that in 1977 there were 153,900 owner-occupied units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied units with two bedroomss and owner-occupied units with three bedrooms is 82,900. Table I shows the standard error of 71,000 is approximately 3,110 and the standard error of 153,900 is approximately 4,310. Therefore, the standard error of the estimated difference of 82,900 is about:

$$5{,}310 = \sqrt{(3{,}110)^2 + (4{,}310)^2}$$

Consequently, the 68-percent confidence interval for the 82,900 difference is from 77,590 to 88,210 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 74,404 to 91,396 housing units, and the 95-percent confidence interval is from 72,280 to 93,520. Thus, we can conclude with 95 percent confidence that the number of 1977 owner-occupied 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 pro-

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 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)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	60.2	60.2	60.2	60.2	60.2	61.5
200	43.1	43.1	43.1	43.1	43.1	43.5
500	23.2	23.2	23.2	23.2	23.8	27.5
700	17.8	17.8	17.8	17.8	20.1	23.2
1,000	13.1	13.1	13.1	13.1	16.8	19.4
2,500	5.7	5.7	5.7	7.4	10.6	12.3
5,000	2.9	2.9	3.8	5.2	7.5	8.7
10,000	1.5	1.5	2.7	3.7	5.3	6.1
25,000	0.6	0.8	1.7	2.3	3.4	3.9
50,000	0.3	0.5	1.2	1.6	2.4	2.7
75,000	0.2	0.4	1.0	1.3	1.9	2.2
100,000	0.15	0.4	0.8	1.2	1.7	1.9
150,000	0.10	0.3	0.7	1.0	1.4	1.6
200,000	0.08	0.3	0.6	0.8	1.2	1.4
250,000	0.06	0.2	0.5	0.7	1.1	1.2
300,000	0.05	0.2	0.5	0.7	1.0	1.1
400,000	0.04	0.2	0.4	0.6	0.8	1.0
500,000	0.03	0.2	0.4	0.5	0.8	0.9
600,000	0.03	0.16	0.3	0.5	0.7	0.8
700,000	0.02	0.15	0.3	0.4	0.6	0.7

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.4 for the total SMSA, for the central city, and for the balance (not in central city) estimates.

cedure 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, read off the confidence interval corresponding to the two points established in step 2.

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

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

Illustration of the computation of the 95-percent confidence interval of a median.—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 3.1. The base of the distribution from which this median was determined is 329,700 housing units.

1. Interpolation in table II shows that the standard error of 50 percent on a base of 329,700 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 table A-1 of part A, it can be seen by cumulating the frequencies for the first two categories that 121,900 owner-occupied housing units, or 37.0 percent, had one or two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) and that an additional 65,900 owner-occupied housing units, or 20.0 percent, had three persons (i.e., 2.5 to 3.5 persons).

By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{47.8-37.0}{20.0}\right) = 3.0$$

Similarly, the upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{52.2-37.0}{20.0}\right) = 3.3$$

Thus, the 95-percent confidence interval ranges from 3.0 to 3.3 persons.

Orlando 1977

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1977 survey App-42
Selection of the 1974 sample App-42
1977 sample reduction App-43
1974-1977 additions to the housing inventory App-43
Coverage improvement sample selection App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1970 Census of Population and Housing App-45
ESTIMATION App-45
1977 housing inventory App-45
1974-1977 lost units App-46
1974 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES App-46
Nonsampling errors App-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-50
Illustration of the computation of the standard error of a difference App-50
Medians App-51
Illustration of the computation of the 95-percent confidence interval of a median ... App-51

SAMPLE DESIGN

Annual Housing Survey.—The estimates for each of the 20 SMSA's are based on data collected from the 1977 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 20 SMSA's, the data were collected for the 11-month period from April 1977 through February 1978 with one-eleventh of the sample units visited each month.

Each group of SMSA's is interviewed for the AHS on a rotating basis, and this group (A-2 SMSA's) was the first to be revisited. Nineteen of the 20 SMSA's were interviewed for the first time in 1974, while the Madison, Wis., SMSA was interviewed for the first time in 1975.

In each group of SMSA's, the largest SMSA from each of the 4 geographic regions is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All remaining SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the A-2 group (1977-1978) are: Boston, Mass.; Detroit, Mich.; Los Angeles-Long Beach, Calif.; and Washington, D.C.-Md.-Va.

The remaining SMSA's in the A-2 group are: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.

In this SMSA, 4,700 housing units were eligible for interview. Of these sample units, 153 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 units eligible for interview, 418 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 1977 survey.—The sample housing units designated for interview in the 1977 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 1974 survey which were not part of the 1977 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 1974 survey which were not part of the 1977 reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1977 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from the list of new construction building permits issued since the 1974 survey. (This sample represented the housing units built in permit-issuing areas, since the 1974 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1974 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1974 survey.)
5. All sample housing units that were selected as part of the 1977 Coverage Improvement Program. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1974 sample.—The sample for the SMSA's which are 100-percent permit-issuing was selected from two sample frames—units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which are not 100-percent permit-issuing included

a sample selected from a third frame—those units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). The following SMSA's are 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Los Angeles-Long Beach, Calif.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 14 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 four largest SMSA's, the overall sampling rate differed by central city and 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 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 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 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-$5,999. . .		
$6,000-$9,999. . .		
$10,000-$14,999 .		
$15,000 and over .		

Thus, for this SMSA, the occupied housing unit records from this 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 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 are 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 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 units in existence at the time of interview in these selected segments are in sample. Thus, units enumerated in the 1970 census as well as units built since the 1970 census are included.

1977 sample reduction.—In previous years, AHS sample units have been divided among 12 panels which were interviewed each month. Due to budget limitations, it became necessary to drop one of these panels from the 1977 survey. The designated sample size for 1977 was thus reduced by one-twelfth.

1974-1977 additions to the housing inventory.—In the permit-issuing universe, a sample of new construction building permits, issued since the 1974 survey, was

selected to represent housing units built in permit-issuing areas since the 1974 survey. Sampling procedures were identical to those used in selecting the 1970-1974 new construction sample. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1974, to identify any units missed in the 1974 survey or any units added since the 1974 survey.

Coverage improvement sample selection.—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 census.
3. Units missed in the 1970 census.
4. 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.

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 A-2 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 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: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Orlando, Fla.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; and Wichita, Kans.

In the remaining 10 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 and the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 870 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, for each 1977-1978 SMSA, 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 1,135 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 no 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 units in multi-unit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multi-unit structures

selected, all units were listed and matched to the 1970 census. Any missed units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply these two procedures in all 20 A-2 SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in three SMSA's. In these SMSA's, a string of 10 structures (instead of 8) that were eligible to be selected for the AHS, were listed. The full application of this technique was employed in 12 SMSA's, with 5 SMSA's not receiving any application of the first procedure. The second procedure was instituted in 13 SMSA's. The following table shows which of these two techniques were applied to the A-2 SMSA's.

The first procedure added an estimated 12,351 units to the coverage of the housing inventory of this SMSA, while the second procedure added no additional 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 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 1977 housing inventory) and estimates pertaining to characteristics of units removed from the housing inventory since 1974 (i.e., 1974-1977 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1977 housing inventory.—The AHS estimates of characteristics of the 1977 housing inventory employed a 1-stage ratio estimation procedure in 19 SMSA's, with a second-stage ratio estimation procedure being employed in 1 SMSA. Prior to the implementation of the ratio estimation procedure, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of 1 or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample 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

SMSA	Coverage improvement technique employed			
	First		Second	Neither
	Partial	Full		
Albany-Schenectady-Troy, N.Y	X		X	
Anaheim-Santa Ana-Garden Grove, Calif	X			
Boston, Mass		X	X	
Dallas, Tex				X
Detroit, Mich		X	X	
Fort Worth, Tex		X	X	
Los Angeles-Long Beach, Calif				X
Madison, Wis		X	X	
Memphis, Tenn.-Ark		X	X	
Minneapolis-St. Paul, Minn				X
Newark, N.J		X	X	
Orlando, Fla		X	X	
Phoenix, Ariz		X	X	
Pittsburgh, Pa				X
Saginaw, Mich		X	X	
Salt Lake City, Utah				X
Spokane, Wash	X			
Tacoma, Wash		X	X	
Washington, D.C.-Md.-Va		X	X	
Wichita, Kans		X	X	

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:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of 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 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 unit within the corresponding ratio estimation category.

The effect of this ratio estimation procedure was to somewhat reduce the variance due to the variation in sample size of strata used in the sample selection of the permit-issuing universe, thereby reducing the sampling error below what would have been obtained by simply weighting the results of the sample by the inverse of the probability of selection. Ordinarily, this would have been controlled by sampling within the strata during the sample selection process. However, prior to the AHS sample selection within each SMSA, units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. Thus, some variation in sample size was introduced during the AHS sample selection process.

For the Pittsburgh, Pa., SMSA, a second-stage ratio estimation procedure was also employed. This procedure involved the ratio estimation of the AHS weighted sample estimate of the October 1977 housing inventory to an independent estimate of the SMSA's October 1977 housing inventory. This ratio estimate factor equaled the following:

$$\frac{\text{SMSA independent estimate of October 1977 housing inventory}}{\text{AHS sample estimate of the housing inventory}}$$

The numerator of the ratio was derived by using the 1970 census estimate of total housing units for the SMSA in conjunction with the estimate of the change in the housing inventory since the census. The estimate of change was based on administrative records from utility companies. The denominator of that ratio was obtained from the weighted estimate of AHS sample units, using the existing weight after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for each sample unit and the resulting product was used as the final weight for tabulation purposes.

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

The second-stage ratio estimation procedure was not employed in all A-2 SMSA's, since it was felt that coverage improvement procedures adequately corrected known deficiencies in the sample where full coverage improvement procedures were employed. The independent estimates were only employed in SMSA's where full coverage improvements were not used and the estimated relative bias of the independent estimate was thought to be low enough (i.e., 2 percent or less over a 10-year period) to preclude the possibility of doing more harm than good to the survey results. A measure of the relative bias of the independent estimate was available based on a comparison between the 1970 census estimate of the April 1970 housing inventory and the independent estimate of the April 1970 housing inventory which was projected over a 10-year period from the 1960 census using the same procedure that produced the October 1977 independent estimates.

1974-1977 lost units.—The 1974-1977 lost unit estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1974 housing inventory, as was described in the 1974 Current Housing Report, Series H-170, **Housing Characteristics for Selected Metropolitan Areas.** Since the 1974-1977 lost units existed, by definition, in the 1974 housing inventory, there was a 1974 housing inventory weight associated with each 1974-1977 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1974-1977 lost units.

1974 estimation procedure.—This report presents data on the housing characteristics of the 1974 housing inventory from the 1974 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1974.

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 to provide correct information on the part of respondents, mistakes in recording or coding the data, and other errors of collection, response, processing, coverage, and estimation for missing data. Nonsampling errors are not unique to sample surveys since they can, and do, occur in complete censuses as well.

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing, the 1977 AHS-SMSA sample, and the 1974 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 included space errors, definitional errors, and occupancy errors. The "content" errors measured the accuracy of the data collected for surveyed 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.—For the 1977 AHS-SMSA sample, a reinterview program was conducted to obtain a measurement of some of the components of the nonsampling error associated with the AHS estimates. A sample of households was revisited and answers to some of the questions on the AHS questionnaire were obtained again. The original interview and the reinterview were assumed to be two independent readings and, thus, were the basis for the measurement of the accuracy of the AHS data collected from interviewed households.

As part of the reinterview, a check was made at each of these households to determine if the following was done during the original interview:

1. The correct unit was visited.
2. The correct number of housing units were interviewed at that address.
3. The correct information on "Year Built" was obtained.
4. The correct information on "Tenure" was obtained.
5. The correct information on "Household Composition" was obtained.
6. The correct information on "Type of Housing Unit" was obtained.
7. The correct information on "Occupancy Status" was obtained.

This check was made for interviewer evaluation and control. That is, tolerance limits were derived to determine which interviewers passed or failed this reinterview check. The results of this study are available and are presented in the following Census Bureau memorandum, "Response Error in the Annual Housing Survey as Measured by Reinterviews—Year IV, 1977-1978 SMSA Sample."

For the A-2 reinterview program, measures of response inconsistency were taken on two subsamples—reconciled and unreconciled. After the question is answered in the reconciled subsample, the interviewer presented the original response and then asked the respondent to decide upon the best answer. Past experience has shown that the level of inconsistency is lower when the interviewer has the response available, and thus, it is believed that there is a downward bias in these estimates. In the unreconciled sample, the interviewer did not have access to previous responses. For the reconciled items, three-fourths showed low levels of response inconsistency while the remaining one-fourth were in the moderate range. For the unreconciled items, approximately one-half showed low levels of response disagreement, with the remaining one-half of the categories in the moderate and high inconsistency range. Although attitudinal items were excluded from the A-2 reinterview program, other reinterview programs have shown that, in general, attitudinal items such as items 102-104 (see facsimile of the 1977 AHS questionnaire, page App-34) have moderate to high levels of inconsistency. Moderate levels of inconsistency indicate that there is some problem with inconsistent reporting, and high levels indicate that improvements are needed in the method used to collect data or that the category concepts themselves are ambiguous.

A second measure of response error, the net difference rate, which is a measure of net error (or bias) in a particular category, was also provided by this reinterview program. None of the categories tested showed any significant evidence of bias.

The results of this study were based on sample data, so there is a sampling error associated with these estimates of nonsampling error. Therefore, the possibility of such errors should be taken into account when considering the results of this study.

A similar study was conducted for the 1974 AHS-SMSA sample. The results of which are presented in the following Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1974."

Coverage errors.—With respect to errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing

units, since, due to the relatively short time span involved, it is possible that construction of these units was not completed at the time the survey was conducted. In which case, they would not have been eligible for interview. In addition to these deficiencies, new construction in special places that do not require building permits, such as military bases, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit. Also, additonal coverage errors may exist in SMSA's where the complete Coverage Improvement Program was not applied (see the table on page App-45 of this appendix).

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all units located inside these ED's would be represented in the sample. However, it has been estimated that the 1974 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. Since these ED's were recanvassed for the 1977 survey, the number of missed housing units may be considerably less for 1977.

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 schedules, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by the standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

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

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

The figures presented in the following tables 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 1974 housing inventory can be found in the AHS Series H-170 reports for 1974.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1977 Housing Inventory and for Estimated Number of 1974-1977 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]	
	SMSA and not in central city	Central city
0	60	40
100	70	60
200	110	90
500	170	140
700	200	170
1,000	240	200
2,500	370	320
5,000	530	450
10,000	740	620
25,000	1,150	920
50,000	1,590	1,150
75,000	1,900	–
100,000	2,130	–
150,000	2,440	–
200,000	2,620	–
250,000	2,690	–

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2 for the total SMSA, 1.4 for the central city, and 1.2 for the balance (not in central city) estimates.

Table I presents the standard errors applicable to estimates of characteristics of the 1977 housing inventory as well as estimates of characteristics of the 1974-1977 lost units. 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 8,550 for the total SMSA, 2,520 for the central city of the SMSA, and 8,080 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 present the standard errors of estimated percentages of the 1977 housing inventory as well as estimated percentages of the 1974-1977 lost units. 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 through 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 1977 there were 130,200 owner-occupied housing units in this SMSA. Interpolation in table I of this appendix shows that the standard error of an estimate of this size is approximately 2,320. 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.	2,130
130,200.	x
150,000.	2,440

The entry for x is determined as follows by vertically interpolating between 2,130 and 2,440:

130,200−100,000 = 30,200
150,000−100,000 = 50,000

$$2{,}130+\frac{30{,}200}{50{,}000}(2{,}440-2{,}130)$$
$$=2{,}320$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 127,880 to 132,520 housing units. There-

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 Lost Units for the Orlando, Fla., SMSA, and the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100.	35.9	35.9	35.9	35.9	35.9	37.4
200.	21.9	21.9	21.9	21.9	22.9	26.5
500.	10.1	10.1	10.1	10.1	14.5	16.7
700.	7.4	7.4	7.4	8.5	12.2	14.1
1,000	5.3	5.3	5.3	7.1	10.2	11.8
2,500	2.2	2.2	3.3	4.5	6.5	7.5
5,000	1.1	1.1	2.3	3.2	4.6	5.3
10,000	0.6	0.7	1.6	2.2	3.2	3.7
25,000	0.2	0.5	1.0	1.4	2.0	2.4
50,000	0.11	0.3	0.7	1.0	1.4	1.7
75,000	0.07	0.3	0.6	0.8	1.2	1.4
100,000. . . .	0.06	0.2	0.5	0.7	1.0	1.2
150,000. . . .	0.04	0.2	0.4	0.6	0.8	1.0
200,000. . . .	0.03	0.2	0.4	0.5	0.7	0.8
250,000. . . .	0.02	0.15	0.3	0.4	0.6	0.7

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

fore, a conclusion that the average estimate, derived from all possible samples, of 1977 owner-occupied housing units lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 126,488 to 133,912 housing units with 90 percent confidence; and that the average estimate lies within the interval from 125,560 to 134,840 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 130,200 owner-occupied housing units, 31,100, or 23.9 percent, had two bedrooms. Interpolation in table II of the appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 23.9 percent is approximately 0.9 percentage points. The following interpolation procedure was used:

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

Base of percentage	Estimated percentage		
	10 or 90	23.9	25 or 75
100,000. . . .	0.7	a	1.0
130,200. . . .		p	
150,000. . . .	0.6	b	0.8

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

$$23.9-10.0 = 13.9$$
$$25.0-10.0 = 15.0$$

$$0.7 + \frac{13.9}{15.0}(1.0-0.7) = 0.98$$

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

$$23.9-10.0 = 13.9$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{13.9}{15.0}(0.8-0.6) = 0.79$$

3. The entry for "p" was then determined by vertical interpolation between 0.98 and 0.78.

$$130{,}200-100{,}000 = 30{,}200$$
$$150{,}000-100{,}000 = 50{,}000$$

$$0.98 + \frac{30{,}200}{50{,}000}(0.79-0.98) = 0.9$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 23.0 to 24.8 percent; the 90-percent confidence interval is from 22.5 to 25.3 percent; and the 95-percent confidence interval is from 22.1 to 25.7 percent.

Differences.—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. However, 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 1974 and 1977 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1974 and 1977 characteristics.

Illustration of the computation of the standard error of a difference.—Table A-1 of part A of this report shows that in 1977 there were 69,800 owner-occupied units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied units with two bedrooms and owner-occupied units with three bedrooms is 38,700. Table I shows the standard error of 69,800 is approximately 1,840 and the standard error of 31,100 is approximately 1,260. Therefore, the standard error of the estimated difference of 38,700 is about:

TABLE III. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 Lost Housing Units for the Central City of the Orlando, Fla., SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100.	29.5	29.5	29.5	29.5	29.5	32.3
200.	17.3	17.3	17.3	17.3	19.8	22.9
500.	7.7	7.7	7.7	8.7	12.5	14.5
700.	5.6	5.6	5.6	7.3	10.6	12.2
1,000	4.0	4.0	4.5	6.1	8.9	10.2
2,500	1.6	1.6	2.8	3.9	5.6	6.5
5,000	0.8	0.9	2.0	2.7	4.0	4.6
10,000	0.4	0.6	1.4	1.9	2.8	3.2
25,000	0.2	0.4	0.9	1.2	1.8	2.0
50,000	0.08	0.3	0.6	0.9	1.3	1.4

[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 1.4.

$$2{,}230 = \sqrt{(1{,}840)^2 + (1{,}260)^2}$$

Consequently, the 68-percent confidence interval for the 38,700 difference is from 36,470 to 40,930 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 35,132 to 42,268 housing units, and the 95-percent confidence interval is from 34,240 to 43,160. Thus, we can conclude with 95 percent confidence that the number of 1977 owner-occupied 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, read off the confidence interval corresponding to the two points established in step 2.

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 130,200 housing units.

1. Interpolation in table II shows that the standard error of 50 percent on a base of 130,200 is approximately 1.1 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 47.8 and 52.2.
3. From table A-1 of part A, it can be seen by cumulating the frequencies for the first two categories that 61,800 owner-occupied housing units, or 47.5 percent, had one or two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) and that an additional 24,300 owner-occupied housing units, or 18.7 percent, had three persons (i.e., 2.5 to 3.5 persons).

By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{47.8-47.5}{18.7}\right) = 2.5$$

Similarly, the upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{52.2-47.5}{18.7}\right) = 2.8$$

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

Phoenix 1977

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1977 survey App-42
Selection of the 1974 sample App-42
1977 sample reduction App-43
1974-1977 additions to the housing inventory App-43
Coverage improvement sample selection App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1970 Census of Population and Housing App-45
ESTIMATION App-45
1977 housing inventory App-45
1974-1977 lost units App-46
1974 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES App-46
Nonsampling errors App-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-50
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

SAMPLE DESIGN

Annual Housing Survey.—The estimates for each of the 20 SMSA's are based on data collected from the 1977 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 20 SMSA's, the data were collected for the 11-month period from April 1977 through February 1978 with one-eleventh of the sample units visited each month.

Each group of SMSA's is interviewed for the AHS on a rotating basis, and this group (A-2 SMSA's) was the first to be revisited. Nineteen of the 20 SMSA's were interviewed for the first time in 1974, while the Madison, Wis., SMSA was interviewed for the first time in 1975.

In each group of SMSA's, the largest SMSA from each of the 4 geographic regions is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All remaining SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the A-2 group (1977-1978) are: Boston, Mass.; Detroit, Mich.; Los Angeles-Long Beach, Calif.; and Washington, D.C.-Md.-Va.

The remaining SMSA's in the A-2 group are: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.

In this SMSA, 5,207 housing units were eligible for interview. Of these sample units, 193 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 units eligible for interview, 358 units were visited but were not eligible for interview because they were condemned, unfit, demolished, converted to group quarters use, etc.

Designation of sample housing units for the 1977 survey.—The sample housing units designated for interview in the 1977 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 1974 survey which were not part of the 1977 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 1974 survey which were not part of the 1977 reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1977 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from the list of new construction building permits issued since the 1974 survey. (This sample represented the housing units built in permit-issuing areas, since the 1974 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1974 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1974 survey.)
5. All sample housing units that were selected as part of the 1977 Coverage Improvement Program. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1974 sample.—The sample for the SMSA's which are 100-percent permit-issuing was selected from two sample frames—units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which are not 100-percent permit-issuing included

a sample selected from a third frame—those units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). The following SMSA's are 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Los Angeles-Long Beach, Calif.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 14 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 four largest SMSA's, the overall sampling rate differed by central city and 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 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 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 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-$5,999. . .		
$6,000-$9,999. . .		
$10,000-$14,999 .		
$15,000 and over .		

Thus, for this SMSA, the occupied housing unit records from this 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 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 are 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 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 units in existence at the time of interview in these selected segments are in sample. Thus, units enumerated in the 1970 census as well as units built since the 1970 census are included.

1977 sample reduction.—In previous years, AHS sample units have been divided among 12 panels which were interviewed each month. Due to budget limitations, it became necessary to drop one of these panels from the 1977 survey. The designated sample size for 1977 was thus reduced by one-twelfth.

1974-1977 additions to the housing inventory.—In the permit-issuing universe, a sample of new construction building permits, issued since the 1974 survey, was

selected to represent housing units built in permit-issuing areas since the 1974 survey. Sampling procedures were identical to those used in selecting the 1970-1974 new construction sample. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1974, to identify any units missed in the 1974 survey or any units added since the 1974 survey.

Coverage improvement sample selection.—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 census.
3. Units missed in the 1970 census.
4. 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.

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 A-2 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 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: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Orlando, Fla.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; and Wichita, Kans.

In the remaining 10 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 and the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 4,510 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, for each 1977-1978 SMSA, 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 9,521 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 no 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 units in multi-unit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multi-unit structures

selected, all units were listed and matched to the 1970 census. Any missed units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply these two procedures in all 20 A-2 SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in three SMSA's. In these SMSA's, a string of 10 structures (instead of 8) that were eligible to be selected for the AHS, were listed. The full application of this technique was employed in 12 SMSA's, with 5 SMSA's not receiving any application of the first procedure. The second procedure was instituted in 13 SMSA's. The following table shows which of these two techniques were applied to the A-2 SMSA's.

The first procedure added an estimated 13,335 units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 435 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 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 1977 housing inventory) and estimates pertaining to characteristics of units removed from the housing inventory since 1974 (i.e., 1974-1977 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1977 housing inventory.—The AHS estimates of characteristics of the 1977 housing inventory employed a 1-stage ratio estimation procedure in 19 SMSA's, with a second-stage ratio estimation procedure being employed in 1 SMSA. Prior to the implementation of the ratio estimation procedure, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of 1 or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample 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

SMSA	Coverage improvement technique employed			
	First		Second	Neither
	Partial	Full		
Albany-Schenectady-Troy, N.Y	X		X	
Anaheim-Santa Ana-Garden Grove, Calif	X			
Boston, Mass		X	X	
Dallas, Tex				X
Detroit, Mich		X	X	
Fort Worth, Tex		X	X	
Los Angeles-Long Beach, Calif				X
Madison, Wis		X	X	
Memphis, Tenn.-Ark		X	X	
Minneapolis-St. Paul, Minn				X
Newark, N.J		X	X	
Orlando, Fla		X	X	
Phoenix, Ariz		X	X	
Pittsburgh, Pa				X
Saginaw, Mich		X	X	
Salt Lake City, Utah				X
Spokane, Wash	X			
Tacoma, Wash		X	X	
Washington, D.C.-Md.-Va		X	X	
Wichita, Kans		X	X	

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:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of 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 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 unit within the corresponding ratio estimation category.

The effect of this ratio estimation procedure was to somewhat reduce the variance due to the variation in sample size of strata used in the sample selection of the permit-issuing universe, thereby reducing the sampling error below what would have been obtained by simply weighting the results of the sample by the inverse of the probability of selection. Ordinarily, this would have been controlled by sampling within the strata during the sample selection process. However, prior to the AHS sample selection within each SMSA, units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. Thus, some variation in sample size was introduced during the AHS sample selection process.

For the Pittsburgh, Pa., SMSA, a second-stage ratio estimation procedure was also employed. This procedure involved the ratio estimation of the AHS weighted sample estimate of the October 1977 housing inventory to an independent estimate of the SMSA's October 1977 housing inventory. This ratio estimate factor equaled the following:

$$\frac{\text{SMSA independent estimate of October 1977 housing inventory}}{\text{AHS sample estimate of the housing inventory}}$$

The numerator of the ratio was derived by using the 1970 census estimate of total housing units for the SMSA in conjunction with the estimate of the change in the housing inventory since the census. The estimate of change was based on administrative records from utility companies. The denominator of that ratio was obtained from the weighted estimate of AHS sample units, using the existing weight after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for each sample unit and the resulting product was used as the final weight for tabulation purposes.

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

The second-stage ratio estimation procedure was not employed in all A-2 SMSA's, since it was felt that coverage improvement procedures adequately corrected known deficiencies in the sample where full coverage improvement procedures were employed. The independent estimates were only employed in SMSA's where full coverage improvements were not used and the estimated relative bias of the independent estimate was thought to be low enough (i.e., 2 percent or less over a 10-year period) to preclude the possibility of doing more harm than good to the survey results. A measure of the relative bias of the independent estimate was available based on a comparison between the 1970 census estimate of the April 1970 housing inventory and the independent estimate of the April 1970 housing inventory which was projected over a 10-year period from the 1960 census using the same procedure that produced the October 1977 independent estimates.

1974-1977 lost units.—The 1974-1977 lost unit estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1974 housing inventory, as was described in the 1974 Current Housing Report, Series H-170, **Housing Characteristics for Selected Metropolitan Areas.** Since the 1974-1977 lost units existed, by definition, in the 1974 housing inventory, there was a 1974 housing inventory weight associated with each 1974-1977 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1974-1977 lost units.

1974 estimation procedure.—This report presents data on the housing characteristics of the 1974 housing inventory from the 1974 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1974.

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 to provide correct information on the part of respondents, mistakes in recording or coding the data, and other errors of collection, response, processing, coverage, and estimation for missing data. Nonsampling errors are not unique to sample surveys since they can, and do, occur in complete censuses as well.

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing, the 1977 AHS-SMSA sample, and the 1974 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 included space errors, definitional errors, and occupancy errors. The "content" errors measured the accuracy of the data collected for surveyed 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.—For the 1977 AHS-SMSA sample, a reinterview program was conducted to obtain a measurement of some of the components of the nonsampling error associated with the AHS estimates. A sample of households was revisited and answers to some of the questions on the AHS questionnaire were obtained again. The original interview and the reinterview were assumed to be two independent readings and, thus, were the basis for the measurement of the accuracy of the AHS data collected from interviewed households.

As part of the reinterview, a check was made at each of these households to determine if the following was done during the original interview:

1. The correct unit was visited.
2. The correct number of housing units were interviewed at that address.
3. The correct information on "Year Built" was obtained.
4. The correct information on "Tenure" was obtained.
5. The correct information on "Household Composition" was obtained.
6. The correct information on "Type of Housing Unit" was obtained.
7. The correct information on "Occupancy Status" was obtained.

This check was made for interviewer evaluation and control. That is, tolerance limits were derived to determine which interviewers passed or failed this reinterview check. The results of this study are available and are presented in the following Census Bureau memorandum, "Response Error in the Annual Housing Survey as Measured by Reinterviews—Year IV, 1977-1978 SMSA Sample."

For the A-2 reinterview program, measures of response inconsistency were taken on two subsamples—reconciled and unreconciled. After the question is answered in the reconciled subsample, the interviewer presented the original response and then asked the respondent to decide upon the best answer. Past experience has shown that the level of inconsistency is lower when the interviewer has the response available, and thus, it is believed that there is a downward bias in these estimates. In the unreconciled sample, the interviewer did not have access to previous responses. For the reconciled items, three-fourths showed low levels of response inconsistency while the remaining one-fourth were in the moderate range. For the unreconciled items, approximately one-half showed low levels of response disagreement, with the remaining one-half of the categories in the moderate and high inconsistency range. Although attitudinal items were excluded from the A-2 reinterview program, other reinterview programs have shown that, in general, attitudinal items such as items 102-104 (see facsimile of the 1977 AHS questionnaire, page App-34) have moderate to high levels of inconsistency. Moderate levels of inconsistency indicate that there is some problem with inconsistent reporting, and high levels indicate that improvements are needed in the method used to collect data or that the category concepts themselves are ambiguous.

A second measure of response error, the net difference rate, which is a measure of net error (or bias) in a particular category, was also provided by this reinterview program. None of the categories tested showed any significant evidence of bias.

The results of this study were based on sample data, so there is a sampling error associated with these estimates of nonsampling error. Therefore, the possibility of such errors should be taken into account when considering the results of this study.

A similar study was conducted for the 1974 AHS-SMSA sample. The results of which are presented in the following Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1974."

Coverage errors.—With respect to errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing

units, since, due to the relatively short time span involved, it is possible that construction of these units was not completed at the time the survey was conducted. In which case, they would not have been eligible for interview. In addition to these deficiencies, new construction in special places that do not require building permits, such as military bases, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit. Also, additional coverage errors may exist in SMSA's where the complete Coverage Improvement Program was not applied (see the table on page App-45 of this appendix).

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all units located inside these ED's would be represented in the sample. However, it has been estimated that the 1974 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. Since these ED's were recanvassed for the 1977 survey, the number of missed housing units may be considerably less for 1977.

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 schedules, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by the standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

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

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

The figures presented in the following tables 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 1974 housing inventory can be found in the AHS Series H-170 reports for 1974.

Table I presents the standard errors applicable to estimates of characteristics of the 1977 housing inventory as well as estimates of characteristics of the 1974-1977 lost units. Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1977 Housing Inventory and for Estimated Number of 1974-1977 Lost 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)

Size of estimate	Standard error[1]	Size of estimate	Standard error[1]
0	110	50,000. .	2,340
100.	110	75,000. .	2,880
200.	150	100,000.	3,330
500.	230	150,000.	4,100
700.	280	200,000.	4,760
1,000 . . .	330	250,000.	5,350
2,500 . . .	520	300,000.	5,900
5,000 . . .	740	400,000.	6,880
10,000 . .	1,040	500,000.	7,770
25,000 . .	1,650		

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

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 21,180 for the total SMSA, 15,290 for the central city of the SMSA, and 14,570 for the balance of the SMSA.

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

Table II presents the standard errors of estimated percentages of the 1977 housing inventory as well as estimated percentages of the 1974-1977 lost units. 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

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 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)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100.	52.0	52.0	52.0	52.0	52.0	52.1
200.	35.2	35.2	35.2	35.2	35.2	36.8
500.	17.8	17.8	17.8	17.8	20.2	23.3
700.	13.4	13.4	13.4	13.4	17.0	19.7
1,000	9.8	9.8	9.8	9.9	14.3	16.5
2,500	4.2	4.2	4.5	6.3	9.0	10.4
5,000	2.1	2.1	3.2	4.4	6.4	7.4
10,000	1.1	1.1	2.3	3.1	4.5	5.2
25,000	0.4	0.7	1.4	2.0	2.9	3.3
50,000	0.2	0.5	1.0	1.4	2.0	2.3
75,000	0.14	0.4	0.8	1.1	1.6	1.9
100,000. . . .	0.11	0.3	0.7	1.0	1.4	1.6
150,000. . . .	0.07	0.3	0.6	0.8	1.2	1.3
200,000. . . .	0.05	0.2	0.5	0.7	1.0	1.2
250,000. . . .	0.04	0.2	0.5	0.6	0.9	1.0
300,000. . . .	0.04	0.2	0.4	0.6	0.8	1.0
400,000. . . .	0.03	0.2	0.4	0.5	0.7	0.8
500,000. . . .	0.02	0.15	0.3	0.4	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, for the central city, and for the balance (not in central city) estimates.

Illustration of the use of the standard error tables.—Table A-1 of part A of this report shows that in 1977 there were 312,900 owner-occupied housing units in this SMSA. Interpolation in table I of this appendix shows that the standard error of an estimate of this size is approximately 6,030. 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.	5,900
312,900.	x
400,000.	6,880

The entry for x is determined as follows by vertically interpolating between 5,900 and 6,880:

$$312{,}900-300{,}000 = 12{,}900$$
$$400{,}000-300{,}000 = 100{,}000$$

$$5{,}900 + \frac{12{,}900}{100{,}000}(6{,}880-5{,}900)$$
$$= 6{,}030$$

Consequently, the 68 percent confidence interval, as shown by these data, is from 306,870 to 318,930 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1977 owner-occupied housing units lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies

within the interval from 303,252 to 322,548 housing units with 90 percent confidence; and that the average estimate lies within the interval from 300,840 to 324,960 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 312,900 owner-occupied housing units, 83,100, or 26.6 percent, had two bedrooms. Interpolation in table II of the appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 26.6 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	26.6	50
300,000. . . .	0.8	a	1.0
312,900. . . .		p	
400,000. . . .	0.7	b	0.8

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

$$26.6-25.0 = 1.6$$
$$50.0-25.0 = 25.0$$

$$0.8 + \frac{1.6}{25.0}(1.0-0.8) = 0.81$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.7 and 0.8.

$$26.6-25.0 = 1.6$$
$$50.0-25.0 = 25.0$$

$$0.7 + \frac{1.6}{25.0}(0.8-0.7) = 0.71$$

3. The entry for "p" was then determined by vertical interpolation between 0.81 and 0.71.

$$312{,}900-300{,}000 = 12{,}900$$
$$400{,}000-300{,}000 = 100{,}000$$

$$0.81 + \frac{12{,}900}{100{,}000}(0.71-0.81) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 25.8 to 27.4 percent; the 90-percent confidence interval is from 25.3 to 27.9 percent; and the 95-percent confidence interval is from 25.0 to 28.2 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. However, 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 1974 and 1977 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1974 and 1977 characteristics.

Illustration of the computation of the standard error of a difference.—Table A-1 of part A of this report shows that in 1977 there were 155,700 owner-occupied units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied units with two bedrooms and owner-occupied units with three bedrooms is 72,600. Table I shows the standard error of 155,700 is approximately 4,180, and the standard error of 83,100 is approximately 3,030. Therefore, the standard error of the estimated difference of 72,600 is about:

$$5{,}160 = \sqrt{(4{,}180)^2 + (3{,}030)^2}$$

Consequently, the 68-percent confidence interval for the 72,600 difference is from 67,400 to 77,760 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 64,344 to 80,856 housing units, and the 95-percent confidence interval is from 62,280 to 82,920. Thus, we can conclude with 95 percent confidence that the number of 1977 owner-occupied 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, read off the confidence interval corresponding to the two points established in step 2.

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 312,900 housing units.

1. Interpolation in table II shows that the standard error of 50 percent on a base of 312,900 is approximately 1.0 percentage points.

2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.0 and 52.0.

3. From table A-1 of part A, it can be seen by cumulating the frequencies for the first two categories that 155,300 owner-occupied housing units, or 50.0 percent, had one or two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) and that an additional 52,800 owner-occupied housing units, or 16.9 percent, had three persons (i.e., 2.5 to 3.5 persons).

By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{48.0-50.0}{16.9} = 2.4$$

Similarly, the upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{52.0-50.0}{16.9} = 2.6$$

Thus, the 95-percent confidence interval ranges from 2.4 to 2.6 persons.

Pittsburgh 1977

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1977 survey App-42
Selection of the 1974 sample App-42
1977 sample reduction App-43
1974-1977 additions to the housing inventory App-44
Coverage improvement sample selection App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1970 Census of Population and Housing App-45
ESTIMATION App-45
1977 housing inventory App-45
1974-1977 lost units App-46
1974 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES App-46
Nonsampling errors App-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-50
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-51

SAMPLE DESIGN

Annual Housing Survey.—The estimates for each of the 20 SMSA's are based on data collected from the 1977 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 20 SMSA's, the data were collected for the 11-month period from April 1977 through February 1978 with one-eleventh of the sample units visited each month.

Each group of SMSA's is interviewed for the AHS on a rotating basis, and this group (A-2 SMSA's) was the first to be revisited. Nineteen of the 20 SMSA's were interviewed for the first time in 1974, while the Madison, Wis., SMSA was interviewed for the first time in 1975.

In each group of SMSA's, the largest SMSA from each of the 4 geographic regions is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All remaining SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the A-2 group (1977-1978) are: Boston, Mass.; Detroit, Mich.; Los Angeles-Long Beach, Calif.; and Washington, D.C.-Md.-Va.

The remaining SMSA's in the A-2 group are: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.

In this SMSA, 4,617 housing units were eligible for interview. Of these sample 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 units eligible for interview, 285 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 1977 survey.—The sample housing units designated for interview in the 1977 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 1974 survey which were not part of the 1977 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 1974 survey which were not part of the 1977 reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1977 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from the list of new construction building permits issued since the 1974 survey. (This sample represented the housing units built in permit-issuing areas, since the 1974 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1974 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1974 survey.)
5. All sample housing units that were selected as part of the 1977 Coverage Improvement Program. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1974 sample.—The sample for the SMSA's which are 100-percent permit-issuing was selected from two sample frames—units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addi-

tion, the sample for those SMSA's which are not 100-percent permit-issuing included a sample selected from a third frame—those units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). The following SMSA's are 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Los Angeles-Long Beach, Calif.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 14 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 four largest SMSA's, the overall sampling rate differed by central city and 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 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 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 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-$5,999. . .		
$6,000-$9,999. . .		
$10,000-$14,999 .		
$15,000 and over .		

Thus, for this SMSA, the occupied housing unit records from this 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 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 are 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 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 units in existence at the time of interview in these selected segments are in sample. Thus, units enumerated in the 1970 census as well as units built since the 1970 census are included.

1977 sample reduction.—In previous years, AHS sample units have been divided among 12 panels which were interviewed each month. Due to budget limitations, it became necessary to drop one of these panels from the 1977 survey. The designated sample size for 1977 was thus reduced by one-twelfth.

1974-1977 additions to the housing inventory.—In the permit-issuing universe, a sample of new construction building permits, issued since the 1974 survey, was selected to represent housing units built in permit-issuing areas since the 1974 survey. Sampling procedures were identical to those used in selecting the 1970-1974 new construction sample. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1974, to identify any units missed in the 1974 survey or any units added since the 1974 survey.

Coverage improvement sample selection.—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 census.
3. Units missed in the 1970 census.
4. 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.

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 A-2 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 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: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Orlando, Fla.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; and Wichita, Kans.

In the remaining 10 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 and the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 2,151 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, for each 1977-1978 SMSA, 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 2,981 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 no 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 units in multi-unit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multi-unit structures selected, all units were listed and matched to the 1970 census. Any missed units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply these two procedures in all 20 A-2 SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in three SMSA's. In these SMSA's, a string of 10 structures (instead of 8) that were eligible to be selected for the AHS, were listed. The full application of this technique was employed in 12 SMSA's, with 5 SMSA's not receiving any application of the first procedure. The second procedure was instituted in 13 SMSA's. The following table shows which of these two techniques were applied to the A-2 SMSA's.

SMSA	Coverage improvement technique employed			
	First		Second	Neither
	Partial	Full		
Albany-Schenectady-Troy, N.Y	X		X	
Anaheim-Santa Ana-Garden Grove, Calif	X			
Boston, Mass		X	X	
Dallas, Tex				X
Detroit, Mich		X	X	
Fort Worth, Tex		X	X	
Los Angeles-Long Beach, Calif				X
Madison, Wis		X	X	
Memphis, Tenn.-Ark		X	X	
Minneapolis-St. Paul, Minn				X
Newark, N.J		X	X	
Orlando, Fla		X	X	
Phoenix, Ariz		X	X	
Pittsburgh, Pa				X
Saginaw, Mich		X	X	
Salt Lake City, Utah				X
Spokane, Wash	X			
Tacoma, Wash		X	X	
Washington, D.C.-Md.-Va		X	X	
Wichita, Kans		X	X	

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 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 1977 housing inventory) and estimates pertaining to characteristics of units removed from the housing inventory since 1974 (i.e., 1974-1977 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1977 housing inventory.—The AHS estimates of characteristics of the 1977 housing inventory employed a 1-stage ratio estimation procedure in 19 SMSA's, with a second-stage ratio estimation procedure being employed in 1 SMSA. Prior to the implementation of the ratio estimation procedure, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of 1 or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample 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:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of 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 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 unit within the corresponding ratio estimation category.

The effect of this ratio estimation procedure was to somewhat reduce the variance due to the variation in sample size of strata used in the sample selection of the permit-issuing universe, thereby reducing the sampling error below what would have been obtained by simply weighting the results of the sample by the inverse of the probability of selection. Ordinarily, this would have been controlled by sampling within the strata during the sample selection process. However, prior to the AHS sample selection within each SMSA, units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. Thus, some variation in sample size was introduced during the AHS sample selection process.

For the Pittsburgh, Pa., SMSA, a second-stage ratio estimation procedure was also employed. This procedure involved the ratio estimation of the AHS weighted sample estimate of the October 1977 housing inventory to an independent estimate of the SMSA's October 1977 housing inventory. This ratio estimate factor equaled the following:

$$\frac{\text{SMSA independent estimate of October 1977 housing inventory}}{\text{AHS sample estimate of the housing inventory}}$$

The numerator of the ratio was derived by using the 1970 census estimate of total housing units for the SMSA in conjunction with the estimate of the change in the housing inventory since the census. The estimate of change was based on administrative records from utility companies. The denominator of that ratio was obtained from the weighted estimate of AHS sample units, using the existing weight after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for each sample unit and the resulting product was used as the final weight for tabulation purposes.

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

The second-stage ratio estimation procedure was not employed in all A-2 SMSA's, since it was felt that coverage improvement procedures adequately corrected known deficiencies in the sample where full coverage improvement procedures were employed. The independent estimates were only employed in SMSA's where full coverage improvements were not used and the estimated relative bias of the independent estimate was thought to be low enough (i.e., 2 percent or less over a 10-year period) to preclude the possibility of doing more harm than good to the survey results. A measure of the relative bias of the independent estimate was available based on a comparison between the 1970 census estimate of the April 1970 housing inventory and the independent estimate of the April 1970 housing inventory which was projected over a 10-year period from the 1960 census using the same procedure that produced the October 1977 independent estimates.

1974-1977 lost units.—The 1974-1977 lost unit estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1974 housing inventory, as was described in the 1974 Current Housing Report, Series H-170, **Housing Characteristics for Selected Metropolitan Areas.** Since the 1974-1977 lost units existed, by definition, in the 1974 housing inventory, there was a 1974 housing inventory weight associated with each 1974-1977 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1974-1977 lost units.

1974 estimation procedure.—This report presents data on the housing characteristics of the 1974 housing inventory from the 1974 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1974.

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 to provide correct information on the part of respondents, mistakes in recording or coding the data, and other errors of collection, response, processing, coverage, and estimation for missing data. Nonsampling errors are not unique to sample surveys since they can, and do, occur in complete censuses as well.

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing, the 1977 AHS-SMSA sample, and the 1974 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 included space errors, definitional errors, and occupancy errors. The "content" errors measured the accuracy of the data collected for surveyed 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.—For the 1977 AHS-SMSA sample, a reinterview program was conducted to obtain a measurement of some of the components of the nonsampling error associated with the AHS estimates. A sample of households was revisited and answers to some of the questions on the AHS questionnaire were obtained again. The original interview and the reinterview were assumed to be two independent readings and, thus, were the basis for the measurement of the accuracy of the AHS data collected from interviewed households.

As part of the reinterview, a check was made at each of these households to determine if the following was done during the original interview:

1. The correct unit was visited.
2. The correct number of housing units were interviewed at that address.
3. The correct information on "Year Built" was obtained.
4. The correct information on "Tenure" was obtained.
5. The correct information on "Household Composition" was obtained.
6. The correct information on "Type of Housing Unit" was obtained.
7. The correct information on "Occupancy Status" was obtained.

This check was made for interviewer evaluation and control. That is, tolerance limits were derived to determine which interviewers passed or failed this reinterview check. The results of this study are available and are presented in the following Census Bureau memorandum, "Response Error in the Annual Housing Survey as Measured by Reinterviews—Year IV, 1977-1978 SMSA Sample."

For the A-2 reinterview program, measures of response inconsistency were taken on two subsamples—reconciled and unreconciled. After the question is answered in the reconciled subsample, the interviewer presented the original response and then asked the respondent to decide upon the best answer. Past experience has shown that the level of inconsistency is lower when the interviewer has the response available, and thus, it is believed that there is a downward bias in these estimates. In the unreconciled sample, the interviewer did not have access to previous responses. For the reconciled items, three-fourths showed low levels of response inconsistency while the remaining one-fourth were in the moderate range. For the unreconciled items, approximately one-half showed low levels of response disagreement, with the remaining one-half of the categories in the moderate and high inconsistency range. Although attitudinal items were excluded from the A-2 reinterview program, other reinterview programs have shown that, in general, attitudinal items such as items 102-104 (see facsimile of the 1977 AHS questionnaire, page App-34) have moderate to high levels of inconsistency. Moderate levels of inconsistency indicate that there is some problem with inconsistent reporting, and high levels indicate that improvements are needed in the method used to collect data or that the category concepts themselves are ambiguous.

A second measure of response error, the net difference rate, which is a measure of net error (or bias) in a particular category, was also provided by this reinterview program. None of the categories tested showed any significant evidence of bias.

The results of this study were based on sample data, so there is a sampling error associated with these estimates of nonsampling error. Therefore, the possibility of such errors should be taken into account when considering the results of this study.

A similar study was conducted for the 1974 AHS-SMSA sample. The results of which are presented in the following Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1974."

Coverage errors.—With respect to errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do

not necessarily represent missed housing units, since, due to the relatively short time span involved, it is possible that construction of these units was not completed at the time the survey was conducted. In which case, they would not have been eligible for interview. In addition to these deficiencies, new construction in special places that do not require building permits, such as military bases, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit. Also, additional coverage errors may exist in SMSA's where the complete Coverage Improvement Program was not applied (see the table on page App-45 of this appendix).

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all units located inside these ED's would be represented in the sample. However, it has been estimated that the 1974 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. Since these ED's were recanvassed for the 1977 survey, the number of missed housing units may be considerably less for 1977.

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 schedules, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by the standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.

2. Approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the average result of all possible samples.

3. Approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimate would include the average result of all possible samples.

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

The figures presented in the following tables 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 1974 housing inventory can be found in the AHS Series H-170 reports for 1974.

Table I presents the standard errors applicable to estimates of characteristics of the 1977 housing inventory as well as estimates of characteristics of the 1974-1977 lost units. 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 31,020 for the total SMSA, 14,940 for the central city of the SMSA, and 27,120 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 presents the standard errors of estimated percentages of the 1977 housing inventory as well as estimated percentages of the 1974-1977 lost units. 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.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1977 Housing Inventory and for Estimated Number of 1974-1977 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]		
	SMSA	In central city	Not in central city
0	190	200	190
100	190	200	190
200	200	200	200
500	310	320	310
700	370	380	370
1,000	440	450	440
2,500	700	710	700
5,000	980	1,000	980
10,000	1,390	1,400	1,390
25,000	2,180	2,150	2,180
50,000	3,050	2,900	3,050
75,000	3,700	3,360	3,700
100,000	4,220	3,640	4,220
150,000	5,050	3,830	5,050
200,000	5,690	3,540	5,690
250,000	6,210	—	6,210
300,000	6,620	—	6,620
400,000	7,210	—	7,210
500,000	7,550	—	7,500
600,000	7,670	—	7,670
700,000	7,580	—	7,580
800,000	7,270	—	—
900,000	6,710	—	—

[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 and the balance (not in central city) estimates and by 1.6 for the central city estimates.

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 1977 there were 552,400 owner-occupied housing units in this SMSA. Interpolation in table I of this appendix shows that the standard error of an estimate of this size is approximately 7,610. 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
500,000	7,550
552,400	x
600,000	7,670

The entry for x is determined as follows by vertically interpolating between 7,550 and 7,670:

$$552{,}400-500{,}000 = 52{,}400$$
$$600{,}000-500{,}000 = 100{,}000$$
$$7{,}550+\frac{52{,}400}{100{,}000}(7{,}670-7{,}550)$$
$$= 7{,}610$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 544,790 to 560,010 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1977 owner-occupied housing units lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 540,224 to 564,576 housing units with 90 percent confidence; and that the average estimate lies within the interval from 537,180 to 567,620 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 552,400 owner-occupied housing units, 160,500, or 29.1 percent, had two bedrooms. Interpolation in table II of the appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 29.1 percent is approximately 0.8 percentage points. The following interpolation procedure was used:

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

Base of percentage	Estimated percentage		
	25 or 75	29.1	50
500,000	0.9	a	1.0
552,400		p	
600,000	0.8	b	0.9

1. The entry for cell "a" is determined by horizontal interpolation between 0.9 and 1.0.

$$29.1-25.0 = 4.1$$
$$50.0-25.0 = 25.0$$

$$0.9+\frac{4.1}{25.0}(1.0-0.9) = 0.9$$

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

$$29.1-25.0 = 4.1$$
$$50.0-25.0 = 25.0$$

$$0.8+\frac{4.1}{25.0}(0.9-0.8) = 0.8$$

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 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)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	66.0	66.0	66.0	66.0	66.0	69.7
200	49.3	49.3	49.3	49.3	49.3	49.3
500	28.0	28.0	28.0	28.0	28.0	31.2
700	21.7	21.7	21.7	21.7	22.8	26.3
1,000	16.3	16.3	16.3	16.3	19.1	22.0
2,500	7.2	7.2	7.2	8.4	12.1	13.9
5,000	3.7	3.7	4.3	5.9	8.5	9.9
10,000	1.9	1.9	3.0	4.2	6.0	7.0
25,000	0.8	0.9	1.9	2.6	3.8	4.4
50,000	0.4	0.6	1.4	1.9	2.7	3.1
75,000	0.3	0.5	1.1	1.5	2.2	2.5
100,000	0.2	0.4	1.0	1.3	1.9	2.2
150,000	0.13	0.4	0.8	1.1	1.6	1.8
200,000	0.10	0.3	0.7	0.9	1.3	1.6
250,000	0.08	0.3	0.6	0.8	1.2	1.4
300,000	0.06	0.3	0.6	0.8	1.1	1.3
400,000	0.05	0.2	0.5	0.7	1.0	1.1
500,000	0.04	0.2	0.4	0.6	0.9	1.0
600,000	0.03	0.2	0.4	0.5	0.8	0.9
700,000	0.03	0.2	0.4	0.5	0.7	0.8
800,000	0.02	0.2	0.3	0.5	0.7	0.8
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 1.3 for the total SMSA and the balance (not in central city) estimates and by 1.6 for the central city estimates.

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

$$552{,}400 - 500{,}000 = 52{,}400$$
$$600{,}000 - 500{,}000 = 100{,}000$$

$$0.9 + \frac{52{,}400}{100{,}000}(0.8 - 0.9) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 28.3 to 29.9 percent; the 90-percent confidence interval is from 27.8 to 30.4 percent; and the 95-percent confidence interval is from 27.5 to 30.7 percent.

Differences.—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. However, 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 1974 and 1977 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1974 and 1977 characteristics.

Illustration of the computation of the standard error of a difference.—Table A-1 of part A of this report shows that in 1977 there were 280,200 owner-occupied units with three bedrooms in this SMSA. Thus, the apparent difference, shown by these data, between owner-occupied units with two bedrooms and owner-occupied units with three bedrooms is 119,700. Table I shows the standard error of 280,200 is approximately 6,460 and the standard error of 160,500 is approximately 5,180. Therefore, the standard error of the estimated difference of 119,700 is about:

$$8{,}280 = \sqrt{(6{,}460)^2 + (5{,}180)^2}$$

Consequently, the 68-percent confidence interval for the 119,700 difference is from 111,420 to 127,980 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of this difference lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 106,452 to 132,948 housing units, and the 95-percent confidence interval is from 103,140 to 136,260. Thus, we can conclude with 95 percent confidence that the number of 1977 owner-occupied 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, read off the confidence interval corresponding to the two points established in step 2.

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 552,400 housing units.

1. Interpolation in table II shows that the standard error of 50 percent on a base of 552,400 is approximately 0.9 percentage points.

2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.2 and 51.8.

3. From table A-1 of part A, it can be seen by cumulating the frequencies for the first two categories that 247,000 owner-occupied housing units, or 44.7 percent, had one or two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) and that an additional 106,700 owner-occupied housing units, or 19.3 percent, had three persons (i.e., 2.5 to 3.5 persons).

By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{48.2-44.7}{19.3}\right) = 2.7$$

Similarly, the upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{51.8-44.7}{19.3}\right) = 2.9$$

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

Saginaw 1977

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1977 survey App-42
Selection of the 1974 sample App-42
1977 sample reduction App-43
1974-1977 additions to the housing inventory App-43
Coverage improvement sample selection App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1970 Census of Population and Housing App-45
ESTIMATION App-45
1977 housing inventory App-45
1974-1977 lost units App-46
1974 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES App-46
Nonsampling errors App-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-50
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

SAMPLE DESIGN

Annual Housing Survey.—The estimates for each of the 20 SMSA's are based on data collected from the 1977 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 20 SMSA's, the data were collected for the 11-month period from April 1977 through February 1978 with one-eleventh of the sample units visited each month.

Each group of SMSA's is interviewed for the AHS on a rotating basis, and this group (A-2 SMSA's) was the first to be revisited. Nineteen of the 20 SMSA's were interviewed for the first time in 1974, while the Madison, Wis., SMSA was interviewed for the first time in 1975.

In each group of SMSA's, the largest SMSA from each of the 4 geographic regions is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All remaining SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the A-2 group (1977-1978) are: Boston, Mass.; Detroit, Mich.; Los Angeles-Long Beach, Calif.; and Washington, D.C.-Md.-Va.

The remaining SMSA's in the A-2 group are: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.

In this SMSA, 4,621 housing units were eligible for interview. Of these sample units, 242 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 units eligible for interview, 330 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 1977 survey.—The sample housing units designated for interview in the 1977 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 1974 survey which were not part of the 1977 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 1974 survey which were not part of the 1977 reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1977 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from the list of new construction building permits issued since the 1974 survey. (This sample represented the housing units built in permit-issuing areas, since the 1974 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1974 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1974 survey.)
5. All sample housing units that were selected as part of the 1977 Coverage Improvement Program. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1974 sample.—The sample for the SMSA's which are 100-percent permit-issuing was selected from two sample frames—units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which are not 100-percent permit-issuing included

a sample selected from a third frame—those units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). The following SMSA's are 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Los Angeles-Long Beach, Calif.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 14 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 four largest SMSA's, the overall sampling rate differed by central city and 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 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 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 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-$5,999. . .		
$6,000-$9,999. . .		
$10,000-$14,999 .		
$15,000 and over .		

Thus, for this SMSA, the occupied housing unit records from this 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 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 are 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 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 units in existence at the time of interview in these selected segments are in sample. Thus, units enumerated in the 1970 census as well as units built since the 1970 census are included.

1977 sample reduction.—In previous years, AHS sample units have been divided among 12 panels which were interviewed each month. Due to budget limitations, it became necessary to drop one of these panels from the 1977 survey. The designated sample size for 1977 was thus reduced by one-twelfth.

1974-1977 additions to the housing inventory.—In the permit-issuing universe, a sample of new construction building permits, issued since the 1974 survey, was

selected to represent housing units built in permit-issuing areas since the 1974 survey. Sampling procedures were identical to those used in selecting the 1970-1974 new construction sample. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1974, to identify any units missed in the 1974 survey or any units added since the 1974 survey.

Coverage improvement sample selection.—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 census.
3. Units missed in the 1970 census.
4. 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.

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 A-2 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 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: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Orlando, Fla.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; and Wichita, Kans.

In the remaining 10 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 and the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 218 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, for each 1977-1978 SMSA, 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 673 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 no 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 units in multi-unit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multi-unit structures

selected, all units were listed and matched to the 1970 census. Any missed units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply these two procedures in all 20 A-2 SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in three SMSA's. In these SMSA's, a string of 10 structures (instead of 8) that were eligible to be selected for the AHS, were listed. The full application of this technique was employed in 12 SMSA's, with 5 SMSA's not receiving any application of the first procedure. The second procedure was instituted in 13 SMSA's. The following table shows which of these two techniques were applied to the A-2 SMSA's.

The first procedure added an estimated 2,002 units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 43 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 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 1977 housing inventory) and estimates pertaining to characteristics of units removed from the housing inventory since 1974 (i.e., 1974-1977 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1977 housing inventory.—The AHS estimates of characteristics of the 1977 housing inventory employed a 1-stage ratio estimation procedure in 19 SMSA's, with a second-stage ratio estimation procedure being employed in 1 SMSA. Prior to the implementation of the ratio estimation procedure, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of 1 or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample 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 estima-

SMSA	Coverage improvement technique employed			
	First		Second	Neither
	Partial	Full		
Albany-Schenectady-Troy, N.Y	X		X	
Anaheim-Santa Ana-Garden Grove, Calif	X			
Boston, Mass		X	X	
Dallas, Tex				X
Detroit, Mich		X	X	
Fort Worth, Tex		X	X	
Los Angeles-Long Beach, Calif				X
Madison, Wis		X	X	
Memphis, Tenn.-Ark		X	X	
Minneapolis-St. Paul, Minn				X
Newark, N.J		X	X	
Orlando, Fla		X	X	
Phoenix, Ariz		X	X	
Pittsburgh, Pa				X
Saginaw, Mich		X	X	
Salt Lake City, Utah				X
Spokane, Wash	X			
Tacoma, Wash		X	X	
Washington, D.C.-Md.-Va		X	X	
Wichita, Kans		X	X	

tion factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of 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 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 unit within the corresponding ratio estimation category.

The effect of this ratio estimation procedure was to somewhat reduce the variance due to the variation in sample size of strata used in the sample selection of the permit-issuing universe, thereby reducing the sampling error below what would have been obtained by simply weighting the results of the sample by the inverse of the probability of selection. Ordinarily, this would have been controlled by sampling within the strata during the sample selection process. However, prior to the AHS sample selection within each SMSA, units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. Thus, some variation in sample size was introduced during the AHS sample selection process.

For the Pittsburgh, Pa., SMSA, a second-stage ratio estimation procedure was also employed. This procedure involved the ratio estimation of the AHS weighted sample estimate of the October 1977 housing inventory to an independent estimate of the SMSA's October 1977 housing inventory. This ratio estimate factor equaled the following:

$$\frac{\text{SMSA independent estimate of October 1977 housing inventory}}{\text{AHS sample estimate of the housing inventory}}$$

The numerator of the ratio was derived by using the 1970 census estimate of total housing units for the SMSA in conjunction with the estimate of the change in the housing inventory since the census. The estimate of change was based on administrative records from utility companies. The denominator of that ratio was obtained from the weighted estimate of AHS sample units, using the existing weight after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for each sample unit and the resulting product was used as the final weight for tabulation purposes.

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

The second-stage ratio estimation procedure was not employed in all A-2 SMSA's, since it was felt that coverage improvement procedures adequately corrected known deficiencies in the sample where full coverage improvement procedures were employed. The independent estimates were only employed in SMSA's where full coverage improvements were not used and the estimated relative bias of the independent estimate was thought to be low enough (i.e., 2 percent or less over a 10-year period) to preclude the possibility of doing more harm than good to the survey results. A measure of the relative bias of the independent estimate was available based on a comparison between the 1970 census estimate of the April 1970 housing inventory and the independent estimate of the April 1970 housing inventory which was projected over a 10-year period from the 1960 census using the same procedure that produced the October 1977 independent estimates.

1974-1977 lost units.—The 1974-1977 lost unit estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1974 housing inventory, as was described in the 1974 Current Housing Report, Series H-170, **Housing Characteristics for Selected Metropolitan Areas.** Since the 1974-1977 lost units existed, by definition, in the 1974 housing inventory, there was a 1974 housing inventory weight associated with each 1974-1977 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1974-1977 lost units.

1974 estimation procedure.—This report presents data on the housing characteristics of the 1974 housing inventory from the 1974 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1974.

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 to provide correct information on the part of respondents, mistakes in recording or coding the data, and other errors of collection, response, processing, coverage, and estimation for missing data. Nonsampling errors are not unique to sample surveys since they can, and do, occur in complete censuses as well.

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing, the 1977 AHS-SMSA sample, and the 1974 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 included space errors, definitional errors, and occupancy errors. The "content" errors measured the accuracy of the data collected for surveyed 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.—For the 1977 AHS-SMSA sample, a reinterview program was conducted to obtain a measurement of some of the components of the nonsampling error associated with the AHS estimates. A sample of households was revisited and answers to some of the questions on the AHS questionnaire were obtained again. The original interview and the reinterview were assumed to be two independent readings and, thus, were the basis for the measurement of the accuracy of the AHS data collected from interviewed households.

As part of the reinterview, a check was made at each of these households to determine if the following was done during the original interview:

1. The correct unit was visited.
2. The correct number of housing units were interviewed at that address.
3. The correct information on "Year Built" was obtained.
4. The correct information on "Tenure" was obtained.
5. The correct information on "Household Composition" was obtained.
6. The correct information on "Type of Housing Unit" was obtained.
7. The correct information on "Occupancy Status" was obtained.

This check was made for interviewer evaluation and control. That is, tolerance limits were derived to determine which interviewers passed or failed this reinterview check. The results of this study are available and are presented in the following Census Bureau memorandum, "Response Error in the Annual Housing Survey as Measured by Reinterviews—Year IV, 1977-1978 SMSA Sample."

For the A-2 reinterview program, measures of response inconsistency were taken on two subsamples—reconciled and unreconciled. After the question is answered in the reconciled subsample, the interviewer presented the original response and then asked the respondent to decide upon the best answer. Past experience has shown that the level of inconsistency is lower when the interviewer has the response available, and thus, it is believed that there is a downward bias in these estimates. In the unreconciled sample, the interviewer did not have access to previous responses. For the reconciled items, three-fourths showed low levels of response inconsistency while the remaining one-fourth were in the moderate range. For the unreconciled items, approximately one-half showed low levels of response disagreement, with the remaining one-half of the categories in the moderate and high inconsistency range. Although attitudinal items were excluded from the A-2 reinterview program, other reinterview programs have shown that, in general, attitudinal items such as items 102-104 (see facsimile of the 1977 AHS questionnaire, page App-34) have moderate to high levels of inconsistency. Moderate levels of inconsistency indicate that there is some problem with inconsistent reporting, and high levels indicate that improvements are needed in the method used to collect data or that the category concepts themselves are ambiguous.

A second measure of response error, the net difference rate, which is a measure of net error (or bias) in a particular category, was also provided by this reinterview program. None of the categories tested showed any significant evidence of bias.

The results of this study were based on sample data, so there is a sampling error associated with these estimates of nonsampling error. Therefore, the possibility of such errors should be taken into account when considering the results of this study.

A similar study was conducted for the 1974 AHS-SMSA sample. The results of which are presented in the following Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1974."

Coverage errors.—With respect to errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units, since, due to the relatively short time span involved, it is possible that

construction of these units was not completed at the time the survey was conducted. In which case, they would not have been eligible for interview. In addition to these deficiencies, new construction in special places that do not require building permits, such as military bases, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit. Also, additonal coverage errors may exist in SMSA's where the complete Coverage Improvement Program was not applied (see the table on page App-45 of this appendix).

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all units located inside these ED's would be represented in the sample. However, it has been estimated that the 1974 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. Since these ED's were recanvassed for the 1977 survey, the number of missed housing units may be considerably less for 1977.

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 schedules, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by the standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.

2. Approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the average result of all possible samples.

3. Approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimate would include the average result of all possible samples.

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

The figures presented in the following tables 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 1974 housing inventory can be found in the AHS Series H-170 reports for 1974.

Table I presents the standard errors applicable to estimates of characteristics

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1977 Housing Inventory and for Estimated Number of 1974-1977 Lost Units for the Saginaw, Mich., SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]	
	SMA and not in central city	Central city
0	20	20
100	40	40
200	60	60
500	90	90
700	110	100
1,000	130	120
2,500	200	190
5,000	280	260
10,000	400	350
25,000	600	430
50,000	790	—
75,000	880	—
100,000	900	—

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

of the 1977 housing inventory as well as estimates of characteristics of the 1974-1977 lost units. Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 3,620 for the total SMSA, 1,460 for the central city of the SMSA, and 3,200 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 presents the standard errors of estimated percentages of the 1977 housing inventory as well as estimated percentages of the 1974-77 lost units. 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 1977 there were 54,800 owner-occupied housing units in this SMSA. Interpolation in table I of this appendix shows that the standard error of an estimate of this size is approximately 810. 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	790
54,800	x
75,000	880

The entry for x is determined as follows by vertically interpolating between 790 and 880:

$$54{,}800-50{,}000 = 4{,}800$$
$$75{,}000-50{,}000 = 25{,}000$$

$$790+\frac{4{,}800}{25{,}000}(880-790) = 810$$

Consequently, the 68 percent confidence interval, as shown by these data, is from 53,990 to 55,610 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1977 owner-occupied housing units lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 53,504 to 56,096 housing units with 90 percent confidence; and that the average estimate lies within

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 Lost Units for the Saginaw, Mich., SMSA, for the Central City of the SMSA, and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	14.2	14.2	14.2	14.2	17.7	20.4
200	7.7	7.7	7.7	8.6	12.5	14.4
500	3.2	3.2	4.0	5.5	7.9	9.1
700	2.3	2.3	3.4	4.6	6.7	7.7
1,000	1.6	1.6	2.8	3.9	5.6	6.4
2,500	0.7	0.8	1.8	2.4	3.5	4.1
5,000	0.3	0.6	1.3	1.7	2.5	2.9
10,000	0.2	0.4	0.9	1.2	1.8	2.0
25,000	0.07	0.3	0.6	0.8	1.1	1.3
50,000	0.03	0.2	0.4	0.5	0.8	0.9
75,000	0.02	0.15	0.3	0.4	0.6	0.7
100,000	0.02	0.13	0.3	0.4	0.6	0.6

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

the interval from 53,180 to 56,420 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 54,800 owner-occupied housing units, 12,800, or 23.4 percent, had two bedrooms. Interpolation in table II of the appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 23.4 percent is approximately 1.6 percentage points. The following interpolation procedure was used:

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

Base of percentage	Estimated percentage		
	10 or 90	23.4	25 or 75
50,000	0.5	a	0.8
54,800		p	
75,000	0.4	b	0.6

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

$$23.4-10.0 = 13.4$$
$$25.0-10.0 = 15.0$$

$$0.5 + \frac{13.4}{15.0}(0.8-0.5) = 0.77$$

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

$$23.4-10.0 = 13.4$$
$$25.0-10.0 = 15.0$$

$$0.4 + \frac{13.4}{15.0}(0.6-0.4) = 0.58$$

3. The entry for "p" was then determined by vertical interpolation between 0.77 and 0.58.

$$54{,}800-50{,}000 = 4{,}800$$
$$75{,}000-50{,}000 = 25{,}000$$

$$0.77 + \frac{4{,}800}{25{,}000}(0.58-0.77) = 0.7$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 22.7 to 24.1 percent; the 90-percent confidence interval is from 22.3 to 24.5 percent; and the 95-percent confidence interval is from 22.0 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. However, 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 1974 and 1977 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1974 and 1977 characteristics.

Illustration of the computation of the standard error of a difference.—Table A-1 of part A of this report shows that in 1977 there were 30,500 owner-occupied units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied units with two bedrooms and owner-occupied units with three bedrooms is 17,700. Table I shows the standard error of 30,500 is approximately 640 and the standard error of 12,800 is approximately 440. Therefore, the standard error of the estimated difference of 17,700 is about:

$$780 = \sqrt{(640)^2 + (440)^2}$$

Consequently, the 68-percent confidence interval for the 17,700 difference is from 16,920 to 18,480 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 16,452 to 18,948 housing units, and the 95-percent confidence interval is from 16,140 to 19,260. Thus, we can conclude with 95 percent confidence that the number of 1977 owner-occupied 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, read off the confidence interval corresponding to the two points established in step 2.

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 3.2. The base of the distribution from which this median was determined is 54,800 housing units.

1. Interpolation in table II shows that the standard error of 50 percent on a base of 54,800 is approximately 0.9 percentage points.

2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.2 and 51.8.

3. From table A-1 of part A, it can be seen by cumulating the frequencies for the first two categories that 21,100 owner-occupied housing units, or 38.5 percent, had one or two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) and that an additional 9,200 owner-occupied housing units, or 16.8 percent, had three persons (i.e., 2.5 to 3.5 persons).

By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{48.2-38.5}{16.8}\right) = 3.1$$

Similarly, the upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{51.8-38.5}{16.8}\right) = 3.3$$

Thus, the 95-percent confidence interval ranges from 3.1 to 3.3 persons.

Salt Lake City 1977

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1977 survey App-42
Selection of the 1974 sample App-42
1977 sample reduction App-43
1974-1977 additions to the housing inventory App-43
Coverage improvement sample selection App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1970 Census of Population and Housing App-45
ESTIMATION App-45
1977 housing inventory App-45
1974-1977 lost units App-46
1974 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES App-46
Nonsampling errors App-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-50
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-51

SAMPLE DESIGN

Annual Housing Survey.—The estimates for each of the 20 SMSA's are based on data collected from the 1977 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 20 SMSA's, the data were collected for the 11-month period from April 1977 through February 1978 with one-eleventh of the sample units visited each month.

Each group of SMSA's is interviewed for the AHS on a rotating basis, and this group (A-2 SMSA's) was the first to be revisited. Nineteen of the 20 SMSA's were interviewed for the first time in 1974, while the Madison, Wis., SMSA was interviewed for the first time in 1975.

In each group of SMSA's, the largest SMSA from each of the 4 geographic regions is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All remaining SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the A-2 group (1977-1978) are: Boston, Mass.; Detroit, Mich.; Los Angeles-Long Beach, Calif.; and Washington, D.C.-Md.-Va.

The remaining SMSA's in the A-2 group are: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.

In this SMSA, 5,127 housing units were eligible for interview. Of these sample units, 121 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 units eligible for interview, 364 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 1977 survey.—The sample housing units designated for interview in the 1977 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 1974 survey which were not part of the 1977 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 1974 survey which were not part of the 1977 reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1977 AHS questionnaire, page App-20.)

3. All sample housing units that were selected from the list of new construction building permits issued since the 1974 survey. (This sample represented the housing units built in permit-issuing areas, since the 1974 survey.)

4. All sample housing units that were added to sample segments in the nonpermit universe since the 1974 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1974 survey.)

5. All sample housing units that were selected as part of the 1977 Coverage Improvement Program. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1974 sample.—The sample for the SMSA's which are 100-percent permit-issuing was selected from two sample frames—units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which are

not 100-percent permit-issuing included a sample selected from a third frame—those units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). The following SMSA's are 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Los Angeles-Long Beach, Calif.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 14 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 four largest SMSA's, the overall sampling rate differed by central city and 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 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 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 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-$5,999. . .		
$6,000-$9,999. . .		
$10,000-$14,999 .		
$15,000 and over .		

Thus, for this SMSA, the occupied housing unit records from this 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 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 are 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 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 units in existence at the time of interview in these selected segments are in sample. Thus, units enumerated in the 1970 census as well as units built since the 1970 census are included.

1977 sample reduction.—In previous years, AHS sample units have been divided among 12 panels which were interviewed each month. Due to budget limitations, it became necessary to drop one of these panels from the 1977 survey. The designated sample size for 1977 was thus reduced by one-twelfth.

1974-1977 additions to the housing inventory.—In the permit-issuing universe, a sample of new construction building

permits, issued since the 1974 survey, was selected to represent housing units built in permit-issuing areas since the 1974 survey. Sampling procedures were identical to those used in selecting the 1970-1974 new construction sample. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1974, to identify any units missed in the 1974 survey or any units added since the 1974 survey.

Coverage improvement sample selection.—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 census.

3. Units missed in the 1970 census.

4. 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.

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 A-2 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 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: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Orlando, Fla.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; and Wichita, Kans.

In the remaining 10 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 and the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 757 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, for each 1977-1978 SMSA, 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 1,697 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 no 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 units in multi-unit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multi-unit structures selected, all units were listed and matched to the 1970 census. Any missed units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply these two procedures in all 20 A-2 SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in three SMSA's. In these SMSA's, a string of 10 structures (instead of 8) that were eligible to be selected for the AHS, were listed. The full application of this technique was employed in 12 SMSA's, with 5 SMSA's not receiving any application of the first procedure. The second procedure was instituted in 13 SMSA's. The following table shows which of these two techniques were applied to the A-2 SMSA's.

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 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 1977 housing inventory) and estimates pertaining to characteristics of units removed from the housing inventory since 1974 (i.e., 1974-1977 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1977 housing inventory.—The AHS estimates of characteristics of the 1977 housing inventory employed a 1-stage ratio estimation procedure in 19 SMSA's, with a second-stage ratio estimation procedure being employed in 1 SMSA. Prior to the implementation of the ratio estimation procedure, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

SMSA	Coverage improvement technique employed			
	First		Second	Neither
	Partial	Full		
Albany-Schenectady-Troy, N.Y	X		X	
Anaheim-Santa Ana-Garden Grove, Calif	X			
Boston, Mass		X	X	
Dallas, Tex				X
Detroit, Mich		X	X	
Fort Worth, Tex		X	X	
Los Angeles-Long Beach, Calif				X
Madison, Wis		X	X	
Memphis, Tenn.-Ark		X	X	
Minneapolis-St. Paul, Minn				X
Newark, N.J		X	X	
Orlando, Fla		X	X	
Phoenix, Ariz		X	X	
Pittsburgh, Pa				X
Saginaw, Mich		X	X	
Salt Lake City, Utah				X
Spokane, Wash	X			
Tacoma, Wash		X	X	
Washington, D.C.-Md.-Va		X	X	
Wichita, Kans		X	X	

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 1 or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample 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 estima-

tion factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of 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 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 unit within the corresponding ratio estimation category.

The effect of this ratio estimation procedure was to somewhat reduce the variance due to the variation in sample size of strata used in the sample selection of the permit-issuing universe, thereby reducing the sampling error below what would have been obtained by simply weighting the results of the sample by the inverse of the probability of selection. Ordinarily, this would have been controlled by sampling within the strata during the sample selection process. However, prior to the AHS sample selection within each SMSA, units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. Thus, some variation in sample size was introduced during the AHS sample selection process.

For the Pittsburgh, Pa., SMSA, a second-stage ratio estimation procedure was also employed. This procedure involved the ratio estimation of the AHS weighted sample estimate of the October 1977 housing inventory to an independent estimate of the SMSA's October 1977 housing inventory. This ratio estimate factor equaled the following:

$$\frac{\text{SMSA independent estimate of October 1977 housing inventory}}{\text{AHS sample estimate of the housing inventory}}$$

The numerator of the ratio was derived by using the 1970 census estimate of total housing units for the SMSA in conjunction with the estimate of the change in the housing inventory since the census. The estimate of change was based on administrative records from utility companies. The denominator of that ratio was obtained from the weighted estimate of AHS sample units, using the existing weight after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for each sample unit and the resulting product was used as the final weight for tabulation purposes.

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

The second-stage ratio estimation procedure was not employed in all A-2 SMSA's, since it was felt that coverage improvement procedures adequately corrected known deficiencies in the sample where full coverage improvement procedures were employed. The independent estimates were only employed in SMSA's where full coverage improvements were not used and the estimated relative bias of the independent estimate was thought to be low enough (i.e., 2 percent or less over a 10-year period) to preclude the possibility of doing more harm than good to the survey results. A measure of the relative bias of the independent estimate was available based on a comparison between the 1970 census estimate of the April 1970 housing inventory and the independent estimate of the April 1970 housing inventory which was projected over a 10-year period from the 1960 census using the same procedure that produced the October 1977 independent estimates.

1974-1977 lost units.—The 1974-1977 lost unit estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1974 housing inventory, as was described in the 1974 Current Housing Report, Series H-170, **Housing Characteristics for Selected Metropolitan Areas.** Since the 1974-1977 lost units existed, by definition, in the 1974 housing inventory, there was a 1974 housing inventory weight associated with each 1974-1977 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1974-1977 lost units.

1974 estimation procedure.—This report presents data on the housing characteristics of the 1974 housing inventory from the 1974 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1974.

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 to provide correct information on the part of respondents, mistakes in recording or coding the data, and other errors of collection, response, processing, coverage, and estimation for missing data. Nonsampling errors are not unique to sample surveys since they can, and do, occur in complete censuses as well.

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing, the 1977 AHS-SMSA sample, and the 1974 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 included space errors, definitional errors, and occupancy errors. The "content" errors measured the accuracy of the data collected for surveyed 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.—For the 1977 AHS-SMSA sample, a reinterview program was conducted to obtain a measurement of some of the components of the nonsampling error associated with the AHS estimates. A sample of households was revisited and answers to some of the questions on the AHS questionnaire were obtained again. The original interview and the reinterview were assumed to be two independent readings and, thus, were the basis for the measurement of the accuracy of the AHS data collected from interviewed households.

As part of the reinterview, a check was made at each of these households to determine if the following was done during the original interview:

1. The correct unit was visited.
2. The correct number of housing units were interviewed at that address.
3. The correct information on "Year Built" was obtained.
4. The correct information on "Tenure" was obtained.
5. The correct information on "Household Composition" was obtained.
6. The correct information on "Type of Housing Unit" was obtained.
7. The correct information on "Occupancy Status" was obtained.

This check was made for interviewer evaluation and control. That is, tolerance limits were derived to determine which interviewers passed or failed this reinterview check. The results of this study are available and are presented in the following Census Bureau memorandum, "Response Error in the Annual Housing Survey as Measured by Reinterviews—Year IV, 1977-1978 SMSA Sample."

For the A-2 reinterview program, measures of response inconsistency were taken on two subsamples—reconciled and unreconciled. After the question is answered in the reconciled subsample, the interviewer presented the original response and then asked the respondent to decide upon the best answer. Past experience has shown that the level of inconsistency is lower when the interviewer has the response available, and thus, it is believed that there is a downward bias in these estimates. In the unreconciled sample, the interviewer did not have access to previous responses. For the reconciled items, three-fourths showed low levels of response inconsistency while the remaining one-fourth were in the moderate range. For the unreconciled items, approximately one-half showed low levels of response disagreement, with the remaining one-half of the categories in the moderate and high inconsistency range. Although attitudinal items were excluded from the A-2 reinterview program, other reinterview programs have shown that, in general, attitudinal items such as items 102-104 (see facsimile of the 1977 AHS questionnaire, page App-34) have moderate to high levels of inconsistency. Moderate levels of inconsistency indicate that there is some problem with inconsistent reporting, and high levels indicate that improvements are needed in the method used to collect data or that the category concepts themselves are ambiguous.

A second measure of response error, the net difference rate, which is a measure of net error (or bias) in a particular category, was also provided by this reinterview program. None of the categories tested showed any significant evidence of bias.

The results of this study were based on sample data, so there is a sampling error associated with these estimates of nonsampling error. Therefore, the possibility of such errors should be taken into account when considering the results of this study.

A similar study was conducted for the 1974 AHS-SMSA sample. The results of which are presented in the following Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1974."

Coverage errors.—With respect to errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units, since, due to the relatively short

time span involved, it is possible that construction of these units was not completed at the time the survey was conducted. In which case, they would not have been eligible for interview. In addition to these deficiencies, new construction in special places that do not require building permits, such as military bases, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit. Also, additional coverage errors may exist in SMSA's where the complete Coverage Improvement Program was not applied (see the table on page App-45 of this appendix).

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all units located inside these ED's would be represented in the sample. However, it has been estimated that the 1974 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. Since these ED's were recanvassed for the 1977 survey, the number of missed housing units may be considerably less for 1977.

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 schedules, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by the standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.

2. Approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the average result of all possible samples.

3. Approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimate would include the average result of all possible samples.

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

The figures presented in the following tables 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 1974

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1977 Housing Inventory and for Estimated Number of 1974-1977 Lost Units for the Salt Lake City, Utah, SMSA, for the Central City of the SMSA, and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]	
	SMSA and in central city	Not in central city
0	50	50
100	70	70
200	100	100
500	150	150
700	180	180
1,000	210	220
2,500	340	340
5,000	480	480
10,000	670	670
25,000	1,040	1,030
50,000	1,430	1,390
75,000	1,690	1,600
100,000	1,870	1,720
150,000	2,100	1,770
200,000	2,180	–
250,000	2,130	–

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

housing inventory can be found in the AHS Series H-170 reports for 1974.

Table I presents the standard errors applicable to estimates of characteristics of the 1977 housing inventory as well as estimates of characteristics of the 1974-1977 lost units. Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 8,420 for the total SMSA, 3,630 for the central city of the SMSA, and 7,510 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 presents the standard errors of estimated percentages of the 1977 housing inventory as well as estimated percentages of the 1974-1977 lost units. 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 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:

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

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 Lost Units for the Salt Lake City, Utah, SMSA, for the Central City of the SMSA, and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100.	31.6	31.6	31.6	31.6	31.6	34.0
200.	18.8	18.8	18.8	18.8	20.8	24.1
500.	8.5	8.5	8.5	9.1	13.2	15.2
700.	6.2	6.2	6.2	7.7	11.1	12.9
1,000	4.4	4.4	4.7	6.5	9.3	10.8
2,500	1.8	1.8	3.0	4.1	5.9	6.8
5,000	0.9	1.0	2.1	2.9	4.2	4.8
10,000	0.5	0.7	1.5	2.0	2.9	3.4
25,000	0.2	0.4	0.9	1.3	1.9	2.2
50,000	0.09	0.3	0.7	0.9	1.3	1.5
75,000	0.06	0.2	0.5	0.7	1.1	1.2
100,000. . . .	0.05	0.2	0.5	0.6	0.9	1.1
150,000. . . .	0.03	0.2	0.4	0.5	0.8	0.9
200,000. . . .	0.02	0.15	0.3	0.5	0.7	0.8
250,000. . . .	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, for the central city, and for the balance (not in central city) estimates.

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 1977 there were 142,600 owner-occupied housing units in this SMSA. Interpolation in table I of this appendix shows that the standard error of an estimate of this size is approximately 2,070. The following interpolation procedure was used:

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

Size of estimate	Standard error
100,000.	1,870
142,600.	x
150,000.	2,100

The entry for x is determined as follows by vertically interpolating between 1,870 and 2,100:

$$142{,}600-100{,}000 = 42{,}600$$
$$150{,}000-100{,}000 = 50{,}000$$

$$1{,}870+\frac{42{,}600}{50{,}000}(2{,}100-1{,}870) = 2{,}070$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 140,530 to 144,670 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples,

of 1977 owner-occupied housing units lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 139,288 to 145,912 housing units with 90 percent confidence; and that the average estimate lies within the interval from 138,460 to 146,740 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 142,600 owner-occupied housing units, 36,600, or 25.7 percent, had two bedrooms. Interpolation in table II of the appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 25.7 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	25.7	50
100,000. . . .	0.9	a	1.1
142,600. . . .		p	
150,000. . . .	0.8	b	0.9

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

$$25.7-25.0 = 0.7$$
$$50.0-25.0 = 25.0$$

$$0.9 + \frac{0.7}{25.0}(1.1-0.9) = 0.91$$

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

$$25.7-25.0 = 0.7$$
$$50.0-25.0 = 25.0$$

$$0.8 + \frac{0.7}{25.0}(0.9-0.8) = 0.80$$

3. The entry for "p" was then determined by vertical interpolation between 0.91 and 0.80.

$$142{,}600-100{,}000 = 42{,}600$$
$$150{,}000-100{,}000 = 50{,}000$$

$$0.91 + \frac{42{,}600}{50{,}000}(0.80-0.90) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 24.9 to 26.5 percent; the 90-percent confidence interval is from 24.4 to 27.0 percent; and the 95-percent confidence interval is from 24.1 to 27.3 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. However, 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 1974 and 1977 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1974 and 1977 characteristics.

Illustration of the computation of the standard error of a difference.—Table A-1 of part A of this report shows that in 1977 there were 60,100 owner-occupied units with three bedrooms in this SMSA, Thus, the apparent difference, as shown by these data, between owner-occupied units with two bedrooms and owner-occupied units with three bedrooms is 23,500. Table I shows the standard error of 60,100 is approximately 1,540 and the standard error of 36,600 is approximately 1,220. Therefore, the standard error of the estimated difference of 23,500 is about:

$$1{,}960 = \sqrt{(1{,}540)^2 + (1{,}220)^2}$$

Consequently, the 68-percent confidence interval for the 23,500 difference is from 21,540 to 25,460 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of this difference lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 20,364 to 26,636 housing units, and the 95-percent confidence interval is from 19,580 to 27,420. Thus, we can conclude with 95 percent confidence that the number of 1977 owner-occupied 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, read off the confidence interval corresponding to the two points established in step 2.

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

A two-standard error confidence interval may be determined by finding the values corresponding to 50 percent plus and minus twice the standard error determined in step 1. For about 95 out of

100 possible samples, the average median from all possible samples would lie between these two values.

Illustration of the computation of the 95-percent confidence interval of a median.—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 3.3. The base of the distribution from which this median was determined is 142,600 housing units.

1. Interpolation in table II shows that the standard error of 50 percent on a base of 142,600 is approximately 0.9 percentage points.

2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.2 and 51.8.

3. From table A-1 of part A, it can be seen by cumulating the frequencies for the first two categories that 51,700 owner-occupied housing units, or 36.3 percent, had one or two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) and that an additional 24,300 owner-occupied housing units, or 17.0 percent had three persons (i.e., 2.5 to 3.5 persons).

By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{48.2-36.3}{17.0}\right) = 3.2$$

Similarly, the upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{51.8-36.3}{17.0}\right) = 3.4$$

Thus, the 95-percent confidence interval ranges from 3.2 to 3.4 persons.

Spokane 1977

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1977 survey App-42
Selection of the 1974 sample App-42
1977 sample reduction App-43
1974-1977 additions to the housing inventory App-43
Coverage improvement sample selection App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1970 Census of Population and Housing App-45
ESTIMATION App-45
1977 housing inventory App-45
1974-1977 lost units App-46
1974 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES App-46
Nonsampling errors App-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-50
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

SAMPLE DESIGN

Annual Housing Survey.—The estimates for each of the 20 SMSA's are based on data collected from the 1977 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 20 SMSA's, the data were collected for the 11-month period from April 1977 through February 1978 with one-eleventh of the sample units visited each month.

Each group of SMSA's is interviewed for the AHS on a rotating basis, and this group (A-2 SMSA's) was the first to be revisited. Nineteen of the 20 SMSA's were interviewed for the first time in 1974, while the Madison, Wis., SMSA was interviewed for the first time in 1975.

In each group of SMSA's, the largest SMSA from each of the 4 geographic regions is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All remaining SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the A-2 group (1977-1978) are: Boston, Mass.; Detroit, Mich.; Los Angeles-Long Beach, Calif.; and Washington, D.C.-Md.-Va.

The remaining SMSA's in the A-2 group are: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.

In this SMSA, 4,868 housing units were eligible for interview. Of these sample units, 213 interviews were not obtained because, for occupied sample units, the occupants refused to be interviewed, were not at home after repeated visits or were unavailable for some other reason; or, for vacant units, no informed respondent could be found after repeated visits. In addition to units eligible for interview, 345 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 1977 survey.—The sample housing units designated for interview in the 1977 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 1974 survey which were not part of the 1977 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 1974 survey which were not part of the 1977 reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1977 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from the list of new construction building permits issued since the 1974 survey. (This sample represented the housing units built in permit-issuing areas since the 1974 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1974 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1974 survey.)
5. All sample housing units that were selected as part of the 1977 Coverage Improvement Program. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1974 sample.—The sample for the SMSA's which are 100-percent permit-issuing was selected from two sample frames—units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which are not 100-percent permit-issuing included

a sample selected from a third frame—those units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). The following SMSA's are 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Los Angeles-Long Beach, Calif.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 14 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 four largest SMSA's, the overall sampling rate differed by central city and 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 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 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 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-$5,999. . .		
$6,000-$9,999. . .		
$10,000-$14,999 .		
$15,000 and over .		

Thus, for this SMSA, the occupied housing unit records from this 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 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 are 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 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 units in existence at the time of interview in these selected segments are in sample. Thus, units enumerated in the 1970 census as well as units built since the 1970 census are included.

1977 sample reduction.—In previous years, AHS sample units have been divided among 12 panels which were interviewed each month. Due to budget limitations, it became necessary to drop one of these panels from the 1977 survey. The designated sample size for 1977 was thus reduced by one-twelfth.

1974-1977 additions to the housing inventory.—In the permit-issuing universe, a sample of new construction building permits, issued since the 1974 survey, was selected to represent housing units built in permit-issuing areas since the 1974

survey. Sampling procedures were identical to those used in selecting the 1970-1974 new construction sample. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1974, to identify any units missed in the 1974 survey or any units added since the 1974 survey.

Coverage improvement sample selection.—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 census.
3. Units missed in the 1970 census.
4. 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.

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 A-2 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 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: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Orlando, Fla.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; and Wichita, Kans.

In the remaining 10 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 and the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 497 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, for each 1977-1978 SMSA, 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 443 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 no 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 units in multi-unit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multi-unit structures selected, all units were listed and matched

to the 1970 census. Any missed units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply these two procedures in all 20 A-2 SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in three SMSA's. In these SMSA's, a string of 10 structures (instead of 8) that were eligible to be selected for the AHS, were listed. The full application of this technique was employed in 12 SMSA's, with 5 SMSA's not receiving any application of the first procedure. The second procedure was instituted in 13 SMSA's. The following table shows which of these two techniques were applied to the A-2 SMSA's.

The first procedure added an estimated 666 units to the coverage of the housing inventory of this SMSA, while the second procedure was not done in Spokane.

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 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 1977 housing inventory) and estimates pertaining to characteristics of units removed from the housing inventory since 1974 (i.e., 1974-1977 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1977 housing inventory.—The AHS estimates of characteristics of the 1977 housing inventory employed a 1-stage ratio estimation procedure in 19 SMSA's, with a second-stage ratio estimation procedure being employed in 1 SMSA. Prior to the implementation of the ratio estimation procedure, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of 1 or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample 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 esti-

SMSA	Coverage improvement technique employed			
	First		Second	Neither
	Partial	Full		
Albany-Schenectady-Troy, N.Y	X		X	
Anaheim-Santa Ana-Garden Grove, Calif	X			
Boston, Mass		X	X	
Dallas, Tex				X
Detroit, Mich		X	X	
Fort Worth, Tex		X	X	
Los Angeles-Long Beach, Calif				X
Madison, Wis		X	X	
Memphis, Tenn.-Ark		X	X	
Minneapolis-St. Paul, Minn				X
Newark, N.J		X	X	
Orlando, Fla		X	X	
Phoenix, Ariz		X	X	
Pittsburgh, Pa				X
Saginaw, Mich		X	X	
Salt Lake City, Utah				X
Spokane, Wash	X			
Tacoma, Wash		X	X	
Washington, D.C.-Md.-Va		X	X	
Wichita, Kans		X	X	

mation factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of 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 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 unit within the corresponding ratio estimation category.

The effect of this ratio estimation procedure was to somewhat reduce the variance due to the variation in sample size of strata used in the sample selection of the permit-issuing universe, thereby reducing the sampling error below what would have been obtained by simply weighting the results of the sample by the inverse of the probability of selection. Ordinarily, this would have been controlled by sampling within the strata during the sample selection process. However, prior to the AHS sample selection within each SMSA, units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. Thus, some variation in sample size was introduced during the AHS sample selection process.

For the Pittsburgh, Pa., SMSA, a second-stage ratio estimation procedure was also employed. This procedure involved the ratio estimation of the AHS weighted sample estimate of the October 1977 housing inventory to an independent estimate of the SMSA's October 1977 housing inventory. This ratio estimate factor equaled the following:

$$\frac{\text{SMSA independent estimate of October 1977 housing inventory}}{\text{AHS sample estimate of the housing inventory}}$$

The numerator of the ratio was derived by using the 1970 census estimate of total housing units for the SMSA in conjunction with the estimate of the change in the housing inventory since the census. The estimate of change was based on administrative records from utility companies. The denominator of that ratio was obtained from the weighted estimate of AHS sample units, using the existing weight after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for each sample unit and the resulting product was used as the final weight for tabulation purposes.

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

The second-stage ratio estimation procedure was not employed in all A-2 SMSA's, since it was felt that coverage improvement procedures adequately corrected known deficiencies in the sample where full coverage improvement procedures were employed. The independent estimates were only employed in SMSA's where full coverage improvements were not used and the estimated relative bias of the independent estimate was thought to be low enough (i.e., 2 percent or less over a 10-year period) to preclude the possibility of doing more harm than good to the survey results. A measure of the relative bias of the independent estimate was available based on a comparison between the 1970 census estimate of the April 1970 housing inventory and the independent estimate of the April 1970 housing inventory which was projected over a 10-year period from the 1960 census using the same procedure that produced the October 1977 independent estimates.

1974-1977 lost units.—The 1974-1977 lost unit estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1974 housing inventory, as was described in the 1974 Current Housing Report, Series H-170, **Housing Characteristics for Selected Metropolitan Areas.** Since the 1974-1977 lost units existed, by definition, in the 1974 housing inventory, there was a 1974 housing inventory weight associated with each 1974-1977 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1974-1977 lost units.

1974 estimation procedure.—This report presents data on the housing characteristics of the 1974 housing inventory from the 1974 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1974.

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 to provide correct information on the part of respondents, mistakes in recording or coding the data, and other errors of collection, response, processing, coverage, and estimation for missing data. Nonsampling errors are not unique to sample surveys since they can, and do, occur in complete censuses as well.

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing, the 1977 AHS-SMSA sample, and the 1974 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 included space errors, definitional errors, and occupancy errors. The "content" errors measured the accuracy of the data collected for surveyed 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.—For the 1977 AHS-SMSA sample, a reinterview program was conducted to obtain a measurement of some of the components of the nonsampling error associated with the AHS estimates. A sample of households was revisited and answers to some of the questions on the AHS questionnaire were obtained again. The original interview and the reinterview were assumed to be two independent readings and, thus, were the basis for the measurement of the accuracy of the AHS data collected from interviewed households.

As part of the reinterview, a check was made at each of these households to determine if the following was done during the original interview:

1. The correct unit was visited.
2. The correct number of housing units were interviewed at that address.
3. The correct information on "Year Built" was obtained.
4. The correct information on "Tenure" was obtained.
5. The correct information on "Household Composition" was obtained.
6. The correct information on "Type of Housing Unit" was obtained.
7. The correct information on "Occupancy Status" was obtained.

This check was made for interviewer evaluation and control. That is, tolerance limits were derived to determine which interviewers passed or failed this reinterview check. The results of this study are available and are presented in the following Census Bureau memorandum, "Response Error in the Annual Housing Survey as Measured by Reinterviews—Year IV, 1977-1978 SMSA Sample."

For the A-2 reinterview program, measures of response inconsistency were taken on two subsamples—reconciled and unreconciled. After the question is answered in the reconciled subsample, the interviewer presented the original response and then asked the respondent to decide upon the best answer. Past experience has shown that the level of inconsistency is lower when the interviewer has the response available, and thus, it is believed that there is a downward bias in these estimates. In the unreconciled sample, the interviewer did not have access to previous responses. For the reconciled items, three-fourths showed low levels of response inconsistency while the remaining one-fourth were in the moderate range. For the unreconciled items, approximately one-half showed low levels of response disagreement with the remaining one-half of the categories in the moderate and high inconsistency range. Although attitudinal items were excluded from the A-2 reinterview program, other reinterview programs have shown that, in general, attitudinal items such as items 102-104 (see facsimile of the 1977 AHS questionnaire, page App-34) have moderate to high levels of inconsistency. Moderate levels of inconsistency indicate that there is some problem with inconsistent reporting, and high levels indicate that improvements are needed in the method used to collect data or that the category concepts themselves are ambiguous.

A second measure of response error, the net difference rate, which is a measure of net error (or bias) in a particular category, was also provided by this reinterview program. None of the categories tested showed any significant evidence of bias.

The results of this study were based on sample data, so there is a sampling error associated with these estimates of nonsampling error. Therefore, the possibility of such errors should be taken into account when considering the results of this study.

A similar study was conducted for the 1974 AHS-SMSA sample. The results of which are presented in the following Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1974."

Coverage errors.—With respect to errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units, since, due to the relatively short time span involved, it is possible that construction of these units was not com-

pleted at the time the survey was conducted. In which case, they would not have been eligible for interview. In addition to these deficiencies, new construction in special places that do not require building permits, such as military bases, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit. Also, additional coverage errors may exist in SMSA's where the complete Coverage Improvement Program was not applied (see the table on page App-45 of this appendix).

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all units located inside these ED's would be represented in the sample. However, it has been estimated that the 1974 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. Since these ED's were recanvassed for the 1977 survey, the number of missed housing units may be considerably less for 1977.

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 schedules, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by the standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

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

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

The figures presented in the following tables 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 1974 housing inventory can be found in the AHS Series H-170 reports for 1974.

Table I presents the standard errors applicable to estimates of characteristics of the 1977 housing inventory as well as estimates of characteristics of the

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1977 Housing Inventory and for Estimated Number of 1974-1977 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]	
	SMSA and not in central city	Central city
0	30	20
100	50	50
200	70	70
500	110	110
700	130	130
1,000	160	160
2,500	250	250
5,000	360	350
10,000	500	480
25,000	780	720
50,000	1,060	890
75,000	1,250	930
100,000	1,390	–
150,000	1,560	–

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

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 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)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100.	20.4	20.4	20.4	20.4	21.9	25.3
200.	11.3	11.3	11.3	11.3	15.5	17.9
500.	4.9	4.9	4.9	6.8	9.8	11.3
700.	3.5	3.5	4.2	5.7	8.3	9.6
1,000	2.5	2.5	3.5	4.8	6.9	8.0
2,500	1.0	1.0	2.2	3.0	4.4	5.1
5,000	0.5	0.7	1.6	2.1	3.1	3.6
10,000	0.3	0.5	1.1	1.5	2.2	2.5
25,000	0.10	0.3	0.7	1.0	1.4	1.6
50,000	0.05	0.2	0.5	0.7	1.0	1.1
75,000	0.03	0.2	0.4	0.6	0.8	0.9
100,000. . . .	0.03	0.2	0.3	0.5	0.7	0.8
150,000. . . .	0.02	0.13	0.3	0.4	0.6	0.7

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

1974-1977 lost units. 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 4,610 for the total SMSA, 2,600 for the central city of the SMSA, and 3,730 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 presents the standard errors of estimated percentages of the 1977 housing inventory as well as estimated percentages of the 1974-1977 lost units. 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 1977 there were 77,300 owner-occupied housing units in this SMSA. Interpolation in 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
75,000	1,250
77,300	x
100,000.	1,390

The entry for x is determined as follows by vertically interpolating between 1,250 and 1,390:

$$77{,}300-75{,}000 = 2{,}300$$
$$100{,}000-75{,}000 = 25{,}000$$

$$1{,}250+\frac{2{,}300}{25{,}000}(1{,}390-1{,}250) = 1{,}260$$

Consequently, the 68 percent confidence interval, as shown by these data, is from 76,040 to 78,560 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1977 owner-occupied housing units lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 75,284 to 79,316 housing units with 90 percent confidence; and that the average estimate lies within the interval from 74,780 to

79,820 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 77,300 owner-occupied housing units, 20,600, or 26.6 percent, had two bedrooms. Interpolation in table II of the appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 26.6 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	26.6	50
75,000	0.8	a	0.9
77,300		p	
100,000. . . .	0.7	b	0.8

1. The entry for cell "a" is determined by horizontal interpolation between 0.8 and 0.9.

$$26.6-25.0 = 1.6$$
$$50.0-25.0 = 25.0$$

$$0.8 + \frac{1.6}{25.0}(0.9-0.8) = .81$$

2. The entry for cell "b" is determined by horizontal interpolation between 0.7 and 0.8.

$$26.6-25.0 = 1.6$$
$$50.0-25.0 = 25.0$$

$$0.7 + \frac{1.6}{25.0}(0.8-0.7) = 0.71$$

3. The entry for "p" was then determined by vertical interpolation between 0.81 and 0.71.

$$77{,}300-75{,}000 = 2{,}300$$
$$100{,}000-75{,}000 = 25{,}000$$

$$0.81 + \frac{2{,}300}{25{,}000}(0.71-0.81) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 25.8 to 27.4 percent; the 90-percent confidence interval is from 25.3 to 27.9 percent; and the 95-percent confidence interval is from 25.0 to 28.2 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. However, 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 1974 and 1977 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1974 and 1977 characteristics.

Illustration of the computation of the standard error of a difference.—Table A-1 of part A of this report shows that in 1977 there were 32,500 owner-occupied units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied units with two bedrooms and owner-occupied units with three bedrooms is 11,900. Table I shows the standard error of 32,500 is approximately 860, and the standard error of 20,600 is approximately 700. Therefore, the standard error of the estimated difference of 11,900 is about:

$$1{,}110 = \sqrt{(860)^2 + (700)^2}$$

Consequently, the 68-percent confidence interval for the 11,900 difference is from 10,790 to 13,010 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 10,124 to 13,676 housing units, and the 95-percent confidence interval is from 9,680 to 14,120. Thus, we can conclude with 95 percent confidence that the number of 1977 owner-occupied 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, read off the confidence interval corresponding to the two points established in step 2.

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 77,300 housing units.

1. Interpolation in table II shows that the standard error of 50 percent on a base of 77,300 is approximately 0.9 percentage points.

2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.2 and 51.8.

3. From table A-1 of part A, it can be seen by cumulating the frequencies for the first two categories that 37,800 owner-occupied housing units, or 48.9 percent, had one or two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) and that an additional 13,300 owner-occupied housing units, or 17.2 percent, had three persons (i.e., 2.5 to 3.5 persons).

By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{48.2-48.9}{17.2}\right) = 2.5$$

Similarly, the upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{51.8-48.9}{17.2}\right) = 2.7$$

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

Tacoma 1977

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1977 survey App-42
Selection of the 1974 sample App-42
1977 sample reduction App-43
1974-1977 additions to the housing inventory App-43
Coverage improvement sample selection App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1970 Census of Population and Housing App-45
ESTIMATION App-45
1977 housing inventory App-45
1974-1977 lost units App-46
1974 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES App-46
Nonsampling errors App-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-50
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

SAMPLE DESIGN

Annual Housing Survey.—The estimates for each of the 20 SMSA's are based on data collected from the 1977 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 20 SMSA's, the data were collected for the 11-month period from April 1977 through February 1978 with one-eleventh of the sample units visited each month.

Each group of SMSA's is interviewed for the AHS on a rotating basis, and this group (A-2 SMSA's) was the first to be revisited. Nineteen of the 20 SMSA's were interviewed for the first time in 1974, while the Madison, Wis., SMSA was interviewed for the first time in 1975.

In each group of SMSA's, the largest SMSA from each of the 4 geographic regions is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All remaining SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the A-2 group (1977-1978) are: Boston, Mass.; Detroit, Mich.; Los Angeles-Long Beach, Calif.; and Washington, D.C.-Md.-Va.

The remaining SMSA's in the A-2 group are: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash., Tacoma, Wash.; and Wichita, Kans.

In this SMSA, 5,061 housing units were eligible for interview. Of these sample units, 247 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 units eligible for interview, 328 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 1977 survey.—The sample housing units designated for interview in the 1977 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 1974 survey which were not part of the 1977 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 1974 survey which were not part of the 1977 reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1977 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from the list of new construction building permits issued since the 1974 survey. (This sample represented the housing units built in permit-issuing areas, since the 1974 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1974 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1974 survey.)
5. All sample housing units that were selected as part of the 1977 Coverage Improvement Program. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1974 sample.—The sample for the SMSA's which are 100-percent permit-issuing was selected from two sample frames—units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which are not 100-percent permit-issuing included

a sample selected from a third frame—those units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). The following SMSA's are 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Los Angeles-Long Beach, Calif.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 14 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 four largest SMSA's, the overall sampling rate differed by central city and 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 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 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 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-$5,999. . .		
$6,000-$9,999. . .		
$10,000-$14,999 .		
$15,000 and over .		

Thus, for this SMSA, the occupied housing unit records from this 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 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 are 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 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 units in existence at the time of interview in these selected segments are in sample. Thus, units enumerated in the 1970 census as well as units built since the 1970 census are included.

1977 sample reduction.—In previous years, AHS sample units have been divided among 12 panels which were interviewed each month. Due to budget limitations, it became necessary to drop one of these panels from the 1977 survey. The designated sample size for 1977 was thus reduced by one-twelfth.

1974-1977 additions to the housing inventory.—In the permit-issuing universe, a sample of new construction building permits, issued since the 1974 survey, was selected to represent housing units built

in permit-issuing areas since the 1974 survey. Sampling procedures were identical to those used in selecting the 1970-1974 new construction sample. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1974, to identify any units missed in the 1974 survey or any units added since the 1974 survey.

Coverage improvement sample selection.—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 census.
3. Units missed in the 1970 census.
4. 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.

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 A-2 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 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: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Orlando, Fla.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; and Wichita, Kans.

In the remaining 10 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 and the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 2,502 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, for each 1977-1978 SMSA, 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 520 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 no 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 units in multi-unit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multi-unit structures selected, all units were listed and matched

to the 1970 census. Any missed units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply these two procedures in all 20 A-2 SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in three SMSA's. In these SMSA's, a string of 10 structures (instead of 8) that were eligible to be selected for the AHS, were listed. The full application of this technique was employed in 12 SMSA's, with 5 SMSA's not receiving any application of the first procedure. The second procedure was instituted in 13 SMSA's. The following table shows which of these two techniques were applied to the A-2 SMSA's.

The first procedure added an estimated 7,084 units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 220 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 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 1977 housing inventory) and estimates pertaining to characteristics of units removed from the housing inventory since 1974 (i.e., 1974-1977 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1977 housing inventory.—The AHS estimates of characteristics of the 1977 housing inventory employed a 1-stage ratio estimation procedure in 19 SMSA's, with a second-stage ratio estimation procedure being employed in 1 SMSA. Prior to the implementation of the ratio estimation procedure, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of 1 or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample 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 esti-

SMSA	Coverage improvement technique employed			
	First		Second	Neither
	Partial	Full		
Albany-Schenectady-Troy, N.Y	X		X	
Anaheim-Santa Ana-Garden Grove, Calif	X			
Boston, Mass		X	X	
Dallas, Tex				X
Detroit, Mich		X	X	
Fort Worth, Tex		X	X	
Los Angeles-Long Beach, Calif				X
Madison, Wis		X	X	
Memphis, Tenn.-Ark		X	X	
Minneapolis-St. Paul, Minn				X
Newark, N.J		X	X	
Orlando, Fla		X	X	
Phoenix, Ariz		X	X	
Pittsburgh, Pa				X
Saginaw, Mich		X	X	
Salt Lake City, Utah				X
Spokane, Wash	X			
Tacoma, Wash		X	X	
Washington, D.C.-Md.-Va		X	X	
Wichita, Kans		X	X	

mation factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of 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 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 unit within the corresponding ratio estimation category.

The effect of this ratio estimation procedure was to somewhat reduce the variance due to the variation in sample size of strata used in the sample selection of the permit-issuing universe, thereby reducing the sampling error below what would have been obtained by simply weighting the results of the sample by the inverse of the probability of selection. Ordinarily, this would have been controlled by sampling within the strata during the sample selection process. However, prior to the AHS sample selection within each SMSA, units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. Thus, some variation in sample size was introduced during the AHS sample selection process.

For the Pittsburgh, Pa., SMSA, a second-stage ratio estimation procedure was also employed. This procedure involved the ratio estimation of the AHS weighted sample estimate of the October 1977 housing inventory to an independent estimate of the SMSA's October 1977 housing inventory. This ratio estimate factor equaled the following:

$$\frac{\text{SMSA independent estimate of October 1977 housing inventory}}{\text{AHS sample estimate of the housing inventory}}$$

The numerator of the ratio was derived by using the 1970 census estimate of total housing units for the SMSA in conjunction with the estimate of the change in the housing inventory since the census. The estimate of change was based on administrative records from utility companies. The denominator of that ratio was obtained from the weighted estimate of AHS sample units, using the existing weight after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for each sample unit and the resulting product was used as the final weight for tabulation purposes.

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

The second-stage ratio estimation procedure was not employed in all A-2 SMSA's, since it was felt that coverage improvement procedures adequately corrected known deficiencies in the sample where full coverage improvement procedures were employed. The independent estimates were only employed in SMSA's where full coverage improvements were not used and the estimated relative bias of the independent estimate was thought to be low enough (i.e., 2 percent or less over a 10-year period) to preclude the possibility of doing more harm than good to the survey results. A measure of the relative bias of the independent estimate was available based on a comparison between the 1970 census estimate of the April 1970 housing inventory and the independent estimate of the April 1970 housing inventory which was projected over a 10-year period from the 1960 census using the same procedure that produced the October 1977 independent estimates.

1974-1977 lost units.—The 1974-1977 lost unit estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1974 housing inventory, as was described in the 1974 Current Housing Report, Series H-170, **Housing Characteristics for Selected Metropolitan Areas.** Since the 1974-1977 lost units existed, by definition, in the 1974 housing inventory, there was a 1974 housing inventory weight associated with each 1974-1977 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1974-1977 lost units.

1974 estimation procedure.—This report presents data on the housing characteristics of the 1974 housing inventory from the 1974 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1974.

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 to provide correct information on the part of respondents, mistakes in recording or coding the data, and other errors of collection, response, processing, coverage, and estimation for missing data. Nonsampling errors are not unique to sample surveys since they can, and do, occur in complete censuses as well.

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing, the 1977 AHS-SMSA sample, and the 1974 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 included space errors, definitional errors, and occupancy errors. The "content" errors measured the accuracy of the data collected for surveyed 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.—For the 1977 AHS-SMSA sample, a reinterview program was conducted to obtain a measurement of some of the components of the nonsampling error associated with the AHS estimates. A sample of households was revisited and answers to some of the questions on the AHS questionnaire were obtained again. The original interview and the reinterview were assumed to be two independent readings and, thus, were the basis for the measurement of the accuracy of the AHS data collected from interviewed households.

As part of the reinterview, a check was made at each of these households to determine if the following was done during the original interview:

1. The correct unit was visited.
2. The correct number of housing units were interviewed at that address.
3. The correct information on "Year Built" was obtained.
4. The correct information on "Tenure" was obtained.
5. The correct information on "Household Composition" was obtained.
6. The correct information on "Type of Housing Unit" was obtained.
7. The correct information on "Occupancy Status" was obtained.

This check was made for interviewer evaluation and control. That is, tolerance limits were derived to determine which interviewers passed or failed this reinterview check. The results of this study are available and are presented in the following Census Bureau memorandum, "Response Error in the Annual Housing Survey as Measured by Reinterviews—Year IV, 1977-1978 SMSA Sample."

For the A-2 reinterview program, measures of response inconsistency were taken on two subsamples—reconciled and unreconciled. After the question is answered in the reconciled subsample, the interviewer presented the original response and then asked the respondent to decide upon the best answer. Past experience has shown that the level of inconsistency is lower when the interviewer has the response available, and thus, it is believed that there is a downward bias in these estimates. In the unreconciled sample, the interviewer did not have access to previous responses. For the reconciled items, three-fourths showed low levels of response inconsistency while the remaining one-fourth were in the moderate range. For the unreconciled items, approximately one-half showed low levels of response disagreement, with the remaining one-half of the categories in the moderate and high inconsistency range. Although attitudinal items were excluded from the A-2 reinterview program, other reinterview programs have shown that, in general, attitudinal items such as items 102-104 (see facsimile of the 1977 AHS questionnaire, page App-34) have moderate to high levels of inconsistency. Moderate levels of inconsistency indicate that there is some problem with inconsistent reporting, and high levels indicate that improvements are needed in the method used to collect data or that the category concepts themselves are ambiguous.

A second measure of response error, the net difference rate, which is a measure of net error (or bias) in a particular category, was also provided by this reinterview program. None of the categories tested showed any significant evidence of bias.

The results of this study were based on sample data, so there is a sampling error associated with these estimates of nonsampling error. Therefore, the possibility of such errors should be taken into account when considering the results of this study.

A similar study was conducted for the 1974 AHS-SMSA sample. The results of which are presented in the following Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1974."

Coverage errors.—With respect to errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units, since, due to the relatively short

time span involved, it is possible that construction of these units was not completed at the time the survey was conducted. In which case, they would not have been eligible for interview. In addition to these deficiencies, new construction in special places that do not require building permits, such as military bases, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit. Also, additional coverage errors may exist in SMSA's where the complete Coverage Improvement Program was not applied (see the table on page App-45 of this appendix).

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all units located inside these ED's would be represented in the sample. However, it has been estimated that the 1974 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. Since these ED's were recanvassed for the 1977 survey, the number of missed housing units may be considerably less for 1977.

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 schedules, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by the standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the average result of all possible samples.

2. Approximately 90 percent of the intervals from 1.6 standard errors below the estimate to 1.6 standard errors above the estimate would include the average result of all possible samples.

3. Approximately 95 percent of the intervals from two standard errors below the estimate to two standard errors above the estimate would include the average result of all possible samples.

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

The figures presented in the following tables 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 1974 housing inventory can be found in the AHS Series H-170 reports for 1974.

Table I presents the standard errors applicable to estimates of characteristics of the 1977 housing inventory as well as estimates of characteristics of the 1974-1977 lost units. 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 5,300 for the total SMSA, 3,520 for the central city of the SMSA, and 3,910 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 presents the standard errors of estimated percentages of the 1977 housing inventory as well as estimated percentages of the 1974-1977 lost units. 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 over-

estimates 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 1977 there were 99,300 owner-occupied housing units in this SMSA. Interpolation in table I of this appendix shows that the standard error of an estimate of this size is approximately 1,610. 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
75,000	1,470
99,300	x
100,000	1,610

The entry for x is determined as follows by vertically interpolating between 1,470 and 1,610:

$$99{,}300-75{,}000 = 24{,}300$$
$$100{,}000-75{,}000 = 25{,}000$$
$$1{,}470+\frac{24{,}300}{25{,}000}(1{,}610-1{,}470) = 1{,}610$$

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1977 Housing Inventory and for Estimated Number of 1974-1977 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]	
	SMSA and central city	Not in central city
0	40	40
100	60	60
200	90	90
500	140	140
700	160	160
1,000	190	190
2,500	310	300
5,000	430	420
10,000	600	590
25,000	930	890
50,000	1,260	1,150
75,000	1,470	1,250
100,000 . . .	1,610	1,240
150,000 . . .	1,740	840
200,000	1,690	—

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

Consequently, the 68-percent confidence interval, as shown by these data, is from 97,690 to 100,910 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1977 owner-occupied housing units lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 96,724 to 101,876 housing units with 90 percent confidence; and that the average estimate lies within the interval from 96,080 to 102,520 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 99,300 owner-occupied housing units, 27,600, or 27.8 percent, had two bedrooms. Interpolation in table II of the appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 27.8 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	27.8	50
75,000	1.0	a	1.1
99,300		p	
100,000 . . .	0.8	b	1.0

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

$$27.8-25.0 = 2.8$$
$$50.0-25.0 = 25.0$$
$$1.0+\frac{2.8}{25.0}(1.1-1.0) = 1.0$$

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

$$27.8-25.0 = 2.8$$
$$50.0-25.0 = 25.0$$
$$0.8+\frac{2.8}{25.0}(1.0-0.8) = .82$$

3. The entry for "p" was then determined by vertical interpolation between 1.0 and 0.82.

$$99{,}300-75{,}000 = 24{,}300$$
$$100{,}000-75{,}000 = 25{,}000$$
$$1.0+\frac{24{,}300}{25{,}000}(0.82-1.0) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 27.0 to 28.6 percent; the 90-percent confidence interval is from 26.5 to 29.1 percent; and the 95-percent confidence interval is from 26.2 to 29.4 percent.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 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)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	27.4	27.4	27.4	27.4	27.4	30.7
200	15.9	15.9	15.9	15.9	18.8	21.7
500	7.0	7.0	7.0	8.2	11.9	13.7
700	5.1	5.1	5.1	7.0	10.1	11.6
1,000	3.6	3.6	4.2	5.8	8.4	9.7
2,500	1.5	1.5	2.7	3.7	5.3	6.1
5,000	0.7	0.9	1.9	2.6	3.8	4.3
10,000	0.4	0.6	1.3	1.8	2.7	3.1
25,000	0.15	0.4	0.8	1.2	1.7	1.9
50,000	0.08	0.3	0.6	0.8	1.2	1.4
75,000	0.05	0.2	0.5	0.7	1.0	1.1
100,000 ...	0.04	0.2	0.4	0.6	0.8	1.0
150,000 ...	0.03	0.2	0.3	0.5	0.7	0.8
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, for the central city, and for the balance (not in central city) estimates.

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. However, 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 1974 and 1977 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1974 and 1977 characteristics.

Illustration of the computation of the standard error of a difference.—Table A-1 of part A of this report shows that in 1977 there were 46,600 owner-occupied units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied units with two bedrooms and owner-occupied units with three bedrooms is 19,000. Table I shows the standard error of 46,600 is approximately 1,220 and the standard error of 27,600 is approximately 960. Therefore, the standard error of the estimated difference of 19,000 is about:

$$1{,}550 = \sqrt{(1{,}220)^2 + (960)^2}$$

Consequently, the 68-percent confidence interval for the 19,000 difference is from 17,450 to 20,550 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 16,520 to 21,480 housing units, and the 95-percent confidence interval is from 15,900 to 22,100. Thus, we can conclude with 95 percent confidence that the number of 1977 owner-occupied 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, read off the confidence interval corresponding to the two points established in step 2.

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 99,300 housing units.

1. Interpolation in table II shows that the standard error of 50 percent on a base of 99,300 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 table A-1 of part A, it can be seen by cumulating the frequencies for the first two categories that 46,200 owner-occupied housing units, or 46.5 percent, had one or two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) and that an additional 17,800 owner-occupied housing units, or 17.9 percent, had three persons (i.e., 2.5 to 3.5 persons).

By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{48.0-46.5}{17.9}\right) = 2.6$$

Similarly, the upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{52.0-46.5}{17.9}\right) = 2.8$$

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

Washington 1977

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1977 survey App-42
Selection of the 1974 sample ... App-42
1977 sample reduction App-43
1974-1977 additions to the housing inventory App-43
Coverage improvement sample selection App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1970 Census of Population and Housing App-45
ESTIMATION App-45
1977 housing inventory App-45
1974-1977 lost units App-46
1974 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES App-46
Nonsampling errors App-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-50
Illustration of the computation of the standard error of a difference App-51
Medians App-51
Illustration of the computation of the 95-percent confidence interval of a median ... App-52

SAMPLE DESIGN

Annual Housing Survey.—The estimates for each of the 20 SMSA's are based on data collected from the 1977 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 20 SMSA's, the data were collected for the 11-month period from April 1977 through February 1978 with one-eleventh of the sample units visited each month.

Each group of SMSA's is interviewed for the AHS on a rotating basis, and this group (A-2 SMSA's) was the first to be revisited. Nineteen of the 20 SMSA's were interviewed for the first time in 1974, while the Madison, Wis., SMSA was interviewed for the first time in 1975.

In each group of SMSA's, the largest SMSA from each of the 4 geographic regions is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All remaining SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the A-2 group (1977-1978) are: Boston, Mass.; Detroit, Mich.; Los Angeles-Long Beach, Calif.; and Washington, D.C.-Md.-Va.

The remaining SMSA's in the A-2 group are: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Saginaw, Mich,; Salt Lake City, Utah; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.

In this SMSA, 13,673 housing units were eligible for interview. Of these sample units, 742 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 units eligible for interview, 1,318 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 1977 survey.—The sample housing units designated for interview in the 1977 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 1974 survey which were not part of the 1977 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 1974 survey which were not part of the 1977 reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1977 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from the list of new construction building permits issued since the 1974 survey. (This sample represented the housing units built in permit-issuing areas, since the 1974 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1974 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1974 survey.)
5. All sample housing units that were selected as part of the 1977 Coverage Improvement Program. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1974 sample.—The sample for the SMSA's which are 100-percent permit-issuing was selected from two sample frames—units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which are not 100-percent permit-issuing in-

cluded a sample selected from a third frame—those units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). The following SMSA's are 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Los Angeles-Long Beach, Calif.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 14 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 four largest SMSA's, the overall sampling rate differed by central city and 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 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 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 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-$5,999. . .		
$6,000-$9,999. . .		
$10,000-$14,999 .		
$15,000 and over .		

Thus, for this SMSA, the occupied housing unit records from this 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 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 are 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 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 units in existence at the time of interview in these selected segments are in sample. Thus, units enumerated in the 1970 census as well as units built since the 1970 census are included.

1977 sample reduction.—In previous years, AHS sample units have been divided among 12 panels which were interviewed each month. Due to budget limitations, it became necessary to drop one of these panels from the 1977 survey. The designated sample size for 1977 was thus reduced by one-twelfth.

1974-1977 additions to the housing inventory.—In the permit-issuing universe, a sample of new construction building permits, issued since the 1974 survey, was selected to represent housing units built

in permit-issuing areas since the 1974 survey. Sampling procedures were identical to those used in selecting the 1970-1974 new construction sample. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1974, to identify any units missed in the 1974 survey or any units added since the 1974 survey.

Coverage improvement sample selection.—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 census.
3. Units missed in the 1970 census.
4. 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.

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 A-2 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 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: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Orlando, Fla.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; and Wichita, Kans.

In the remaining 10 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 and the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 9,327 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, for each 1977-1978 SMSA, 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 1,015 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 no 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 units in multi-unit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multi-unit structures selected, all units were listed and matched

to the 1970 census. Any missed units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply these two procedures in all 20 A-2 SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in three SMSA's. In these SMSA's, a string of 10 structures (instead of 8) that were eligible to be selected for the AHS, were listed. The full application of this technique was employed in 12 SMSA's, with 5 SMSA's not receiving any application of the first procedure. The second procedure was instituted in 13 SMSA's. The following table shows which of these two techniques were applied to the A-2 SMSA's:

SMSA	Coverage improvement technique employed			
	First		Second	Neither
	Partial	Full		
Albany-Schenectady-Troy, N.Y	X		X	
Anaheim-Santa Ana-Garden Grove, Calif	X			
Boston, Mass		X	X	
Dallas, Tex				X
Detroit, Mich		X	X	
Fort Worth, Tex		X	X	
Los Angeles-Long Beach, Calif				X
Madison, Wis		X	X	
Memphis, Tenn.-Ark		X	X	
Minneapolis-St. Paul, Minn				X
Newark, N.J		X	X	
Orlando, Fla		X	X	
Phoenix, Ariz		X	X	
Pittsburgh, Pa				X
Saginaw, Mich		X	X	
Salt Lake City, Utah				X
Spokane, Wash	X			
Tacoma, Wash		X	X	
Washington, D.C.-Md.-Va		X	X	
Wichita, Kans		X	X	

The first procedure added an estimated 23,220 units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 898 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 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 1977 housing inventory) and estimates pertaining to characteristics of units removed from the housing inventory since 1974 (i.e., 1974-1977 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1977 housing inventory.—The AHS estimates of characteristics of the 1977 housing inventory employed a 1-stage ratio estimation procedure in 19 SMSA's, with a second-stage ratio estimation procedure being employed in 1 SMSA. Prior to the implementation of the ratio estimation procedure, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of 1 or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample 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 estima-

tion factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of 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 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 unit within the corresponding ratio estimation category.

The effect of this ratio estimation procedure was to somewhat reduce the variance due to the variation in sample size of strata used in the sample selection of the permit-issuing universe, thereby reducing the sampling error below what would have been obtained by simply weighting the results of the sample by the inverse of the probability of selection. Ordinarily, this would have been controlled by sampling within the strata during the sample selection process. However, prior to the AHS sample selection within each SMSA, units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. Thus, some variation in sample size was introduced during the AHS sample selection process.

For the Pittsburgh, Pa., SMSA, a second-stage ratio estimation procedure was also employed. This procedure involved the ratio estimation of the AHS weighted sample estimate of the October 1977 housing inventory to an independent estimate of the SMSA's October 1977 housing inventory. This ratio estimate factor equaled the following:

$$\frac{\text{SMSA independent estimate of October 1977 housing inventory}}{\text{AHS sample estimate of the housing inventory}}$$

The numerator of the ratio was derived by using the 1970 census estimate of total housing units for the SMSA in conjunction with the estimate of the change in the housing inventory since the census. The estimate of change was based on administrative records from utility companies. The denominator of that ratio was obtained from the weighted estimate of AHS sample units, using the existing weight after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for each sample unit and the resulting product was used as the final weight for tabulation purposes.

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

The second-stage ratio estimation procedure was not employed in all A-2 SMSA's, since it was felt that coverage improvement procedures adequately corrected known deficiencies in the sample where full coverage improvement procedures were employed. The independent estimates were only employed in SMSA's where full coverage improvements were not used and the estimated relative bias of the independent estimate was thought to be low enough (i.e., 2 percent or less over a 10-year period) to preclude the possibility of doing more harm than good to the survey results. A measure of the relative bias of the independent estimate was available based on a comparison between the 1970 census estimate of the April 1970 housing inventory and the independent estimate of the April 1970 housing inventory which was projected over a 10-year period from the 1960 census using the same procedure that produced the October 1977 independent estimates.

1974-1977 lost units.—The 1974-1977 lost unit estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1974 housing inventory, as was described in the 1974 Current Housing Report, Series H-170, **Housing Characteristics for Selected Metropolitan Areas.** Since the 1974-1977 lost units existed, by definition, in the 1974 housing inventory, there was a 1974 housing inventory weight associated with each 1974-1977 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1974-1977 lost units.

1974 estimation procedure.—This report presents data on the housing characteristics of the 1974 housing inventory from the 1974 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1974.

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 to provide correct information on the part of respondents, mistakes in recording or coding the data, and other errors of collection, response, processing, coverage, and estimation for missing data. Nonsampling errors are not unique to sample surveys since they can, and do, occur in complete censuses as well.

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing, the 1977 AHS-SMSA sample, and the 1974 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 included space errors, definitional errors, and occupancy errors. The "content" errors measured the accuracy of the data collected for surveyed 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.—For the 1977 AHS-SMSA sample, a reinterview program was conducted to obtain a measurement of some of the components of the nonsampling error associated with the AHS estimates. A sample of households was revisited and answers to some of the questions on the AHS questionnaire were obtained again. The original interview and the reinterview were assumed to be two independent readings and, thus, were the basis for the measurement of the accuracy of the AHS data collected from interviewed households.

As part of the reinterview, a check was made at each of these households to determine if the following was done during the original interview:

1. The correct unit was visited.
2. The correct number of housing units were interviewed at that address.
3. The correct information on "Year Built" was obtained.
4. The correct information on "Tenure" was obtained.
5. The correct information on "Household Composition" was obtained.
6. The correct information on "Type of Housing Unit" was obtained.
7. The correct information on "Occupancy Status" was obtained.

This check was made for interviewer evaluation and control. That is, tolerance limits were derived to determine which interviewers passed or failed this reinterview check. The results of this study are available and are presented in the following Census Bureau memorandum, "Response Error in the Annual Housing Survey as Measured by Reinterviews—Year IV, 1977-1978 SMSA Sample."

For the A-2 reinterview program, measures of response inconsistency were taken on two subsamples—reconciled and unreconciled. After the question is answered in the reconciled subsample, the interviewer presented the original response and then asked the respondent to decide upon the best answer. Past experience has shown that the level of inconsistency is lower when the interviewer has the response available, and thus, it is believed that there is a downward bias in these estimates. In the unreconciled sample, the interviewer did not have access to previous responses. For the reconciled items, three-fourths showed low levels of response inconsistency while the remaining one-fourth were in the moderate range. For the unreconciled items, approximately one-half showed low levels of response disagreement, with the remaining one-half of the categories in the moderate and high inconsistency range. Although attitudinal items were excluded from the A-2 reinterview program, other reinterview programs have shown that, in general, attitudinal items such as items 102-104 (see facsimile of the 1977 AHS questionnaire, page App-34) have moderate to high levels of inconsistency. Moderate levels of inconsistency indicate that there is some problem with inconsistent reporting, and high levels indicate that improvements are needed in the method used to collect data or that the category concepts themselves are ambiguous.

A second measure of response error, the net difference rate, which is a measure of net error (or bias) in a particular category, was also provided by this reinterview program. None of the categories tested showed any significant evidence of bias.

The results of this study were based on sample data, so there is a sampling error associated with these estimates of nonsampling error. Therefore, the possibility of such errors should be taken into account when considering the results of this study.

A similar study was conducted for the 1974 AHS-SMSA sample. The results of which are presented in the following Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1974."

Coverage errors.—With respect to errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units, since, due to the relatively short time span involved, it is possible that construction of these units was not com-

pleted at the time the survey was conducted. In which case, they would not have been eligible for interview. In addition to these deficiencies, new construction in special places that do not require building permits, such as military bases, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit. Also, additional coverage errors may exist in SMSA's where the complete Coverage Improvement Program was not applied (see the table on page App-45 of this appendix).

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all units located inside these ED's would be represented in the sample. However, it has been estimated that the 1974 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. Since these ED's were recanvassed for the 1977 survey, the number of missed housing units may be considerably less for 1977.

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 schedules, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by the standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

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

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

The figures presented in the following tables 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 1974 housing inventory can be found in the AHS Series H-170 reports for 1974.

Table I presents the standard errors applicable to estimates of characteristics of the 1977 housing inventory as well as estimates of characteristics of the 1974-1977 lost units. Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 19,870 for the total SMSA, 8,620 for the central city of the SMSA, and 17,690 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 present the standard errors of estimated percentages of the 1977 housing inventory as well as estimated percentages of the 1974-1977 lost units. 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

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1977 Housing Inventory and for Estimated Number of 1974-1977 Lost Units for the Washington, D.C.-Md.-Va., SMSA, for the Central City, and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city
0	100	50	140
100	100	70	140
200	140	100	160
500	230	150	260
700	270	180	310
1,000	320	210	370
2,500	510	340	580
5,000	720	470	820
10,000	1,010	670	1,160
25,000	1,590	1,030	1,820
50,000	2,230	1,400	2,540
75,000	2,700	1,640	3,070
100,000	3,090	1,800	3,500
150,000	3,720	1,970	4,170
200,000	4,210	1,960	4,680
250,000	4,610	1,780	5,070
300,000	4,940	1,340	5,380
400,000	5,450	—	5,770
500,000	5,800	—	5,930
600,000	6,010	—	5,860
700,000	6,100	—	5,560
800,000	6,070	—	4,980
900,000	5,920	—	4,020
1,000,000	5,640	—	—
1,100,000	5,210	—	—
1,200,000	4,600	—	—

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

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 Lost Units for the Washington, D.C.-Md.-Va. SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	50.7	50.7	50.7	50.7	50.7	50.7
200	33.9	33.9	33.9	33.9	33.9	35.8
500	17.0	17.0	17.0	17.0	19.6	22.7
700	12.8	12.8	12.8	12.8	16.6	19.2
1,000	9.3	9.3	9.3	9.6	13.9	16.0
2,500	3.9	3.9	4.4	6.1	8.8	10.1
5,000	2.0	2.0	3.1	4.3	6.2	7.2
10,000	1.0	1.0	2.2	3.0	4.4	5.1
25,000	.4	.6	1.4	1.9	2.8	3.2
50,000	.2	.5	1.0	1.4	2.0	2.3
75,000	.14	.4	.8	1.1	1.6	1.9
100,000	.10	.3	.7	1.0	1.4	1.6
150,000	.07	.3	.6	.8	1.1	1.3
200,000	.05	.3	.5	.7	1.0	1.1
250,000	.04	.2	.4	.6	.9	1.0
300,000	.03	.18	.4	.6	.8	.9
400,000	.03	.16	.3	.5	.7	.8
500,000	.02	.14	.3	.4	.6	.7
600,000	.02	.13	.3	.4	.6	.7
700,000	.01	.12	.3	.4	.5	.6
800,000	.01	.11	.2	.3	.5	.6
900,000	.01	.11	.2	.3	.5	.5
1,000,000	.01	.10	.2	.3	.4	.5
1,100,000	.01	.10	.2	.3	.4	.5
1,200,000	.01	.09	.2	.3	.4	.5

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.3.

Illustration of the use of the standard error tables.—Table A-1 of part A of this report shows that in 1977 there were 543,800 owner-occupied housing units in this SMSA. Interpolation in table I of this appendix shows that the standard error of an estimate of this size is approximately 5,890. 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.	5,800
543,800.	x
600,000.	6,010

The entry for x is determined as follows by vertically interpolating between 5,800 and 6,010:

$$543{,}800-500{,}000 = 43{,}800$$
$$600{,}000-500{,}000 = 100{,}000$$

$$5{,}800 + \frac{43{,}800}{100{,}000}(6{,}010-5{,}800) = 5{,}890$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 537,910 to 549,690 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1977 owner-occupied housing units lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 534,376 to 553,224 housing units with 90 percent confidence; and that the average estimate lies within the interval from 555,580 to 532,020 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 543,800 owner-occupied housing units, 74,500, or 13.7 percent, had two bedrooms. Interpolation in table II of the appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 13.7 percent is approximately 0.4 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	13.7	25 or 75
500,000. . . .	0.4	a	0.6
543,800. . . .		p	
600,000. . . .	0.4	b	0.6

TABLE III. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 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]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100.	31.3	31.3	31.3	31.3	31.3	33.8
200.	18.6	18.6	18.6	18.6	20.7	23.9
500.	8.4	8.4	8.4	9.1	13.1	15.1
700.	6.1	6.1	6.1	7.7	11.1	12.8
1,000	4.4	4.4	4.7	6.4	9.3	10.7
2,500	1.8	1.8	2.9	4.1	5.9	6.8
5,000	.9	1.0	2.1	2.9	4.1	4.8
10,000	.5	.7	1.5	2.0	2.9	3.4
25,000	.18	.4	.9	1.3	1.9	2.1
50,000	.09	.3	.7	.9	1.3	1.5
75,000	.06	.2	.5	.7	1.1	1.2
100,000. . . .	.05	.2	.5	.6	.9	1.1
150,000. . . .	.03	.17	.4	.5	.8	.9
200,000. . . .	.02	.15	.3	.5	.7	.8
250,000. . . .	.02	.13	.3	.4	.6	.7
300,000. . . .	.02	.12	.3	.4	.5	.6

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.3.

1. The entry for cell "a" is determined by horizontal interpolation between 0.4 and 0.6.

$$13.7-10.0 = 3.7$$
$$25.0-10.0 = 15.0$$

$$0.4 + \frac{3.7}{15.0}(0.6-0.4) = 0.4$$

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

$$13.7-10.0 = 3.7$$
$$25.0-10.0 = 15.0$$

$$0.4 + \frac{3.7}{15.0}(0.6-0.4) = 0.4$$

3. The entry for "p" was then determined by vertical interpolation between 0.4 and 0.4.

$$543{,}800-500{,}000 = 43{,}800$$
$$600{,}000-500{,}000 = 100{,}000$$

$$0.4 + \frac{43{,}800}{100{,}000}(0.4-0.4) = 0.4$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 13.3 to 14.1 percent; the 90-percent confidence interval is from 13.1 to 14.3 percent; and the 95-percent interval is from 12.9 to 14.5 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 esti-

TABLE IV. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 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]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	57.6	57.6	57.6	57.6	57.6	58.3
200	40.4	40.4	40.4	40.4	40.4	41.2
500	21.3	21.3	21.3	21.3	22.6	26.1
700	16.2	16.2	16.2	16.2	19.1	22.0
1,000	12.0	12.0	12.0	12.0	16.0	18.4
2,500	5.1	5.1	5.1	7.0	10.1	11.7
5,000	2.6	2.6	3.6	4.9	7.1	8.2
10,000	1.3	1.3	2.5	3.5	5.0	5.8
25,000	.5	.7	1.6	2.2	3.2	3.7
50,000	.3	.5	1.1	1.6	2.3	2.6
75,000	.18	.4	.9	1.3	1.8	2.1
100,000	.14	.4	.8	1.1	1.6	1.8
150,000	.09	.3	.7	.9	1.3	1.5
200,000	.07	.3	.6	.8	1.1	1.3
250,000	.05	.2	.5	.7	1.0	1.2
300,000	.05	.2	.5	.6	.9	1.1
400,000	.03	.18	.4	.6	.8	.9
500,000	.03	.16	.4	.5	.7	.8
600,000	.02	.15	.3	.5	.7	.8
700,000	.02	.14	.3	.4	.6	.7
800,000	.02	.13	.3	.4	.6	.7
900,000	.02	.12	.3	.4	.5	.6

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

mates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. However, 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 1974 and 1977 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1974 and 1977 characteristics.

Illustration of the computation of the standard error of a difference.—Table A-1 of part A of this report shows that in 1977 there were 250,600 owner-occupied units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied units with two bedrooms and owner-occupied units with three bedrooms is 176,100. Table I shows the standard error of 250,600 is approximately 4,610, and the standard error of 74,500 is approximately 2,690. Therefore, the standard error of the estimated difference of 176,100 is about:

$$5{,}340 = \sqrt{(4{,}610)^2 + (2{,}690)^2}$$

Consequently, the 68-percent confidence interval for the 176,100 difference is from 170,760 to 181,440 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 167,556 to 184,644 housing units, and the 95-percent confidence interval is from 165,420 to 186,780. Thus, we can conclude with 95 percent confidence that the number of 1977 owner-occupied 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, read off the confidence interval corresponding to the two points established in step 2.

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

A two-standard error confidence interval may be determined by finding the values corresponding to 50 percent plus and minus twice the standard error determined in step 1. For about 95 out of

100 possible samples, the average median from all possible samples would lie between these two values.

Illustration of the computation of the 95-percent confidence interval of a median.—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 3.1. The base of the distribution from which this median was determined is 543,800 housing units.

1. Interpolation in table II shows that the standard error of 50 percent on a base of 543,800 is approximately 0.7 percentage points.

2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.6 and 51.4.

3. From table A-1 of part A, it can be seen by cumulating the frequencies for the first two categories that 206,900 owner-occupied housing units, or 38.0 percent, had one or two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) and that an additional 104,200 owner-occupied housing units, or 19.2 percent, had three persons (i.e., 2.5 to 3.5 persons).

By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{48.6-38.0}{19.2}\right) = 3.0$$

Similarly, the upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{51.4-38.0}{19.2}\right) = 3.2$$

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

Wichita 1977

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1977 survey App-42
Selection of the 1974 sample App-42
1977 sample reduction App-43
1974-1977 additions to the housing inventory App-43
Coverage improvement sample selection App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1970 Census of Population and Housing App-45
ESTIMATION App-45
1977 housing inventory App-45
1974-1977 lost units App-46
1974 estimation procedure App-46
Ratio estimation procedure of the 1970 Census of Population and Housing App-46
RELIABILITY OF THE ESTIMATES App-46
Nonsampling errors App-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-50
Illustration of the computation of the standard error of a difference App-50
Medians App-51
Illustration of the computation of the 95-percent confidence interval of a median ... App-51

SAMPLE DESIGN

Annual Housing Survey.—The estimates for each of the 20 SMSA's are based on data collected from the 1977 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development. In each of the 20 SMSA's, the data were collected for the 11-month period from April 1977 through February 1978 with one-eleventh of the sample units visited each month.

Each group of SMSA's is interviewed for the AHS on a rotating basis, and this group (A-2 SMSA's) was the first to be revisited. Nineteen of the 20 SMSA's were interviewed for the first time in 1974, while the Madison, Wis., SMSA was interviewed for the first time in 1975.

In each group of SMSA's, the largest SMSA from each of the 4 geographic regions is represented by a sample of about 15,000 designated housing units evenly divided between the central city and the balance of the respective SMSA. All remaining SMSA's are each represented by a sample of about 5,000 designated housing units distributed proportionately between the central city and the balance of the respective SMSA based on the distribution of total housing units in each sector.

The largest SMSA's in the A-2 group (1977-1978) are: Boston, Mass.; Detroit, Mich.; Los Angeles-Long Beach, Calif.; and Washington, D.C.-Md.-Va.

The remaining SMSA's in the A-2 group are: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Dallas, Tex.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark; Minneapolis-St. Paul, Minn.; Newark, N.J.; Orlando, Fla.; Phoenix, Ariz.; Pittsburgh, Pa.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; Tacoma, Wash.; and Wichita, Kans.

In this SMSA, 4,915 housing units were eligible for interview. Of these sample units, 129 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 units eligible for interview, 428 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 1977 survey.—The sample housing units designated for interview in the 1977 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 1974 survey which were not part of the 1977 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 1974 survey which were not part of the 1977 reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1977 AHS questionnaire, page App-20.)
3. All sample housing units that were selected from the list of new construction building permits issued since the 1974 survey. (This sample represented the housing units built in permit-issuing areas since the 1974 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1974 survey. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1974 survey.)
5. All sample housing units that were selected as part of the 1977 Coverage Improvement Program. (This sample represented most of the housing units which, until 1977, did not have a chance of selection.)

Selection of the 1974 sample.—The sample for the SMSA's which are 100-percent permit-issuing was selected from two sample frames—units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which are not 100-percent permit-issuing included

a sample selected from a third frame—those units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). The following SMSA's are 100-percent permit-issuing: Anaheim-Santa Ana-Garden Grove, Calif.; Boston, Mass.; Los Angeles-Long Beach, Calif.; Newark, N.J.; Phoenix, Ariz.; and Washington, D.C.-Md.-Va. The remaining 14 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 four largest SMSA's, the overall sampling rate differed by central city and 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 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 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 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-$5,999. . .		
$6,000-$9,999. . .		
$10,000-$14,999 .		
$15,000 and over .		

Thus, for this SMSA, the occupied housing unit records from this 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 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 are 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 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 units in existence at the time of interview in these selected segments are in sample. Thus, units enumerated in the 1970 census as well as units built since the 1970 census are included.

1977 sample reduction.—In previous years, AHS sample units have been divided among 12 panels which were interviewed each month. Due to budget limitations, it became necessary to drop one of these panels from the 1977 survey. The designated sample size for 1977 was thus reduced by one-twelfth.

1974-1977 additions to the housing inventory.—In the permit-issuing universe, a sample of new construction building permits, issued since the 1974 survey, was selected to represent housing units built

in permit-issuing areas since the 1974 survey. Sampling procedures were identical to those used in selecting the 1970-1974 new construction sample. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1974, to identify any units missed in the 1974 survey or any units added since the 1974 survey.

Coverage improvement sample selection.—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 census.
3. Units missed in the 1970 census.
4. 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.

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 A-2 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 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: Albany-Schenectady-Troy, N.Y.; Anaheim-Santa Ana-Garden Grove, Calif.; Fort Worth, Tex.; Madison, Wis.; Memphis, Tenn.-Ark.; Orlando, Fla.; Saginaw, Mich.; Salt Lake City, Utah; Spokane, Wash.; and Wichita, Kans.

In the remaining 10 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 and the Memphis, Tenn.-Ark., SMSA to supplement the sample described above. These procedures added an estimated 718 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, for each 1977-1978 SMSA, 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 1,885 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 no 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 units in multi-unit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multi-unit structures

selected, all units were listed and matched to the 1970 census. Any missed units were then assigned for interview.

Based on a cost-benefit analysis, a decision was made not to apply these two procedures in all 20 A-2 SMSA's. A partial application of the first procedure, which did not include structures missed in the 1970 census, was instituted in three SMSA's. In these SMSA's, a string of 10 structures (instead of 8) that were eligible to be selected for the AHS, were listed. The full application of this technique was employed in 12 SMSA's, with 5 SMSA's not receiving any application of the first procedure. The second procedure was instituted in 13 SMSA's. The following table shows which of these two techniques were applied to the A-2 SMSA's.

The first procedure added an estimated 2,160 units to the coverage of the housing inventory of this SMSA, while the second procedure added an estimated 148 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 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 1977 housing inventory) and estimates pertaining to characteristics of units removed from the housing inventory since 1974 (i.e., 1974-1977 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1977 housing inventory.—The AHS estimates of characteristics of the 1977 housing inventory employed a 1-stage ratio estimation procedure in 19 SMSA's, with a second-stage ratio estimation procedure being employed in 1 SMSA. Prior to the implementation of the ratio estimation procedure, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

Within each sector of each SMSA, a noninterview factor was computed separately for 50 noninterview cells for sample housing units from the permit-issuing universe (where the cells consisted of 1 or more of the different strata used in the stratification of the universe as previously described). In addition, within each sector, separate noninterview factors were computed for one noninterview cell for conventional new construction sample housing units from both the permit-issuing universe and the coverage improvement universe, one noninterview cell for mobile homes from both the nonpermit universe and the coverage improvement universe, and one noninterview cell for other sample 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 estima-

SMSA	Coverage improvement technique employed			
	First		Second	Neither
	Partial	Full		
Albany-Schenectady-Troy, N.Y	X		X	
Anaheim-Santa Ana-Garden Grove, Calif	X			
Boston, Mass		X	X	
Dallas, Tex				X
Detroit, Mich		X	X	
Fort Worth, Tex		X	X	
Los Angeles-Long Beach, Calif				X
Madison, Wis		X	X	
Memphis, Tenn.-Ark		X	X	
Minneapolis-St. Paul, Minn				X
Newark, N.J		X	X	
Orlando, Fla		X	X	
Phoenix, Ariz		X	X	
Pittsburgh, Pa				X
Saginaw, Mich		X	X	
Salt Lake City, Utah				X
Spokane, Wash	X			
Tacoma, Wash		X	X	
Washington, D.C.-Md.-Va		X	X	
Wichita, Kans		X	X	

tion factor for each cell was equal to the following:

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

For each SMSA, the numerators of the ratios were obtained from the 1970 Census of Population and Housing 20-percent file of 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 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 unit within the corresponding ratio estimation category.

The effect of this ratio estimation procedure was to somewhat reduce the variance due to the variation in sample size of strata used in the sample selection of the permit-issuing universe, thereby reducing the sampling error below what would have been obtained by simply weighting the results of the sample by the inverse of the probability of selection. Ordinarily, this would have been controlled by sampling within the strata during the sample selection process. However, prior to the AHS sample selection within each SMSA, units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. Thus, some variation in sample size was introduced during the AHS sample selection process.

For the Pittsburgh, Pa., SMSA, a second-stage ratio estimation procedure was also employed. This procedure involved the ratio estimation of the AHS weighted sample estimate of the October 1977 housing inventory to an independent estimate of the SMSA's October 1977 housing inventory. This ratio estimate factor equaled the following:

$$\frac{\text{SMSA independent estimate of October 1977 housing inventory}}{\text{AHS sample estimate of the housing inventory}}$$

The numerator of the ratio was derived by using the 1970 census estimate of total housing units for the SMSA in conjunction with the estimate of the change in the housing inventory since the census. The estimate of change was based on administrative records from utility companies. The denominator of that ratio was obtained from the weighted estimate of AHS sample units, using the existing weight after the first-stage ratio estimation procedure. The computed second-stage ratio estimation factor was then applied to the existing weight for each sample unit and the resulting product was used as the final weight for tabulation purposes.

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

The second-stage ratio estimation procedure was not employed in all A-2 SMSA's, since it was felt that coverage improvement procedures adequately corrected known deficiencies in the sample where full coverage improvement procedures were employed. The independent estimates were only employed in SMSA's where full coverage improvements were not used and the estimated relative bias of the independent estimate was thought to be low enough (i.e., 2 percent or less over a 10-year period) to preclude the possibility of doing more harm than good to the survey results. A measure of the relative bias of the independent estimate was available based on a comparison between the 1970 census estimate of the April 1970 housing inventory and the independent estimate of the April 1970 housing inventory which was projected over a 10-year period from the 1960 census using the same procedure that produced the October 1977 independent estimates.

1974-1977 lost units.—The 1974-1977 lost unit estimates employed the three-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1974 housing inventory, as was described in the 1974 Current Housing Report, Series H-170, **Housing Characteristics for Selected Metropolitan Areas.** Since the 1974-1977 lost units existed, by definition, in the 1974 housing inventory, there was a 1974 housing inventory weight associated with each 1974-1977 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1974-1977 lost units.

1974 estimation procedure.—This report presents data on the housing characteristics of the 1974 housing inventory from the 1974 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a three-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1974.

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 to provide correct information on the part of respondents, mistakes in recording or coding the data, and other errors of collection, response, processing, coverage, and estimation for missing data. Nonsampling errors are not unique to sample surveys since they can, and do, occur in complete censuses as well.

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing, the 1977 AHS-SMSA sample, and the 1974 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 included space errors, definitional errors, and occupancy errors. The "content" errors measured the accuracy of the data collected for surveyed 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.—For the 1977 AHS-SMSA sample, a reinterview program was conducted to obtain a measurement of some of the components of the nonsampling error associated with the AHS estimates. A sample of households was revisited and answers to some of the questions on the AHS questionnaire were obtained again. The original interview and the reinterview were assumed to be two independent readings and, thus, were the basis for the measurement of the accuracy of the AHS data collected from interviewed households.

As part of the reinterview, a check was made at each of these households to determine if the following was done during the original interview:

1. The correct unit was visited.
2. The correct number of housing units were interviewed at that address.
3. The correct information on "Year Built" was obtained.
4. The correct information on "Tenure" was obtained.
5. The correct information on "Household Composition" was obtained.
6. The correct information on "Type of Housing Unit" was obtained.
7. The correct information on "Occupancy Status" was obtained.

This check was made for interviewer evaluation and control. That is, tolerance limits were derived to determine which interviewers passed or failed this reinterview check. The results of this study are available and are presented in the following Census Bureau memorandum, "Response Error in the Annual Housing Survey as Measured by Reinterviews—Year IV, 1977-1978 SMSA Sample."

For the A-2 reinterview program, measures of response inconsistency were taken on two subsamples—reconciled and unreconciled. After the question is answered in the reconciled subsample, the interviewer presented the original response and then asked the respondent to decide upon the best answer. Past experience has shown that the level of inconsistency is lower when the interviewer has the response available, and thus, it is believed that there is a downward bias in these estimates. In the unreconciled sample, the interviewer did not have access to previous responses. For the reconciled items, three-fourths showed low levels of response inconsistency while the remaining one-fourth were in the moderate range. For the unreconciled items, approximately one-half showed low levels of response disagreement, with the remaining one-half of the categories in the moderate and high inconsistency range. Although attitudinal items were excluded from the A-2 reinterview program, other reinterview programs have shown that, in general, attitudinal items such as items 102-104 (see facsimile of the 1977 AHS questionnaire, page App-34) have moderate to high levels of inconsistency. Moderate levels of inconsistency indicate that there is some problem with inconsistent reporting, and high levels indicate that improvements are needed in the method used to collect data or that the category concepts themselves are ambiguous.

A second measure of response error, the net difference rate, which is a measure of net error (or bias) in a particular category, was also provided by this reinterview program. None of the categories tested showed any significant evidence of bias.

The results of this study were based on sample data, so there is a sampling error associated with these estimates of nonsampling error. Therefore, the possibility of such errors should be taken into account when considering the results of this study.

A similar study was conducted for the 1974 AHS-SMSA sample. The results of which are presented in the following Census Bureau memorandum, "Reinterview Results for Annual Housing Survey—SMSA Sample: 1974."

Coverage errors.—With respect to errors of coverage and estimation for missing data, the AHS new construction sample had deficiencies in the representation of conventional new construction. Due to time constraints, only those building permits issued more than 5 months before the survey ended were eligible to be sampled to represent conventional new construction in permit-issuing areas for this SMSA. However, these permits issued during the last 5 months of the survey do not necessarily represent missed housing units, since, due to the relatively short time span involved, it is possible that construction of these units was not com-

pleted at the time the survey was conducted. In which case, they would not have been eligible for interview. In addition to these deficiencies, new construction in special places that do not require building permits, such as military bases, are also not adequately represented.

The Coverage Improvement Program also had certain deficiencies. It appears that the listing procedure used to correct deficiencies 3-6 (see the coverage improvement section of this appendix) was not very effective in finding nonresidential conversions. Such conversions were primarily in business districts, whereas the listing procedure started from a residential unit. Also, additional coverage errors may exist in SMSA's where the complete Coverage Improvement Program was not applied (see the table on page App-45 of this appendix).

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all units located inside these ED's would be represented in the sample. However, it has been estimated that the 1974 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. Since these ED's were recanvassed for the 1977 survey, the number of missed housing units may be considerably less for 1977.

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 schedules, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling errors, but it does not measure, as such, any systematic biases in the data. Therefore, the accuracy of the estimates depends on both the sampling and nonsampling errors measured by the standard error, biases, and any additional nonsampling errors not measured by the standard error. The sample estimate and its estimated standard error enable one to construct interval estimates in which the interval includes the average result of all possible samples with a known probability. For example, if all possible samples were selected, each of these surveyed under essentially the same general conditions, and an estimate and its estimated standard error were calculated from each sample, then:

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

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

The figures presented in the following tables 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 1974 housing inventory can be found in the AHS Series H-170 reports for 1974.

Table I presents the standard errors applicable to estimates of characteristics of the 1977 housing inventory as well as estimates of characteristics of the

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1977 Housing Inventory and for Estimated Number of 1974-1977 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]		
	SMSA	Central city	Not in central city
0	30	30	30
100.	60	50	60
200.	80	80	80
500.	120	120	130
700.	150	140	160
1,000 . . .	170	170	190
2,500 . . .	270	270	300
5,000 . . .	390	370	430
10,000 . .	550	520	620
25,000 . .	850	770	1,070
50,000 . .	1,170	970	1,700
75,000 . .	1,390	1,020	–
100,000. .	1,550	950	–
150,000. .	1,780	–	–
200,000. .	1,890	–	–

[1] For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2 for the total SMSA, 1.3 for the central city, and 1.1 for the balance (not in central city) estimates.

1974-1977 lost units. Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 7,010 for the total SMSA, 3,890 for the central city of the SMSA, and 5,730 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 present the standard errors of estimated percentages of the 1977 housing inventory as well as estimated percentages of the 1974-1977 lost units. 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 through III are 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 1977 there were 95,400 owner-occupied housing units in this SMSA. Interpolation in table I of this appendix shows that the standard error of an estimate of this size is approximately 1,520. The following interpolation procedure was used:

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

Size of estimate	Standard error
75,000	1,390
95,400	x
100,000.	1,550

The entry for x is determined as follows by vertically interpolating between 1,390 and 1,550:

$$95{,}400-75{,}000 = 20{,}400$$
$$100{,}000-75{,}000 = 25{,}000$$

$$1{,}390+\frac{20{,}400}{25{,}000}(1{,}550-1{,}390)$$
$$= 1{,}520$$

Consequently, the 68 percent confidence interval, as shown by these data, is from 93,880 to 96,920 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1977 owner-occupied housing units lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 92,968 to 97,832 housing units with 90 percent confidence; and that the average estimate lies within the interval from 92,360 to

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 Lost Units for the Wichita, Kans., SMSA, and the Central City of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100.	23.3	23.3	23.3	23.3	23.8	27.5
200.	13.2	13.2	13.2	13.2	16.9	19.5
500.	5.7	5.7	5.7	7.4	10.7	12.3
700.	4.2	4.2	4.5	6.2	9.0	10.4
1,000	2.9	2.9	3.8	5.2	7.5	8.7
2,500	1.2	1.2	2.4	3.3	4.8	5.5
5,000	0.6	0.8	1.7	2.3	3.4	3.9
10,000	0.3	0.5	1.2	1.7	2.4	2.8
25,000	0.12	0.3	0.8	1.0	1.5	1.7
50,000	0.06	0.2	0.5	0.7	1.1	1.2
75,000	0.04	0.2	0.4	0.6	0.9	1.0
100,000. . . .	0.03	0.2	0.4	0.5	0.8	0.9
150,000. . . .	0.02	0.14	0.3	0.4	0.6	0.7
200,000. . . .	0.02	0.12	0.3	0.4	0.5	0.6

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2 for the total SMSA and 1.3 for the central city estimates.

98,440 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 95,400 owner-occupied housing units, 31,600, or 33.1 percent, had two bedrooms. Interpolation in table II of the appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 33.1 percent is approximately 0.8 percentage points. The following interpolation procedure was used:

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

Base of percentage	Estimated percentage		
	25 or 75	33.1	50
75,000	0.9	a	1.0
95,400		p	
100,000. . . .	0.8	b	0.9

1. The entry for cell "a" is determined by horizontal interpolation between 0.9 and 1.0.

$$33.1-25.0 = 8.1$$
$$50.0-25.0 = 25.0$$
$$0.9 + \frac{8.1}{25.0}(1.0-0.9) = 0.93$$

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

$$33.1-25.0 = 8.1$$
$$50.0-25.0 = 25.0$$
$$0.8 + \frac{8.1}{25.0}(0.9-0.8) = 0.83$$

3. The entry for "p" was then determined by vertical interpolation between 0.93 and 0.83.

$$95{,}400-75{,}000 = 20{,}400$$
$$100{,}000-75{,}000 = 25{,}000$$
$$0.93 + \frac{20{,}400}{25{,}000}(0.83-0.93) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 32.3 to 33.9 percent; the 90-percent confidence interval is from 31.8 to 34.4 percent; and the 95-percent confidence interval is from 31.5 to 34.7 percent.

TABLE III. Standard Errors for Estimated Percentages of Housing Units in the 1977 Housing Inventory and for Estimated Percentages of 1974-1977 Lost Housing Units for the Balance (Not in Central City) of the Wichita, Kans., SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100.	25.4	25.4	25.4	25.4	25.4	29.2
200.	14.6	14.6	14.6	14.6	17.9	20.7
500.	6.4	6.4	6.4	7.8	11.3	13.1
700.	4.6	4.6	4.8	6.6	9.6	11.0
1,000	3.3	3.3	4.0	5.5	8.0	9.2
2,500	1.3	1.3	2.5	3.5	5.1	5.8
5,000	0.7	0.8	1.8	2.5	3.6	4.1
10,000	0.3	0.6	1.3	1.8	2.5	2.9
25,000	0.14	0.4	0.8	1.1	1.6	1.8
50,000	0.07	0.3	0.6	0.8	1.1	1.3

[1] Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.1 for the balance (not in central city) estimates.

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. However, 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 1974 and 1977 AHS-SMSA samples, a positive correlation should be expected when making comparisons between 1974 and 1977 characteristics.

Illustration of the computation of the standard error of a difference.—Table A-1 of part A of this report shows that in 1977 there were 46,000 owner-occupied units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied units with two bedrooms and owner-occupied units with three bedrooms is 14,400. Table I shows the standard error of 46,000 is approximately 1,120 and the standard error of 31,600 is approximately 930. Therefore, the standard error of the estimated difference of 14,400 is about:

$$1{,}460 = \sqrt{(1{,}120)^2 + (930)^2}$$

Consequently, the 68-percent confidence interval for the 14,400 difference is from 12,940 to 15,860 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 confi-

dence interval is from 12,064 to 16,736 housing units, and the 95-percent confidence interval is from 11,480 to 17,320. Thus, we can conclude with 95 percent confidence that the number of 1977 owner-occupied 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, read off the confidence interval corresponding to the two points established in step 2.

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 95,400 housing units.

1. Interpolation in table II shows that the standard error of 50 percent on a base of 95,400 is approximately 0.9 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50 percent twice the standard error determined in step 1. This yields percentage limits of 48.2 and 51.8.
3. From table A-1 of part A, it can be seen by cumulating the frequencies for the first two categories that 45,700 owner-occupied housing units, or 47.9 percent, had one or two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) and that an additional 18,200 owner-occupied housing units, or 19.1 percent, had three persons (i.e., 2.5 to 3.5 persons).

By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{48.2-47.9}{19.1}\right) = 2.5$$

Similarly, the upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\left(\frac{51.8-47.9}{19.1}\right) = 2.7$$

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