1982 Metropolitan Appendix B: Source and Accuracy

Table of Contents

Atlanta 3

Cincinnati 14

Columbus 25

Kansas City 36

New Orleans 47

Paterson-Clifton-Passaic 58

Philadelphia 70

Riverside-San Bernardino-Ontario 81

Rochester 93

San Antonio 104

San Diego 115

San Francisco-Oakland 126

Atlanta 1982

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1982 survey App-42
Selection of the 1978 AHS-SMSA sample App-43
Sample selection for the 1978 Coverage Improvement Program . App-44
Coverage Improvement for deficiency 1 ..... App-44
Coverage Improvement for deficiency 2 ..... App-44
Coverage Improvement for deficiencies 3-6 ..... App-44
1978-1982 additions to the housing inventory App-45
1970 Census of Population and Housing App-45
ESTIMATION App-45
1982 housing inventory App-45
1978-1982 lost housing units ... App-46
1978 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-47
Rounding errors App-47
Sampling errors for the AHS-SMSA sample App-47
Illustration of the use of the standard error tables App-48
Differences App-49
Illustration of the computation of the standard error of a difference App-49
Medians App-49
Illustration of the computation of the 95 percent confidence interval of a median ... App-50
Standard error tables App-51

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 12 SMSA's in this report series (H-170-82) are based on data collected from the 1982 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 12 SMSA's selected for interview during 1982 were interviewed previously in 1975 and 1978 (see the list of SMSA reports from the AHS in the introduction of this report).

Three of the larger SMSA's (Atlanta, GA; Philadelphia, PA-NJ; and San Francisco-Oakland, CA) in the 1982 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1975 and 1978 surveys, evenly divided between the central city and the balance of the respective SMSA. The 9 remaining SMSA's (Cincinnati, OH-KY-IN; Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Paterson-Clifton-Passaic, NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; and San Diego, CA) in the 1982 group were represented by a sample of 5,000 designated housing units in the 1975 and 1978 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1975 and 1978 surveys, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panels 1 through 3 were dropped from all SMSA's for the 1982 survey, excluding the central city of the Paterson, NJ, SMSA. Due to additional budget limitations for the 1982 survey, it also became necessary to drop part of the sample of housing units from panels 4 through 12 in the Atlanta, GA; New Orleans, LA; Philadelphia, PA-NJ; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. The effect of this reduction was to maintain a sample size of 4,250 in all SMSA's. The interviewing was done during the period April 1982 through December 1982 with one-ninth of the sample housing units being visited each month.

In this SMSA, 4,062 housing units were eligible for interview. Of these sample housing units, 174 interviews were not obtained because, for occupied sample units, the occupants refused to be interviewed, were not at home after repeated visits, or were unavailable for some other reason; or, for vacant units, no informed respondent could be found after repeated visits. In addition to the 4,062 housing units eligible for interview, 185 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 1982 survey—The sample housing units designated to be interviewed in the 1982 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 1978 survey and remained in sample after the 1982 panel reduction.
2. All sample housing units that were type A noninterviews (i.e., units eligible to be interviewed) or type B noninterviews (i.e., units not eligible for interview at the time of the survey but which could become eligible in the future) in the 1978 survey and remained in sample after the 1982 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1982 AHS questionnaire, page App-18.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented the housing units built in permit-issuing areas since the 1978 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1978 survey.)
5. All sample housing units that were selected as part of the 1978 Coverage Improvement Program. (This sample represented most of the housing units which, until 1978, did not have a chance of selection.)

Selection of the 1978 AHS-SMSA sample—The sample for the SMSA's which, in 1970, were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). In 1970, the following four SMSA's were 100-percent permit-issuing: Paterson-Clifton-Passaic, NJ; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA. The remaining eight SMSA's contain a sample from the nonpermit universe.

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

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA during the 1970 Census of Population and Housing. This file contained records for occupied housing units, vacant housing units, and housing units in certain special places or group quarters. Sampling operations were done separately for the special place and group quarters records, and for the occupied and vacant housing unit records. Before the sample was selected from the occupied and vacant housing unit records, the occupied housing unit records were stratified by race of head (non-Black/Black) and the vacant records were stratified into four categories pertaining to the value or rent associated with the vacant housing units. The occupied housing unit records were further stratified so that each unit was assigned to 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household income	Tenure	
	Owner—Family size	Renter—Family size
	1 2 3 4 5+	1 2 3 4 5+
Under $3,000		
$3,000 to $5,999 . . .		
$6,000 to $9,999 . . .		
$10,000 to $14,999 .		
$15,000 and over . . .		

Thus, for this SMSA, the occupied housing unit records from the permit-issuing universe were assigned to 1 of 100 strata for either the central city or for the balance, and the vacant housing unit records were assigned to 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. However, whenever a record was selected to be in sample, the housing unit record adjacent to it on the file was also selected to be in sample, thereby insuring the necessary designated sample size.

Before the sample was selected from the group quarters and special place records, the records were stratified by census tract and census enumeration district (ED) within the central city and within the balance of the SMSA. A sample of special place records was then selected by a procedure that produced one-quarter of the desired sample size. However, at the time of the survey, the housing units at each of the special places were listed and subsampled at a rate which produced an expected four sample units, thereby insuring the necessary designated sample size.

The second frame from which this SMSA sample was selected was a list of new construction building permits issued since 1970 (i.e., the new construction universe). The sample selection from the list of new construction building permits was an independent operation within this SMSA. Prior to sample selection, the

list of permits was chronologically stratified by the date the permits were issued, and clusters of an expected four (usually adjacent) housing units were formed. These clusters were then sampled for inclusion at the overall sampling rate.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS sample was selected from a frame consisting of areas not under the jurisdiction of permit-issuing offices (i.e., the nonpermit universe). The first step in the sampling operation for the nonpermit universe was the selection (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

The sample ED's were then divided into segments; i.e., small land areas with well-defined boundaries having an expected size of four, or a multiple of four, housing units. At the time of the survey, those segments that did not have an expected size of four were further subdivided to produce an expected four sample housing units.

The next step was the selection of one of these segments within each sample ED. All housing units in existence at the time of interview in these selected segments were eligible for sample. Thus, housing units enumerated in the 1970 census as well as housing units built since the 1970 census are included.

Sample Selection for the 1978 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

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

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 SMSA. 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. 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 Rochester, NY, SMSA.

In the remaining 11 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. These procedures added an estimated 11,401 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 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 1,567 units to the coverage of the housing inventory of this SMSA.

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

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

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample unit were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and units within these structures were interviewed.

In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. This procedure added an estimated 14,332 units to the coverage of the housing inventory of this SMSA.

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

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 1,202 units to the coverage of the housing inventory of this SMSA.

1978-1982 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1978 survey, was selected to represent housing units built in permit-issuing areas since the 1978 survey. Sampling procedures were identical to those used in selecting the 1970-1978 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1977, to identify any housing units missed in the 1978 survey or any housing units added since the 1978 survey.

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 1982 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 1982 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1978 (i.e., 1978-1982 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1982 housing inventory—The AHS estimates of characteristics of the 1982 housing inventory were produced using a one-stage ratio estimation procedure for the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's. A two-stage ratio estimation procedure was used for the Columbus, OH; Kansas City, KS-MO; New Orleans, LA; Philadelphia, PA-NJ; San Antonio, TX; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. A three-stage ratio estimation procedure was used for the San Bernardino-Riverside-Ontario, CA, SMSA.

Prior to the implementation of the ratio estimation procedures, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

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

The following ratio estimation procedure was employed for all sample housing units from the permit-issuing universe. This factor was computed separately for all sample housing units within each permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

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

This ratio estimation procedure was introduced to correct the probabilities of selection for samples in each of the strata used in the sample selection of the permit-issuing universe. Prior to the AHS sample selection within each SMSA, housing units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. The same probability of selection was then applied to the remaining units to select the AHS sample. Since the number of housing units deleted from the AHS

universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the San Bernardino-Riverside-Ontario, CA, SMSA, the next ratio estimation procedure was employed for new construction mobile home units. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of mobile home units for each sector of the SMSA (central city and balance) by year built groups (1970-1974 and 1975-1982) to an independent estimate of mobile home units for the corresponding sector and year built group. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of mobile homes for the corresponding sector and year built group}}{\text{Sample estimate of mobile homes for the corresponding sector and year built group}}$$

The computed ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell.

The next ratio estimation procedure was applied in the Columbus, OH; Kansas City, KS-MO; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1982 housing inventory in each sector of the SMSA (central city and balance) to an independent estimate of the total housing units in each sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1982 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for the 12 SMSA's in 1982. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1982 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units using the existing sample weights were compared to the corresponding independent estimates for the central city, balance, and total for each of the 12 SMSA's, and the estimates which showed the most likely level of net growth since the 1980 census in both the central city and balance, as well as, the total SMSA were used in this ratio estimation procedure.

As a result of this analysis, the independent estimates were used in the Kansas City, KS-MO; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco, CA, SMSA's. The sample estimates were used in the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's.

For the Columbus, OH and San Antonio, TX, SMSA's the independent estimates were used for the total SMSA only. The independent estimates were then proportioned between the central city and the balance of each SMSA according to the central city/balance distribution of the total housing units as given by the sample estimates. Due to the central city boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimates did not reflect the 1970 central city/balance SMSA definitions. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definitions were used.

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

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

1978-1982 lost housing units—The 1978-1982 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1978 housing inventory as was described in the 1978 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas*. Since the 1978-1982 lost housing units existed, by definition, in the 1978 housing inventory, there was a 1978 housing inventory weight associated with each 1978-1982 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1978-1982 lost housing units.

1978 estimation procedure—This report presents data on the housing characteristics of the 1978 housing inventory from the 1978 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1978.

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

RELIABILITY OF THE ESTIMATES

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

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

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing and the 1982 AHS-SMSA sample.

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

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

AHS-SMSA—The AHS-SMSA evaluation study was administered to a sample of units selected from the regular interviews containing recent movers who had previous residences within the SMSA. The purpose of the study was to evaluate the accuracy of responses regarding housing characteristics of recent mover's previous residence. The results of which are presented in the Census Bureau memorandum "Analysis of the Results from the AHS-SMSA Group AA-2 1982 Evaluation Study." (June 17, 1983)

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

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

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all housing units located inside these ED's would be represented in the sample. However, it has been estimated that the 1975 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. It should be noted that since these ED's were recanvassed for the 1978 and 1982 surveys, the number of missed housing units may be considerably less for 1982.

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

Rounding errors—For errors associated with processing, rounding of estimates introduces another source of error in the data, the severity of which depends on the statistics being measured. The effect of rounding is significant relative to the sampling error only for small percentages or small medians, when these figures are derived from relatively large bases (e.g., median number of persons per household). This means that confidence intervals formed from the standard errors given may be distorted, and this should be taken into account when considering the results of the survey.

Sampling errors for the AHS-SMSA sample—The particular sample used for this survey is one of a large number of possible samples of the same size that could have been selected using the same sample design. Even if the same questionnaires, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling

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

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

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

The figures presented in the tables that follow (page App-51) are approximations to the standard errors of various estimates shown in this report for this SMSA. In order to derive standard errors that would be applicable to a wide variety of items and also could be prepared at a moderate cost, a number of approximations were required. As a result, the tables of standard errors provide an indication of the order of magnitude of the standard errors rather than precise standard errors for any specific item. Standard errors applicable to estimates of characteristics of the 1978 housing inventory can be found in the AHS Series H-170 reports for 1978.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1982 housing inventory as well as estimates of characteristics of the 1978-1982 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 30,500 for the total SMSA, 17,230 for the central city of the SMSA, and 25,210 for the balance of the SMSA.

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

Table II (page App-51) presents the standard errors of estimated percentages for the 1982 housing inventory as well as estimated percentages of the 1978-1982 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

Included in tables I and II are estimates of standard errors for estimates of zero and zero percent. These estimates of standard errors are considered as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is obtained.

For ratios, 100 (x/y), where x is not a subclass of y, table II underestimates the standard error of the ratio when there is little or no correlation between x and y. For this type of ratio, a better approximation of the standard error may be obtained by letting the standard error of the ratio be approximately equal to:

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

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

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1982 there were 374,200 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 7,520. 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
300,000	6,930
374,200	x
400,000	7,730

The entry of "x" is determined as follows by vertically interpolating between 6,930 and 7,730.

$$374{,}200 - 300{,}000 = 74{,}200$$
$$400{,}000 - 300{,}000 = 100{,}000$$

$$6{,}930 + \frac{74{,}200}{100{,}000}\,(7{,}730 - 6{,}930) = 7{,}520$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 366,680 to 381,720 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1982 owner-occupied housing units lies within a range computed in this way would be correct for roughly

68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 362,170 to 386,230 housing units with 90 percent confidence; and that the average estimate lies within the interval from 359,160 to 389,240 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 374,200 owner-occupied housing units, 79,600, or 21.3 percent, had two bedrooms. Interpolation using table II of this appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 21.3 percent is approximately 0.9 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	21.3	25 or 75
300,000	0.8	a	1.1
374,200		p	
400,000	0.7	b	1.0

1. The entry for cell "a" is determined by horizontal interpolation between 0.8 and 1.1.

$$21.3-10.0=11.3$$
$$25.0-10.0=15.0$$

$$0.8+\frac{11.3}{15.0}\ (1.1-0.8)=1.0$$

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

$$21.3-10.0=11.3$$
$$25.0-10.0=15.0$$

$$0.7+\frac{11.3}{15.0}\ (1.0-0.7)=0.9$$

3. The entry for "p" is then determined by vertical interpolation between1.0 and 0.9.

$$374{,}200-300{,}000=74{,}200$$
$$400{,}000-300{,}000=100{,}000$$

$$1.0+\frac{74{,}200}{100{,}000}\ (0.9-1.0)=0.9$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 20.4 to 22.2 percent; the 90-percent confidence interval is from 19.9 to 22.7 percent; and the 95-percent confidence interval is from 19.5 to 23.1 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. If there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1978 and 1982 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1978 and 1982 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1982 there were 194,100 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 114,500. Table I shows the standard error of 79,600 is approximately 3,820, and the standard error of 194,100 is approximately 5,770. Therefore, the standard error of the estimated difference of 114,500 is about 6,920.

$$6{,}920=\sqrt{(3{,}820)^2+(5{,}770)^2}$$

Consequently, the 68-percent confidence interval for the 114,500 difference is from 107,580 to 121,420 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 103,430 to 125,570 housing units, and the 95-percent confidence interval is from 100,660 to 128,340 housing units. Thus, we can conclude with 95 percent confidence that the number of 1982 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied units with two bedrooms since the 95-percent confidence interval does not include zero or negative values.

Medians—For medians presented in certain tables, the sampling error depends on the size of the base and on the distribution upon which the median is based. An approximate method for measuring the reliability of the estimated median is to determine an interval about the estimated median so that there is a stated degree of confidence that the average median from all possible samples lies within the interval. The following procedure may be used to estimate confidence limits of a median based on sample data:

1. From table II, determine the standard error of a 50-percent characteristic on the base of the median.
2. Add to and subtract from 50 percent, the standard error determined in step 1.
3. Using the distribution of the characteristics, determine the confidence interval corresponding to the two points established in step 2. To find the lower endpoint of the confidence

interval, it is necessary to know into which interval of the distribution the lower percentage limit falls. Similarly, to find the upper endpoint to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median— Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.8. The base of the distribution from which this median was determined is 374,200 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 374,200 is approximately 1.2 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50-percent twice the standard error determined in step 1. This yields percentage limits of 47.6 and 52.4.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with three persons (for purposes of calculating the median, the category of three persons is considered to be from 2.5 to 3.5 persons) contains the 47.6 percent derived in step 2. About 163,100 housing units or 43.6 percent fall below this interval, and 77,600 housing units or 20.7 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(47.6-43.6)}{20.7} = 2.7$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.4 percent derived in step 2. About 163,100 housing units or 43.6 percent fall below this interval, and 77,600 housing units or 20.7 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(52.4-43.6)}{20.7} = 2.9$$

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

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1982 Housing Inventory and for Estimated Number of 1978-1982 Lost Units for the Atlanta, GA, 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]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	190	180	180	100,000	4,260	3,670	4,130
100	190	180	180	150,000	5,140	4,080	4,970
200	200	190	190	179,700	5,580	4,180	5,380
500	310	300	300	200,000	5,850	—	5,630
700	370	360	360	250,000	6,430	—	6,170
1,000	440	430	430	300,000	6,930	—	6,620
2,500	690	670	680	400,000	7,730	—	7,320
5,000	980	950	950	500,000	8,320	—	7,800
10,000	1,380	1,330	1,350	512,900	8,390	—	7,840
25,000	2,180	2,060	2,120	600,000	8,750	—	—
50,000	3,060	2,810	2,970	692,500	9,030	—	—
75,000	3,720	3,310	3,610				

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

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1982 Housing Inventory and for Estimated Percentages of 1978-1982 Lost Housing Units for the Atlanta, GA, SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	65.8	65.8	65.8	65.8	65.8	69.4	75,000	0.3	0.5	1.1	1.5	2.2	2.5
200	49.0	49.0	49.0	49.0	49.0	49.1	100,000	0.2	0.4	1.0	1.3	1.9	2.2
500	27.8	27.8	27.8	27.8	27.8	31.0	150,000	0.13	0.4	0.8	1.1	1.6	1.8
700	21.6	21.6	21.6	21.6	22.7	26.2	200,000	0.10	0.3	0.7	0.9	1.3	1.6
1,000	16.1	16.1	16.1	16.1	19.0	21.9	250,000	0.08	0.3	0.6	0.8	1.2	1.4
2,500	7.2	7.2	7.2	8.3	12.0	13.9	300,000	0.06	0.3	0.6	0.8	1.1	1.3
5,000	3.7	3.7	4.3	5.9	8.5	9.8	400,000	0.05	0.2	0.5	0.7	1.0	1.1
10,000	1.9	1.9	3.0	4.2	6.0	6.9	500,000	0.04	0.2	0.4	0.6	0.9	1.0
25,000	0.8	0.9	1.9	2.6	3.8	4.4	600,000	0.03	0.2	0.4	0.5	0.8	0.9
50,000	0.4	0.6	1.4	1.9	2.7	3.1	700,000	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 1.3 for the total SMSA and 1.2 for the central city and balance (not in central city) estimates.

Cincinnati 1982

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1982 survey App-42
Selection of the 1978 AHS-SMSA sample App-43
Sample selection for the 1978 Coverage Improvement Program . App-44
Coverage Improvement for deficiency 1 ..... App-44
Coverage Improvement for deficiency 2 ..... App-44
Coverage Improvement for deficiencies 3-6 ..... App-44
1978-1982 additions to the housing inventory App-45
1970 Census of Population and Housing App-45
ESTIMATION App-45
1982 housing inventory App-45
1978-1982 lost housing units ... App-46
1978 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-47
Rounding errors App-47
Sampling errors for the AHS-SMSA sample App-47
Illustration of the use of the standard error tables App-48
Differences App-49
Illustration of the computation of the standard error of a difference App-49
Medians App-49
Illustration of the computation of the 95-percent confidence interval of a median ... App-50
Standard error tables App-51

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 12 SMSA's in this report series (H-170-82) are based on data collected from the 1982 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 12 SMSA's selected for interview during 1982 were interviewed previously in 1975 and 1978 (see the list of SMSA reports from the AHS in the introduction of this report).

Three of the larger SMSA's (Atlanta, GA; Philadelphia, PA-NJ; and San Francisco-Oakland, CA) in the 1982 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1975 and 1978 surveys, evenly divided between the central city and the balance of the respective SMSA. The 9 remaining SMSA's (Cincinnati, OH-KY-IN; Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Paterson-Clifton-Passaic, NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; and San Diego, CA) in the 1982 group were represented by a sample of 5,000 designated housing units in the 1975 and 1978 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1975 and 1978 surveys, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panels 1 through 3 were dropped from all SMSA's for the 1982 survey, excluding the central cities of the Paterson-Clifton-Passaic, NJ, SMSA. Due to additional budget limitations for the 1982 survey, it also became necessary to drop part of the sample of housing units from panels 4 through 12 in the Atlanta, GA; New Orleans, LA; Philadelphia, PA-NJ; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. The effect of this reduction was to maintain a sample size of 4,250 in all SMSA's. The interviewing was done during the period April 1982 through December 1982 with one-ninth of the sample housing units being visited each month.

In this SMSA, 3,989 housing units were eligible for interview. Of these sample housing units, 120 interviews were not obtained because, for occupied sample units, the occupants refused to be interviewed, were not at home after repeated visits, or were unavailable for some other reason; or, for vacant units, no informed respondent could be found after repeated visits. In addition to the 3,989 housing units eligible for interview, 239 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 1982 survey—The sample housing units designated to be interviewed in the 1982 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 1978 survey and remained in sample after the 1982 panel reduction.
2. All sample housing units that were type A noninterviews (i.e., units eligible to be interviewed) or type B noninterviews (i.e., units not eligible for interview at the time of the survey but which could become eligible in the future) in the 1978 survey and remained in sample after the 1982 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1982 AHS questionnaire, page App-18.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented the housing units built in permit-issuing areas since the 1978 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1978 survey.)
5. All sample housing units that were selected as part of the 1978 Coverage Improvement Program. (This sample represented most of the housing units which, until 1978, did not have a chance of selection.)

Selection of the 1978 AHS-SMSA sample—The sample for the SMSA's which, in 1970, were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). In 1970, the following four SMSA's were 100-percent permit-issuing: Paterson-Clifton-Passaic, NJ; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA. The remaining eight SMSA's contain a sample from the nonpermit universe.

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

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA during the 1970 Census of Population and Housing. This file contained records for occupied housing units, vacant housing units, and housing units in certain special places or group quarters. Sampling operations were done separately for the special place and group quarters records, and for the occupied and vacant housing unit records. Before the sample was selected from the occupied and vacant housing unit records, the occupied housing unit records were stratified by race of head (non-Black/Black) and the vacant records were stratified into four categories pertaining to the value or rent associated with the vacant housing units. The occupied housing unit records were further stratified so that each unit was assigned to 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household income	Tenure	
	Owner—Family size 1 2 3 4 5+	Renter—Family size 1 2 3 4 5+
Under $3,000		
$3,000 to $5,999 . . .		
$6,000 to $9,999 . . .		
$10,000 to $14,999 .		
$15,000 and over . . .		

Thus, for this SMSA, the occupied housing unit records from the permit-issuing universe were assigned to 1 of 100 strata for either the central city or for the balance, and the vacant housing unit records were assigned to 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. However, whenever a record was selected to be in sample, the housing unit record adjacent to it on the file was also selected to be in sample, thereby insuring the necessary designated sample size.

Before the sample was selected from the group quarters and special place records, the records were stratified by census tract and census enumeration district (ED) within the central city and within the balance of the SMSA. A sample of special place records was then selected by a procedure that produced one-quarter of the desired sample size. However, at the time of the survey, the housing units at each of the special places were listed and subsampled at a rate which produced an expected four sample units, thereby insuring the necessary designated sample size.

The second frame from which this SMSA sample was selected was a list of new construction building permits issued since 1970 (i.e., the new construction universe). The sample selection from the list of new construction building permits was an independent operation within this SMSA. Prior to sample selection, the

list of permits was chronologically stratified by the date the permits were issued, and clusters of an expected four (usually adjacent) housing units were formed. These clusters were then sampled for inclusion at the overall sampling rate.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS sample was selected from a frame consisting of areas not under the jurisdiction of permit-issuing offices (i.e., the nonpermit universe). The first step in the sampling operation for the nonpermit universe was the selection (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

The sample ED's were then divided into segments; i.e., small land areas with well-defined boundaries having an expected size of four, or a multiple of four, housing units. At the time of the survey, those segments that did not have an expected size of four were further subdivided to produce an expected four sample housing units.

The next step was the selection of one of these segments within each sample ED. All housing units in existence at the time of interview in these selected segments were eligible for sample. Thus, housing units enumerated in the 1970 census as well as housing units built since the 1970 census are included.

Sample Selection for the 1978 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

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

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 SMSA. 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. 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 Rochester, NY, SMSA.

In the remaining 11 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. These procedures added an estimated 2,491 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 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 2,257 units to the coverage of the housing inventory of this SMSA.

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

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

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample unit were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and units within these structures were interviewed.

In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. This procedure added an estimated 14,610 units to the coverage of the housing inventory of this SMSA.

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

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 4,993 units to the coverage of the housing inventory of this SMSA.

1978-1982 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1978 survey, was selected to represent housing units built in permit-issuing areas since the 1978 survey. Sampling procedures were identical to those used in selecting the 1970-1978 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1977, to identify any housing units missed in the 1978 survey or any housing units added since the 1978 survey.

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 1982 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 1982 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1978 (i.e., 1978-1982 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1982 housing inventory—The AHS estimates of characteristics of the 1982 housing inventory were produced using a one-stage ratio estimation procedure for the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's. A two-stage ratio estimation procedure was used for the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. A three-stage ratio estimation procedure was used for the San Bernardino-Riverside-Ontario, CA, SMSA.

Prior to the implementation of the ratio estimation procedures, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

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

The following ratio estimation procedure was employed for all sample housing units from the permit-issuing universe. This factor was computed separately for all sample housing units within each permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

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

This ratio estimation procedure was introduced to correct the probabilities of selection for samples in each of the strata used in the sample selection of the permit-issuing universe. Prior to the AHS sample selection within each SMSA, housing units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. The same probability of selection was then applied to the remaining units to select the AHS sample. Since the number of housing units deleted from the AHS

universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the San Bernardino-Riverside-Ontario, CA, SMSA, the next ratio estimation procedure was employed for new construction mobile home units. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of mobile home units for each sector of the SMSA (central city and balance) by year built groups (1970-1974 and 1975-1982) to an independent estimate of mobile home units for the corresponding sector and year built group. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of mobile homes for the corresponding sector and year built group}}{\text{Sample estimate of mobile homes for the corresponding sector and year built group}}$$

The computed ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell.

The next ratio estimation procedure was applied in the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1982 housing inventory in each sector of the SMSA (central city and balance) to an independent estimate of the total housing units in each sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1982 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for the 12 SMSA's in 1982. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1982 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units using the existing sample weights were compared to the corresponding independent estimates for the central city, balance, and total for each of the 12 SMSA's, and the estimates which showed the most likely level of net growth since the 1980 census in both the central city and balance, as well as, the total SMSA were used in this ratio estimation procedure.

As a result of this analysis, the independent estimates were used in the Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco, CA, SMSA's. The sample estimates were used in the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's.

For the Columbus, OH and San Antonio, TX, SMSA's the independent estimates were used for the total SMSA only. The independent estimates were then proportioned between the central city and the balance of each SMSA according to the central city/balance distribution of the total housing units as given by the sample estimates. Due to the central city boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimates did not reflect the 1970 central city/balance SMSA definitions. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definitions were used.

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

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

1978-1982 lost housing units—The 1978-1982 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1978 housing inventory as was described in the 1978 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas*. Since the 1978-1982 lost housing units existed, by definition, in the 1978 housing inventory, there was a 1978 housing inventory weight associated with each 1978-1982 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1978-1982 lost housing units.

1978 estimation procedure—This report presents data on the housing characteristics of the 1978 housing inventory from the 1978 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1978.

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

RELIABILITY OF THE ESTIMATES

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

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

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing and the 1982 AHS-SMSA sample.

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

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

AHS-SMSA—The AHS-SMSA evaluation study was administered to a sample of units selected from the regular interviews containing recent movers who had previous residences within the SMSA. The purpose of the study was to evaluate the accuracy of responses regarding housing characteristics of recent mover's previous residence. The results of which are presented in the Census Bureau memorandum "Analysis of the Results from the AHS-SMSA Group AA-2 1982 Evaluation Study" (June 17, 1983).

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

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

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all housing units located inside these ED's would be represented in the sample. However, it has been estimated that the 1975 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. It should be noted that since these ED's were recanvassed for the 1978 and 1982 surveys, the number of missed housing units may be considerably less for 1982.

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

Rounding errors—For errors associated with processing, rounding of estimates introduces another source of error in the data, the severity of which depends on the statistics being measured. The effect of rounding is significant relative to the sampling error only for small percentages or small medians, when these figures are derived from relatively large bases (e.g., median number of persons per household). This means that confidence intervals formed from the standard errors given may be distorted, and this should be taken into account when considering the results of the survey.

Sampling errors for the AHS-SMSA sample—The particular sample used for this survey is one of a large number of possible samples of the same size that could have been selected using the same sample design. Even if the same questionnaires, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling

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

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

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

The figures presented in the tables that follow (page App-51) are approximations to the standard errors of various estimates shown in this report for this SMSA. In order to derive standard errors that would be applicable to a wide variety of items and also could be prepared at a moderate cost, a number of approximations were required. As a result, the tables of standard errors provide an indication of the order of magnitude of the standard errors rather than precise standard errors for any specific item. Standard errors applicable to estimates of characteristics of the 1978 housing inventory can be found in the AHS Series H-170 reports for 1978.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1982 housing inventory as well as estimates of characteristics of the 1978-1982 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 20,810 for the total SMSA, 9,550 for the central city of the SMSA, and 18,490 for the balance of the SMSA.

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

Table II (page App-51) presents the standard errors of estimated percentages for the 1982 housing inventory as well as estimated percentages of the 1978-1982 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

Included in tables I and II are estimates of standard errors for estimates of zero and zero percent. These estimates of standard errors are considered as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is obtained.

For ratios, 100 (x/y), where x is not a subclass of y, table II underestimates the standard error of the ratio when there is little or no correlation between x and y. For this type of ratio, a better approximation of the standard error may be obtained by letting the standard error of the ratio be approximately equal to:

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

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

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1982 there were 316,500 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 5,360. The following interpolation procedure was used.

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

Size of estimate	Standard error
300,000	5,320
316,500	x
377,000	5,500

The entry of "x" is determined as follows by vertically interpolating between 5,320 and 5,500.

$$316{,}500-300{,}000 = 16{,}500$$
$$377{,}000-300{,}000 = 77{,}000$$

$$5{,}320 + \frac{16{,}500}{77{,}000}(5{,}500-5{,}320) = 5{,}360$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 311,140 to 321,860 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1982 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 307,920 to 325,080 housing units with 90 percent confidence; and that the average estimate lies within the interval from 305,780 to 327,220 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 316,500 owner-occupied housing units, 77,900, or 24.6 percent, had two bedrooms. Interpolation using table II of this appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 24.6 percent is approximately 1.0 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	24.6	25 or 75
300,000	0.7	a	1.0
316,500		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.

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

$$0.7 + \frac{14.6}{15.0}\ (1.0-0.7) = 1.0$$

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

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

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

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

$$316{,}500-300{,}000 = 16{,}500$$
$$400{,}000-300{,}000 = 100{,}000$$

$$1.0 + \frac{16{,}500}{100{,}000}\ (0.8-1.0) = 1.0$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 23.6 to 25.6 percent; the 90-percent confidence interval is from 23.0 to 26.2 percent; and the 95-percent confidence interval is from 22.6 to 26.6 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. If there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1978 and 1982 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1978 and 1982 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1982 there were 148,500 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 70,600. Table I shows the standard error of 77,900 is approximately 3,240, and the standard error of 148,500 is approximately 4,250. Therefore, the standard error of the estimated difference of 70,600 is about 5,340.

$$5{,}340 = \sqrt{(3{,}240)^2 + (4{,}250)^2}$$

Consequently, the 68-percent confidence interval for the 70,600 difference is from 65,260 to 75,940 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 62,060 to 79,140 housing units, and the 95-percent confidence interval is from 59,920 to 81,280 housing units. Thus, we can conclude with 95 percent confidence that the number of 1982 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied units with two bedrooms since the 95-percent confidence interval does not include zero or negative values.

Medians—For medians presented in certain tables, the sampling error depends on the size of the base and on the distribution upon which the median is based. An approximate method for measuring the reliability of the estimated median is to determine an interval about the estimated median so that there is a stated degree of confidence that the average median from all possible samples lies within the interval. The following procedure may be used to estimate confidence limits of a median based on sample data:

1. From table II, determine the standard error of a 50-percent characteristic on the base of the median.
2. Add to and subtract from 50 percent, the standard error determined in step 1.
3. Using the distribution of the characteristics, determine the confidence interval corresponding to the two points established in step 2. To find the lower endpoint of the confidence

interval, it is necessary to know into which interval of the distribution the lower percentage limit falls. Similarly, to find the upper endpoint to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.8. The base of the distribution from which this median was determined is 316,500 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 316,500 is approximately 1.1 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50-percent twice the standard error determined in step 1. This yields percentage limits of 47.8 and 52.2.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with three persons (for purposes of calculating the median, the category of three persons is considered to be from 2.5 to 3.5 persons) contains the 47.8 percent derived in step 2. About 141,400 housing units or 44.7 percent fall below this interval, and 54,900 housing units or 17.3 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(47.8-44.7)}{17.3} = 2.7$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.2 percent derived in step 2. About 141,400 housing units or 44.7 percent fall below this interval, and 54,900 housing units or 17.3 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\,\frac{(52.2-44.7)}{17.3} = 2.9$$

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

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1982 Housing Inventory and for Estimated Number of 1978-1982 Lost Units for the Cincinnati, OH-KY-IN, SMSA, for the Central City of the SMSA and for the Balance (Not in Central Cities) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	150	140	150	75,000	3,190	2,860	3,110
100	150	140	150	100,000	3,620	3,120	3,490
200	170	170	170	150,000	4,270	3,330	4,020
500	270	270	270	170,300	4,480	3,310	4,160
700	320	320	320	200,000	4,740	—	4,320
1,000	390	380	390	250,000	5,080	—	4,440
2,500	610	600	610	300,000	5,320	—	4,400
5,000	860	840	860	377,000	5,500	—	4,000
10,000	1,210	1,180	1,220	400,000	5,510	—	—
25,000	1,900	1,820	1,890	500,000	5,370	—	—
50,000	2,650	2,460	2,610	547,300	5,180	—	—

[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 balance (not in central city) and 1.4 for the central city estimates.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1982 Housing Inventory and for Estimated Percentages of 1978-1982 Lost Housing Units for the Cincinnati, OH-KY-IN, SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
200	43.0	43.0	43.0	43.0	43.0	43.4	75,000	0.2	0.4	1.0	1.3	1.9	2.2
500	23.1	23.1	23.1	32.1	23.8	27.4	100,000	0.2	0.4	0.8	1.2	1.7	1.9
700	17.7	17.7	17.7	17.7	20.1	23.2	150,000	0.10	0.3	0.7	1.0	1.4	1.6
1,000	13.1	13.1	13.1	13.1	16.8	19.4	200,000	0.08	0.3	0.6	0.8	1.2	1.4
2,500	5.7	5.7	5.7	7.4	10.6	12.3	250,000	0.06	0.2	0.5	0.7	1.1	1.2
5,000	2.9	2.9	3.8	5.2	7.5	8.7	300,000	0.05	0.2	0.5	0.7	1.0	1.1
10,000	1.5	1.5	2.7	3.7	5.3	6.1	400,000	0.04	0.2	0.4	0.6	0.8	1.0
25,000	0.6	0.8	1.7	2.3	3.4	3.9	500,000	0.03	0.2	0.4	0.5	0.8	0.9
50,000	0.3	0.5	1.2	1.6	2.4	2.7	600,000	0.03	0.2	0.3	0.5	0.7	0.8

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

Columbus 1982

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1982 survey App-42
Selection of the 1978 AHS-SMSA sample App-43
Sample selection for the 1978 Coverage Improvement Program . App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1978-1982 additions to the housing inventory App-45
1970 Census of Population and Housing App-45
ESTIMATION App-45
1982 housing inventory App-45
1978-1982 lost housing units ... App-46
1978 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-47
Rounding errors App-47
Sampling errors for the AHS-SMSA sample App-47
Illustration of the use of the standard error tables App-48
Differences App-49
Illustration of the computation of the standard error of a difference App-49
Medians App-49
Illustration of the computation of the 95 percent confidence interval of a median ... App-50
Standard error tables App-51

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 12 SMSA's in this report series (H-170-82) are based on data collected from the 1982 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 12 SMSA's selected for interview during 1982 were interviewed previously in 1975 and 1978 (see the list of SMSA reports from the AHS in the introduction of this report).

Three of the larger SMSA's (Atlanta, GA; Philadelphia, PA-NJ; and San Francisco-Oakland, CA) in the 1982 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1975 and 1978 surveys, evenly divided between the central city and the balance of the respective SMSA. The 9 remaining SMSA's (Cincinnati, OH-KY-IN; Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Paterson-Clifton-Passaic, NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; and San Diego, CA) in the 1982 group were represented by a sample of 5,000 designated housing units in the 1975 and 1978 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1975 and 1978 surveys, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panels 1 through 3 were dropped from all SMSA's for the 1982 survey, excluding the central cities of the Paterson-Clifton-Passaic, NJ, SMSA. Due to additional budget limitations for the 1982 survey, it also became necessary to drop part of the sample of housing units from panels 4 through 12 in the Atlanta, GA; New Orleans, LA; Philadelphia, PA-NJ; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. The effect of this reduction was to maintain a sample size of 4,250 in all SMSA's. The interviewing was done during the period April 1982 through December 1982 with one-ninth of the sample housing units being visited each month.

In this SMSA, 4,034 housing units were eligible for interview. Of these sample housing units, 157 interviews were not obtained because, for occupied sample units, the occupants refused to be interviewed, were not at home after repeated visits, or were unavailable for some other reason; or, for vacant units, no informed respondent could be found after repeated visits. In addition to the 4,034 housing units eligible for interview, 216 units were visited but were not eligible for interview because they were condemned, unfit, demolished, converted to group quarters use, etc.

Designation of sample housing units for the 1982 survey—The sample housing units designated to be interviewed in the 1982 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 1978 survey and remained in sample after the 1982 panel reduction.
2. All sample housing units that were type A noninterviews (i.e., units eligible to be interviewed) or type B noninterviews (i.e., units not eligible for interview at the time of the survey but which could become eligible in the future) in the 1978 survey and remained in sample after the 1982 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1982 AHS questionnaire, page App-18.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented the housing units built in permit-issuing areas since the 1978 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1978 survey.)
5. All sample housing units that were selected as part of the 1978 Coverage Improvement Program. (This sample represented most of the housing units which, until 1978, did not have a chance of selection.)

Selection of the 1978 AHS-SMSA sample—The sample for the SMSA's which, in 1970, were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). In 1970, the following four SMSA's were 100-percent permit-issuing: Paterson-Clifton-Passaic, NJ; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA. The remaining eight SMSA's contain a sample from the nonpermit universe.

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

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA during the 1970 Census of Population and Housing. This file contained records for occupied housing units, vacant housing units, and housing units in certain special places or group quarters. Sampling operations were done separately for the special place and group quarters records, and for the occupied and vacant housing unit records. Before the sample was selected from the occupied and vacant housing unit records, the occupied housing unit records were stratified by race of head (non-Black/Black) and the vacant records were stratified into four categories pertaining to the value or rent associated with the vacant housing units. The occupied housing unit records were further stratified so that each unit was assigned to 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household income	Tenure	
	Owner—Family size	Renter—Family size
	1 2 3 4 5+	1 2 3 4 5+
Under $3,000		
$3,000 to $5,999 . . .		
$6,000 to $9,999 . . .		
$10,000 to $14,999 .		
$15,000 and over . . .		

Thus, for this SMSA, the occupied housing unit records from the permit-issuing universe were assigned to 1 of 100 strata for either the central city or for the balance, and the vacant housing unit records were assigned to 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. However, whenever a record was selected to be in sample, the housing unit record adjacent to it on the file was also selected to be in sample, thereby insuring the necessary designated sample size.

Before the sample was selected from the group quarters and special place records, the records were stratified by census tract and census enumeration district (ED) within the central city and within the balance of the SMSA. A sample of special place records was then selected by a procedure that produced one-quarter of the desired sample size. However, at the time of the survey, the housing units at each of the special places were listed and subsampled at a rate which produced an expected four sample units, thereby insuring the necessary designated sample size.

The second frame from which this SMSA sample was selected was a list of new construction building permits issued since 1970 (i.e., the new construction universe). The sample selection from the list of new construction building permits was an independent operation within this SMSA. Prior to sample selection, the

list of permits was chronologically stratified by the date the permits were issued, and clusters of an expected four (usually adjacent) housing units were formed. These clusters were then sampled for inclusion at the overall sampling rate.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS sample was selected from a frame consisting of areas not under the jurisdiction of permit-issuing offices (i.e., the nonpermit universe). The first step in the sampling operation for the nonpermit universe was the selection (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

The sample ED's were then divided into segments; i.e., small land areas with well-defined boundaries having an expected size of four, or a multiple of four, housing units. At the time of the survey, those segments that did not have an expected size of four were further subdivided to produce an expected four sample housing units.

The next step was the selection of one of these segments within each sample ED. All housing units in existence at the time of interview in these selected segments were eligible for sample. Thus, housing units enumerated in the 1970 census as well as housing units built since the 1970 census are included.

Sample Selection for the 1978 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

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

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 SMSA. 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. 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 Rochester, NY, SMSA.

In the remaining 11 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. These procedures added an estimated 3,378 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 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 676 units to the coverage of the housing inventory of this SMSA.

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

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

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample unit were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and units within these structures were interviewed.

In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. This procedure added an estimated 8,913 units to the coverage of the housing inventory of this SMSA.

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

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure did not add any units to the coverage of the housing inventory of this SMSA.

1978-1982 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1978 survey, was selected to represent housing units built in permit-issuing areas since the 1978 survey. Sampling procedures were identical to those used in selecting the 1970-1978 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1977, to identify any housing units missed in the 1978 survey or any housing units added since the 1978 survey.

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 1982 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 1982 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1978 (i.e., 1978-1982 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1982 housing inventory—The AHS estimates of characteristics of the 1982 housing inventory were produced using a one-stage ratio estimation procedure for the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's. A two-stage ratio estimation procedure was used for the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. A three-stage ratio estimation procedure was used for the San Bernardino-Riverside-Ontario, CA, SMSA.

Prior to the implementation of the ratio estimation procedures, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

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

The following ratio estimation procedure was employed for all sample housing units from the permit-issuing universe. This factor was computed separately for all sample housing units within each permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

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

This ratio estimation procedure was introduced to correct the probabilities of selection for samples in each of the strata used in the sample selection of the permit-issuing universe. Prior to the AHS sample selection within each SMSA, housing units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. The same probability of selection was then applied to the remaining units to select the AHS sample. Since the number of housing units deleted from the AHS

universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the San Bernardino-Riverside-Ontario, CA, SMSA, the next ratio estimation procedure was employed for new construction mobile home units. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of mobile home units for each sector of the SMSA (central city and balance) by year built groups (1970-1974 and 1975-1982) to an independent estimate of mobile home units for the corresponding sector and year built group. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of mobile homes for the corresponding sector and year built group}}{\text{Sample estimate of mobile homes for the corresponding sector and year built group}}$$

The computed ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell.

The next ratio estimation procedure was applied in the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1982 housing inventory in each sector of the SMSA (central city and balance) to an independent estimate of the total housing units in each sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1982 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for the 12 SMSA's in 1982. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1982 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units using the existing sample weights were compared to the corresponding independent estimates for the central city, balance, and total for each of the 12 SMSA's, and the estimates which showed the most likely level of net growth since the 1980 census in both the central city and balance, as well as, the total SMSA were used in this ratio estimation procedure.

As a result of this analysis, the independent estimates were used in the Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco, CA, SMSA's. The sample estimates were used in the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's.

For the Columbus, OH and San Antonio, TX, SMSA's the independent estimates were used for the total SMSA only. The independent estimates were then proportioned between the central city and the balance of each SMSA according to the central city/balance distribution of the total housing units as given by the sample estimates. Due to the central city boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimates did not reflect the 1970 central city/balance SMSA definitions. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definitions were used.

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

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

1978-1982 lost housing units—The 1978-1982 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1978 housing inventory as was described in the 1978 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas*. Since the 1978-1982 lost housing units existed, by definition, in the 1978 housing inventory, there was a 1978 housing inventory weight associated with each 1978-1982 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1978-1982 lost housing units.

1978 estimation procedure—This report presents data on the housing characteristics of the 1978 housing inventory from the 1978 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1978.

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

RELIABILITY OF THE ESTIMATES

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

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

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing and the 1982 AHS-SMSA sample.

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

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

AHS-SMSA—The AHS-SMSA evaluation study was administered to a sample of units selected from the regular interviews containing recent movers who had previous residences within the SMSA. The purpose of the study was to evaluate the accuracy of responses regarding housing characteristics of recent mover's previous residence. The results of which are presented in the Census Bureau memorandum "Analysis of the Results from the AHS-SMSA Group AA-2 1982 Evaluation Study" (June 17, 1983).

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

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

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all housing units located inside these ED's would be represented in the sample. However, it has been estimated that the 1975 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. It should be noted that since these ED's were recanvassed for the 1978 and 1982 surveys, the number of missed housing units may be considerably less for 1982.

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

Rounding errors—For errors associated with processing, rounding of estimates introduces another source of error in the data, the severity of which depends on the statistics being measured. The effect of rounding is significant relative to the sampling error only for small percentages or small medians, when these figures are derived from relatively large bases (e.g., median number of persons per household). This means that confidence intervals formed from the standard errors given may be distorted, and this should be taken into account when considering the results of the survey.

Sampling errors for the AHS-SMSA sample—The particular sample used for this survey is one of a large number of possible samples of the same size that could have been selected using the same sample design. Even if the same questionnaires, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling

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

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

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

The figures presented in the tables that follow (page App-51) are approximations to the standard errors of various estimates shown in this report for this SMSA. In order to derive standard errors that would be applicable to a wide variety of items and also could be prepared at a moderate cost, a number of approximations were required. As a result, the tables of standard errors provide an indication of the order of magnitude of the standard errors rather than precise standard errors for any specific item. Standard errors applicable to estimates of characteristics of the 1978 housing inventory can be found in the AHS Series H-170 reports for 1978.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1982 housing inventory as well as estimates of characteristics of the 1978-1982 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 15,340 for the total SMSA, 8,310 for the central city of the SMSA, and 12,900 for the balance of the SMSA.

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

Table II (page App-51) presents the standard errors of estimated percentages for the 1982 housing inventory as well as estimated percentages of the 1978-1982 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

Included in tables I and II are estimates of standard errors for estimates of zero and zero percent. These estimates of standard errors are considered as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is obtained.

For ratios, 100 (x/y), where x is not a subclass of y, table II underestimates the standard error of the ratio when there is little or no correlation between x and y. For this type of ratio, a better approximation of the standard error may be obtained by letting the standard error of the ratio be approximately equal to:

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

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

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1982 there were 211,900 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 3,190. 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	3,210
211,900	x
227,100	3,170

The entry of "x" is determined as follows by vertically interpolating between 3,210 and 3,170.

$$211{,}900 - 200{,}000 = 11{,}900$$
$$227{,}100 - 200{,}000 = 27{,}100$$

$$3{,}210 + \frac{11{,}900}{27{,}100}(3{,}170 - 3{,}210) = 3{,}190$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 208,710 to 215,090 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1982 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 206,800 to 217,000 housing units with 90 percent confidence; and that the average estimate lies within the interval from 205,520 to 218,280 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 211,900 owner-occupied housing units, 44,800, or 21.1 percent, had two bedrooms. Interpolation using table II of this appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 21.1 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	21.1	25 or 75
200,000	0.7	a	1.0
211,900		p	
250,000	0.6	b	0.9

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

$$21.1-10.0 = 11.1$$
$$25.0-10.0 = 15.0$$

$$0.7 + \frac{11.1}{15.0}(1.0-0.7) = 0.9$$

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

$$21.1-10.0 = 11.1$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{11.1}{15.0}(0.9-0.6) = 0.8$$

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

$$211{,}900-200{,}000 = 11{,}900$$
$$250{,}000-200{,}000 = 50{,}000$$

$$0.9 + \frac{11{,}900}{50{,}000}(0.8-0.9) = 0.9$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 20.2 to 22.0 percent; the 90-percent confidence interval is from 19.7 to 22.5 percent; and the 95-percent confidence interval is from 19.3 to 22.9 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. If there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1978 and 1982 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1978 and 1982 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1982 there were 117,900 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 73,100. Table I shows the standard error of 44,800 is approximately 2,030, and the standard error of 117,900 is approximately 2,920. Therefore, the standard error of the estimated difference of 73,100 is about 3,560.

$$3{,}560 = \sqrt{(2{,}030)^2 + (2{,}920)^2}$$

Consequently, the 68-percent confidence interval for the 73,100 difference is from 69,540 to 76,660 housing units. Therefore, a conclusion that the average estimate derived from all possible samples, of this difference, lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 67,400 to 78,800 housing units, and the 95-percent confidence interval is from 65,980 to 80,220 housing units. Thus, we can conclude with 95 percent confidence that the number of 1982 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied units with two bedrooms since the 95-percent confidence interval does not include zero or negative values.

Medians—For medians presented in certain tables, the sampling error depends on the size of the base and on the distribution upon which the median is based. An approximate method for measuring the reliability of the estimated median is to determine an interval about the estimated median so that there is a stated degree of confidence that the average median from all possible samples lies within the interval. The following procedure may be used to estimate confidence limits of a median based on sample data:

1. From table II, determine the standard error of a 50-percent characteristic on the base of the median.
2. Add to and subtract from 50 percent, the standard error determined in step 1.
3. Using the distribution of the characteristics, determine the confidence interval corresponding to the two points established in step 2. To find the lower endpoint of the confidence

interval, it is necessary to know into which interval of the distribution the lower percentage limit falls. Similarly, to find the upper endpoint to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median— Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.6. The base of the distribution from which this median was determined is 211,900 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 211,900 is approximately 1.2 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50-percent twice the standard error determined in step 1. This yields percentage limits of 47.6 and 52.4.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with three persons (for purposes of calculating the median, the category of three persons is considered to be from 2.5 to 3.5 persons) contains the 47.6 percent derived in step 2. About 100,500 housing units or 47.4 percent fall below this interval, and 41,800 housing units or 19.7 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(47.6-47.4)}{19.7} = 2.5$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.4 percent derived in step 2. About 100,500 housing units or 47.4 percent fall below this interval, and 41,800 housing units or 19.7 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(52.4-47.4)}{19.7} = 2.8$$

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

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1982 Housing Inventory and for Estimated Number of 1978-1982 Lost Units for the Columbus, OH, SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	110	100	110	50,000	2,150	2,010	1,970
100	110	100	110	75,000	2,530	2,290	2,130
200	150	140	150	100,000	2,810	2,410	2,080
500	230	230	240	150,000	3,130	2,300	1,160
700	270	270	280	163,200	3,170	2,190	—
1,000	330	320	330	200,000	3,210	1,580	—
2,500	510	510	520	227,100	3,170	—	—
5,000	720	710	740	250,000	3,090	—	—
10,000	1,020	1,000	1,020	300,000	2,710	—	—
25,000	1,570	1,520	1,540	390,300	—	—	—

[1]For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2 for the total SMSA and balance (not in central city) and 1.3 for the central city estimates. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero where the estimate for the total SMSA was derived from an independent estimate.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1982 Housing Inventory and for Estimated Percentages of 1978-1982 Lost Housing Units for the Columbus, OH, SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	52.7	52.7	52.7	52.7	52.7	52.8	50,000	0.2	0.5	1.0	1.4	2.0	2.4
200	35.8	35.8	35.8	35.8	35.8	37.3	75,000	0.15	0.4	0.8	1.2	1.7	1.9
500	18.2	18.2	18.2	18.2	20.5	23.6	100,000	0.11	0.3	0.7	1.0	1.4	[illegible]
700	13.7	13.7	13.7	13.7	17.3	20.0	150,000	0.07	0.3	0.6	0.8	1.2	[illegible]
1,000	10.0	10.0	10.0	10.0	14.5	16.7	200,000	0.06	0.2	0.5	0.7	[illegible]	[illegible]
2,500	4.3	4.3	4.6	6.3	9.1	10.6	250,000	0.04	0.2	0.5	0.6	[illegible]	[illegible]
5,000	2.2	2.2	3.3	4.5	6.5	7.5	300,000	0.04	0.2	0.4	[illegible]	[illegible]	[illegible]
10,000	1.1	1.1	2.3	3.2	4.6	5.3	400,000	0.03	0.2	0.4	[illegible]	[illegible]	[illegible]
25,000	0.4	0.7	1.5	2.0	2.9	3.3							

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

Kansas City 1982

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1982 survey App-42
Selection of the 1978 AHS-SMSA sample App-43
Sample selection for the 1978 Coverage Improvement Program . App-44
Coverage Improvement for deficiency 1 ..... App-44
Coverage Improvement for deficiency 2 ..... App-44
Coverage Improvement for deficiencies 3-6 ..... App-44
1978-1982 additions to the housing inventory App-45
1970 Census of Population and Housing App-45
ESTIMATION App-45
1982 housing inventory App-45
1978-1982 lost housing units ... App-46
1978 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-47
Rounding errors App-47
Sampling errors for the AHS-SMSA sample App-47
Illustration of the use of the standard error tables App-48
Differences App-49
Illustration of the computation of the standard error of a difference App-49
Medians App-49
Illustration of the computation of the 95-percent confidence interval of a median ... App-50
Standard error tables App-51

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 12 SMSA's in this report series (H-170-82) are based on data collected from the 1982 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 12 SMSA's selected for interview during 1982 were interviewed previously in 1975 and 1978 (see the list of SMSA reports from the AHS in the introduction of this report).

Three of the larger SMSA's (Atlanta, GA; Philadelphia, PA-NJ; and San Francisco-Oakland, CA) in the 1982 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1975 and 1978 surveys, evenly divided between the central city and the balance of the respective SMSA. The 9 remaining SMSA's (Cincinnati, OH-KY-IN; Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Paterson-Clifton-Passaic, NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; and San Diego, CA) in the 1982 group were represented by a sample of 5,000 designated housing units in the 1975 and 1978 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1975 and 1978 surveys, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panels 1 through 3 were dropped from all SMSA's for the 1982 survey, excluding the central cities of the Paterson-Clifton-Passaic, NJ, SMSA. Due to additional budget limitations for the 1982 survey, it also became necessary to drop part of the sample of housing units from panels 4 through 12 in the Atlanta, GA; New Orleans, LA; Philadelphia, PA-NJ; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. The effect of this reduction was to maintain a sample size of 4,250 in all SMSA's. The interviewing was done during the period April 1982 through December 1982 with one-ninth of the sample housing units being visited each month.

In this SMSA, 3,871 housing units were eligible for interview. Of these sample housing units, 80 interviews were not obtained because, for occupied sample units, the occupants refused to be interviewed, were not at home after repeated visits, or were unavailable for some other reason; or, for vacant units, no informed respondent could be found after repeated visits. In addition to the 3,871 housing units eligible for interview, 214 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 1982 survey—The sample housing units designated to be interviewed in the 1982 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 1978 survey and remained in sample after the 1982 panel reduction.
2. All sample housing units that were type A noninterviews (i.e., units eligible to be interviewed) or type B noninterviews (i.e., units not eligible for interview at the time of the survey but which could become eligible in the future) in the 1978 survey and remained in sample after the 1982 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1982 AHS questionnaire, page App-18.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented the housing units built in permit-issuing areas since the 1978 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1978 survey.)
5. All sample housing units that were selected as part of the 1978 Coverage Improvement Program. (This sample represented most of the housing units which, until 1978, did not have a chance of selection.)

Selection of the 1978 AHS-SMSA sample—The sample for the SMSA's which, in 1970, were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). In 1970, the following four SMSA's were 100-percent permit-issuing: Paterson-Clifton-Passaic, NJ; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA. The remaining eight SMSA's contain a sample from the nonpermit universe.

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

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA during the 1970 Census of Population and Housing. This file contained records for occupied housing units, vacant housing units, and housing units in certain special places or group quarters. Sampling operations were done separately for the special place and group quarters records, and for the occupied and vacant housing unit records. Before the sample was selected from the occupied and vacant housing unit records, the occupied housing unit records were stratified by race of head (non-Black/Black) and the vacant records were stratified into four categories pertaining to the value or rent associated with the vacant housing units. The occupied housing unit records were further stratified so that each unit was assigned to 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household income	Tenure	
	Owner—Family size	Renter—Family size
	1 2 3 4 5+	1 2 3 4 5+
Under $3,000		
$3,000 to $5,999 . . .		
$6,000 to $9,999 . . .		
$10,000 to $14,999 .		
$15,000 and over . . .		

Thus, for this SMSA, the occupied housing unit records from the permit-issuing universe were assigned to 1 of 100 strata for either the central city or for the balance, and the vacant housing unit records were assigned to 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. However, whenever a record was selected to be in sample, the housing unit record adjacent to it on the file was also selected to be in sample, thereby insuring the necessary designated sample size.

Before the sample was selected from the group quarters and special place records, the records were stratified by census tract and census enumeration district (ED) within the central city and within the balance of the SMSA. A sample of special place records was then selected by a procedure that produced one-quarter of the desired sample size. However, at the time of the survey, the housing units at each of the special places were listed and subsampled at a rate which produced an expected four sample units, thereby insuring the necessary designated sample size.

The second frame from which this SMSA sample was selected was a list of new construction building permits issued since 1970 (i.e., the new construction universe). The sample selection from the list of new construction building permits was an independent operation within this SMSA. Prior to sample selection, the

list of permits was chronologically stratified by the date the permits were issued, and clusters of an expected four (usually adjacent) housing units were formed. These clusters were then sampled for inclusion at the overall sampling rate.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS sample was selected from a frame consisting of areas not under the jurisdiction of permit-issuing offices (i.e., the nonpermit universe). The first step in the sampling operation for the nonpermit universe was the selection (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

The sample ED's were then divided into segments; i.e., small land areas with well-defined boundaries having an expected size of four, or a multiple of four, housing units. At the time of the survey, those segments that did not have an expected size of four were further subdivided to produce an expected four sample housing units.

The next step was the selection of one of these segments within each sample ED. All housing units in existence at the time of interview in these selected segments were eligible for sample. Thus, housing units enumerated in the 1970 census as well as housing units built since the 1970 census are included.

Sample Selection for the 1978 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

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

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 SMSA. 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. 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 Rochester, NY, SMSA.

In the remaining 11 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. These procedures added an estimated 4,589 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 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 2,399 units to the coverage of the housing inventory of this SMSA.

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

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

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample unit were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and units within these structures were interviewed.

In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. This procedure added an estimated 7,028 units to the coverage of the housing inventory of this SMSA.

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

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 961 units to the coverage of the housing inventory of this SMSA.

1978-1982 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1978 survey, was selected to represent housing units built in permit-issuing areas since the 1978 survey. Sampling procedures were identical to those used in selecting the 1970-1978 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1977, to identify any housing units missed in the 1978 survey or any housing units added since the 1978 survey.

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 1982 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 1982 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1978 (i.e., 1978-1982 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1982 housing inventory—The AHS estimates of characteristics of the 1982 housing inventory were produced using a one-stage ratio estimation procedure for the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's. A two-stage ratio estimation procedure was used for the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. A three-stage ratio estimation procedure was used for the San Bernardino-Riverside-Ontario, CA, SMSA.

Prior to the implementation of the ratio estimation procedures, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

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

The following ratio estimation procedure was employed for all sample housing units from the permit-issuing universe. This factor was computed separately for all sample housing units within each permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

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

This ratio estimation procedure was introduced to correct the probabilities of selection for samples in each of the strata used in the sample selection of the permit-issuing universe. Prior to the AHS sample selection within each SMSA, housing units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. The same probability of selection was then applied to the remaining units to select the AHS sample. Since the number of housing units deleted from the AHS

universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the San Bernardino-Riverside-Ontario, CA, SMSA, the next ratio estimation procedure was employed for new construction mobile home units. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of mobile home units for each sector of the SMSA (central city and balance) by year built groups (1970-1974 and 1975-1982) to an independent estimate of mobile home units for the corresponding sector and year built group. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of mobile homes for the corresponding sector and year built group}}{\text{Sample estimate of mobile homes for the corresponding sector and year built group}}$$

The computed ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell.

The next ratio estimation procedure was applied in the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1982 housing inventory in each sector of the SMSA (central city and balance) to an independent estimate of the total housing units in each sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1982 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for the 12 SMSA's in 1982. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1982 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units using the existing sample weights were compared to the corresponding independent estimates for the central city, balance, and total for each of the 12 SMSA's, and the estimates which showed the most likely level of net growth since the 1980 census in both the central city and balance, as well as, the total SMSA were used in this ratio estimation procedure.

As a result of this analysis, the independent estimates were used in the Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco, CA, SMSA's. The sample estimates were used in the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's.

For the Columbus, OH and San Antonio, TX, SMSA's the independent estimates were used for the total SMSA only. The independent estimates were then proportioned between the central city and the balance of each SMSA according to the central city/balance distribution of the total housing units as given by the sample estimates. Due to the central city boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimates did not reflect the 1970 central city/balance SMSA definitions. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definitions were used.

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

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

1978-1982 lost housing units—The 1978-1982 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1978 housing inventory as was described in the 1978 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1978-1982 lost housing units existed, by definition, in the 1978 housing inventory, there was a 1978 housing inventory weight associated with each 1978-1982 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1978-1982 lost housing units.

1978 estimation procedure—This report presents data on the housing characteristics of the 1978 housing inventory from the 1978 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1978.

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

RELIABILITY OF THE ESTIMATES

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

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

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing and the 1982 AHS-SMSA sample.

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

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

AHS-SMSA—The AHS-SMSA evaluation study was administered to a sample of units selected from the regular interviews containing recent movers who had previous residences within the SMSA. The purpose of the study was to evaluate the accuracy of responses regarding housing characteristics of recent mover's previous residence. The results of which are presented in the Census Bureau memorandum "Analysis of the Results from the AHS-SMSA Group AA-2 1982 Evaluation Study" (June 17, 1983).

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

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

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all housing units located inside these ED's would be represented in the sample. However, it has been estimated that the 1975 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. It should be noted that since these ED's were recanvassed for the 1978 and 1982 surveys, the number of missed housing units may be considerably less for 1982.

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

Rounding errors—For errors associated with processing, rounding of estimates introduces another source of error in the data, the severity of which depends on the statistics being measured. The effect of rounding is significant relative to the sampling error only for small percentages or small medians, when these figures are derived from relatively large bases (e.g., median number of persons per household). This means that confidence intervals formed from the standard errors given may be distorted, and this should be taken into account when considering the results of the survey.

Sampling errors for the AHS-SMSA sample—The particular sample used for this survey is one of a large number of possible samples of the same size that could have been selected using the same sample design. Even if the same questionnaires, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling

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

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

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

The figures presented in the tables that follow (page App-51) are approximations to the standard errors of various estimates shown in this report for this SMSA. In order to derive standard errors that would be applicable to a wide variety of items and also could be prepared at a moderate cost, a number of approximations were required. As a result, the tables of standard errors provide an indication of the order of magnitude of the standard errors rather than precise standard errors for any specific item. Standard errors applicable to estimates of characteristics of the 1978 housing inventory can be found in the AHS Series H-170 reports for 1978.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1982 housing inventory as well as estimates of characteristics of the 1978-1982 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 20,070 for the total SMSA, 13,960 for the central city of the SMSA, and 14,420 for the balance of the SMSA.

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

Table II (page App-51) presents the standard errors of estimated percentages for the 1982 housing inventory as well as estimated percentages of the 1978-1982 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

Included in tables I and II are estimates of standard errors for estimates of zero and zero percent. These estimates of standard errors are considered as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is obtained.

For ratios, 100 (x/y), where x is not a subclass of y, table II underestimates the standard error of the ratio when there is little or no correlation between x and y. For this type of ratio, a better approximation of the standard error may be obtained by letting the standard error of the ratio be approximately equal to:

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

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

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1982 there were 329,300 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 4,310. 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	4,410
329,300	x
338,700	4,280

The entry of "x" is determined as follows by vertically interpolating between 4,410 and 4,280.

$$329{,}300 - 300{,}000 = 29{,}300$$
$$338{,}700 - 300{,}000 = 38{,}700$$

$$4{,}410 + \frac{29{,}300}{38{,}700}(4{,}280 - 4{,}410) = 4{,}310$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 324,990 to 333,610 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1982 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 322,400 to 336,200 housing units with 90 percent confidence; and that the average estimate lies within the interval from 320,680 to 337,920 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 329,300 owner-occupied housing units, 86,300, or 26.2 percent, had two bedrooms. Interpolation using table II of this appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 26.2 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	26.2	50
300,000	1.0	a	1.2
329,300		p	
400,000	0.9	b	1.0

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

$$26.2-25.0 = 1.2$$
$$50.0-25.0 = 25.0$$

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

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

$$26.2-25.0 = 1.2$$
$$50.0-25.0 = 25.0$$

$$0.9 + \frac{1.2}{25.0}(1.0-0.9) = 0.9$$

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

$$329{,}300-300{,}000 = 29{,}300$$
$$400{,}000-300{,}000 = 100{,}000$$

$$1.0 + \frac{29{,}300}{100{,}000}(0.9-1.0) = 1.0$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 25.2 to 27.2 percent; the 90-percent confidence interval is from 24.6 to 27.8 percent; and the 95-percent confidence interval is from 24.2 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. 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 1978 and 1982 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1978 and 1982 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1982 there were 170,900 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 84,600. Table I shows the standard error of 86,300 is approximately 3,270, and the standard error of 170,900 is approximately 4,140. Therefore, the standard error of the estimated difference of 84,600 is about 5,280.

$$5{,}280 = \sqrt{(3{,}270)^2 + (4{,}140)^2}$$

Consequently, the 68-percent confidence interval for the 84,600 difference is from 79,320 to 89,880 housing units. Therefore, a conclusion that the average estimate derived from all possible samples, of this difference, lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 76,150 to 93,050 housing units, and the 95-percent confidence interval is from 74,040 to 95,160 housing units. Thus, we can conclude with 95 percent confidence that the number of 1982 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied units with two bedrooms since the 95-percent confidence interval does not include zero or negative values.

Medians—For medians presented in certain tables, the sampling error depends on the size of the base and on the distribution upon which the median is based. An approximate method for measuring the reliability of the estimated median is to determine an interval about the estimated median so that there is a stated degree of confidence that the average median from all possible samples lies within the interval. The following procedure may be used to estimate confidence limits of a median based on sample data:

1. From table II, determine the standard error of a 50-percent characteristic on the base of the median.
2. Add to and subtract from 50 percent, the standard error determined in step 1.
3. Using the distribution of the characteristics, determine the confidence interval corresponding to the two points established in step 2. To find the lower endpoint of the confidence

interval, it is necessary to know into which interval of the distribution the lower percentage limit falls. Similarly, to find the upper endpoint to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.5. The base of the distribution from which this median was determined is 329,300 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 329,300 is approximately 1.1 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50-percent twice the standard error determined in step 1. This yields percentage limits of 47.8 and 52.2.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) contains the 47.8 percent derived in step 2. About 55,500 housing units or 16.9 percent fall below this interval, and 108,200 housing units or 32.9 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$1.5 + (2.5-1.5)\frac{(47.8-16.9)}{32.9} = 2.4$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.2 percent derived in step 2. About 163,700 housing units or 49.7 percent fall below this interval, and 56,900 housing units or 17.3 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\frac{(52.2-49.7)}{17.3} = 2.6$$

Thus, the 95-percent confidence interval ranges from 2.4 to 2.6 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1982 Housing Inventory and for Estimated Number of 1978-1982 Lost Units for the Kansas City, MO-KS, SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	150	140	160	75,000	3,100	2,510	3,060
100	150	140	160	100,000	3,480	2,570	3,360
200	170	170	180	150,000	4,000	2,140	3,650
500	270	260	280	192,800	4,280	—	3,640
700	320	310	330	200,000	4,310	—	3,620
1,000	390	370	400	250,000	4,440	—	3,230
2,500	610	580	630	300,000	4,410	—	2,340
5,000	860	820	890	338,700	4,280	—	—
10,000	1,210	1,140	1,250	400,000	3,840	—	—
25,000	1,880	1,730	1,920	500,000	2,100	—	—
50,000	2,600	2,250	2,610	531,500	—	—	—

[1]For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.3 for the total SMSA and the central city and 1.2 for the balance (not in central city) estimates. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero where the estimate for the total SMSA was derived from an independent estimate.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1982 Housing Inventory and for Estimated Percentages of 1978-1982 Lost Housing Units for the Kansas City, MO-KS, SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
200	44.4	44.4	44.4	44.4	44.4	44.7	75,000	0.2	0.5	1.0	1.4	2.0	2.3
500	24.2	24.2	24.2	24.2	24.5	28.3	100,000	0.2	0.4	0.9	1.2	1.7	2.0
700	18.6	18.6	18.6	18.6	20.7	23.9	150,000	0.11	0.3	0.7	1.0	1.4	1.6
1,000	13.8	13.8	13.8	13.8	17.3	20.0	200,000	0.08	0.3	0.6	0.8	1.2	1.4
2,500	6.0	6.0	6.0	7.6	10.9	12.6	250,000	0.06	0.3	0.6	0.8	1.1	1.3
5,000	3.1	3.1	3.9	5.4	7.7	8.9	300,000	0.05	0.2	0.5	0.7	1.0	1.2
10,000	1.6	1.6	2.8	3.8	5.5	6.3	400,000	0.04	0.2	0.4	0.6	0.9	1.0
25,000	0.6	0.8	1.7	2.4	3.5	4.0	500,000	0.03	0.2	0.4	0.5	0.8	0.9
50,000	0.3	0.6	1.2	1.7	2.4	2.8	600,000	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.

New Orleans 1982

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN . . . App-42
Annual Housing Survey . . . App-42
Designation of sample housing units for the 1982 survey . . . App-42
Selection of the 1978 AHS-SMSA sample . . . App-43
Sample selection for the 1978 Coverage Improvement Program . App-44
Coverage Improvement for deficiency 1 . . . App-44
Coverage Improvement for deficiency 2 . . . App-44
Coverage Improvement for deficiencies 3-6 . . . App-44
1978-1982 additions to the housing inventory . . . App-45
1970 Census of Population and Housing . . . App-45
ESTIMATION . . . App-45
1982 housing inventory . . . App-45
1978-1982 lost housing units . . . App-46
1978 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-47
Rounding errors . . . App-47
Sampling errors for the AHS-SMSA sample . . . App-47
Illustration of the use of the standard error tables . . . App-48
Differences . . . App-49
Illustration of the computation of the standard error of a difference . . . App-49
Medians . . . App-49
Illustration of the computation of the 95-percent confidence interval of a median . . . App-50
Standard error tables . . . App-51

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 12 SMSA's in this report series (H-170-82) are based on data collected from the 1982 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 12 SMSA's selected for interview during 1982 were interviewed previously in 1975 and 1978 (see the list of SMSA reports from the AHS in the introduction of this report).

Three of the larger SMSA's (Atlanta, GA; Philadelphia, PA-NJ; and San Francisco-Oakland, CA) in the 1982 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1975 and 1978 surveys, evenly divided between the central city and the balance of the respective SMSA. The 9 remaining SMSA's (Cincinnati, OH-KY-IN; Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Paterson-Clifton-Passaic, NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; and San Diego, CA) in the 1982 group were represented by a sample of 5,000 designated housing units in the 1975 and 1978 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1975 and 1978 surveys, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panels 1 through 3 were dropped from all SMSA's for the 1982 survey, excluding the central cities of the Paterson-Clifton-Passaic, NJ, SMSA. Due to additional budget limitations for the 1982 survey, it also became necessary to drop part of the sample of housing units from panels 4 through 12 in the Atlanta, GA; New Orleans, LA; Philadelphia, PA-NJ; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. The effect of this reduction was to maintain a sample size of 4,250 in all SMSA's. The interviewing was done during the period April 1982 through December 1982 with one-ninth of the sample housing units being visited each month.

In this SMSA, 4,026 housing units were eligible for interview. Of these sample housing units, 165 interviews were not obtained because, for occupied sample units, the occupants refused to be interviewed, were not at home after repeated visits, or were unavailable for some other reason; or, for vacant units, no informed respondent could be found after repeated visits. In addition to the 4,026 housing units eligible for interview, 255 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 1982 survey—The sample housing units designated to be interviewed in the 1982 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 1978 survey and remained in sample after the 1982 panel reduction.
2. All sample housing units that were type A noninterviews (i.e., units eligible to be interviewed) or type B noninterviews (i.e., units not eligible for interview at the time of the survey but which could become eligible in the future) in the 1978 survey and remained in sample after the 1982 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1982 AHS questionnaire, page App-18.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented the housing units built in permit-issuing areas since the 1978 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1978 survey.)
5. All sample housing units that were selected as part of the 1978 Coverage Improvement Program. (This sample represented most of the housing units which, until 1978, did not have a chance of selection.)

Selection of the 1978 AHS-SMSA sample—The sample for the SMSA's which, in 1970, were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). In 1970, the following four SMSA's were 100-percent permit-issuing: Paterson-Clifton-Passaic, NJ; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA. The remaining eight SMSA's contain a sample from the nonpermit universe.

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

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA during the 1970 Census of Population and Housing. This file contained records for occupied housing units, vacant housing units, and housing units in certain special places or group quarters. Sampling operations were done separately for the special place and group quarters records, and for the occupied and vacant housing unit records. Before the sample was selected from the occupied and vacant housing unit records, the occupied housing unit records were stratified by race of head (non-Black/Black) and the vacant records were stratified into four categories pertaining to the value or rent associated with the vacant housing units. The occupied housing unit records were further stratified so that each unit was assigned to 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household income	Tenure	
	Owner—Family size	Renter—Family size
	1 2 3 4 5+	1 2 3 4 5+
Under $3,000		
$3,000 to $5,999 . . .		
$6,000 to $9,999 . . .		
$10,000 to $14,999 .		
$15,000 and over . . .		

Thus, for this SMSA, the occupied housing unit records from the permit-issuing universe were assigned to 1 of 100 strata for either the central city or for the balance, and the vacant housing unit records were assigned to 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. However, whenever a record was selected to be in sample, the housing unit record adjacent to it on the file was also selected to be in sample, thereby insuring the necessary designated sample size.

Before the sample was selected from the group quarters and special place records, the records were stratified by census tract and census enumeration district (ED) within the central city and within the balance of the SMSA. A sample of special place records was then selected by a procedure that produced one-quarter of the desired sample size. However, at the time of the survey, the housing units at each of the special places were listed and subsampled at a rate which produced an expected four sample units, thereby insuring the necessary designated sample size.

The second frame from which this SMSA sample was selected was a list of new construction building permits issued since 1970 (i.e., the new construction universe). The sample selection from the list of new construction building permits was an independent operation within this SMSA. Prior to sample selection, the

list of permits was chronologically stratified by the date the permits were issued, and clusters of an expected four (usually adjacent) housing units were formed. These clusters were then sampled for inclusion at the overall sampling rate.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS sample was selected from a frame consisting of areas not under the jurisdiction of permit-issuing offices (i.e., the nonpermit universe). The first step in the sampling operation for the nonpermit universe was the selection (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

The sample ED's were then divided into segments; i.e., small land areas with well-defined boundaries having an expected size of four, or a multiple of four, housing units. At the time of the survey, those segments that did not have an expected size of four were further subdivided to produce an expected four sample housing units.

The next step was the selection of one of these segments within each sample ED. All housing units in existence at the time of interview in these selected segments were eligible for sample. Thus, housing units enumerated in the 1970 census as well as housing units built since the 1970 census are included.

Sample Selection for the 1978 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

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

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 SMSA. 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. 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 Rochester, NY, SMSA.

In the remaining 11 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. These procedures added an estimated 1,225 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 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 504 units to the coverage of the housing inventory of this SMSA.

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

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

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample unit were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and units within these structures were interviewed.

In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. This procedure added an estimated 8,119 units to the coverage of the housing inventory of this SMSA.

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

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 1,887 units to the coverage of the housing inventory of this SMSA.

1978-1982 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1978 survey, was selected to represent housing units built in permit-issuing areas since the 1978 survey. Sampling procedures were identical to those used in selecting the 1970-1978 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1977, to identify any housing units missed in the 1978 survey or any housing units added since the 1978 survey.

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 1982 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 1982 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1978 (i.e., 1978-1982 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1982 housing inventory—The AHS estimates of characteristics of the 1982 housing inventory were produced using a one-stage ratio estimation procedure for the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's. A two-stage ratio estimation procedure was used for the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. A three-stage ratio estimation procedure was used for the San Bernardino-Riverside-Ontario, CA, SMSA.

Prior to the implementation of the ratio estimation procedures, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

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

The following ratio estimation procedure was employed for all sample housing units from the permit-issuing universe. This factor was computed separately for all sample housing units within each permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

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

This ratio estimation procedure was introduced to correct the probabilities of selection for samples in each of the strata used in the sample selection of the permit-issuing universe. Prior to the AHS sample selection within each SMSA, housing units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. The same probability of selection was then applied to the remaining units to select the AHS sample. Since the number of housing units deleted from the AHS

universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the San Bernardino-Riverside-Ontario, CA, SMSA, the next ratio estimation procedure was employed for new construction mobile home units. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of mobile home units for each sector of the SMSA (central city and balance) by year built groups (1970-1974 and 1975-1982) to an independent estimate of mobile home units for the corresponding sector and year built group. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of mobile homes for the corresponding sector and year built group}}{\text{Sample estimate of mobile homes for the corresponding sector and year built group}}$$

The computed ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell.

The next ratio estimation procedure was applied in the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1982 housing inventory in each sector of the SMSA (central city and balance) to an independent estimate of the total housing units in each sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1982 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for the 12 SMSA's in 1982. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1982 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units using the existing sample weights were compared to the corresponding independent estimates for the central city, balance, and total for each of the 12 SMSA's, and the estimates which showed the most likely level of net growth since the 1980 census in both the central city and balance, as well as, the total SMSA were used in this ratio estimation procedure.

As a result of this analysis, the independent estimates were used in the Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco, CA, SMSA's. The sample estimates were used in the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's.

For the Columbus, OH and San Antonio, TX, SMSA's the independent estimates were used for the total SMSA only. The independent estimates were then proportioned between the central city and the balance of each SMSA according to the central city/balance distribution of the total housing units as given by the sample estimates. Due to the central city boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimates did not reflect the 1970 central city/balance SMSA definitions. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definitions were used.

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

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

1978-1982 lost housing units—The 1978-1982 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1978 housing inventory as was described in the 1978 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1978-1982 lost housing units existed, by definition, in the 1978 housing inventory, there was a 1978 housing inventory weight associated with each 1978-1982 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1978-1982 lost housing units.

1978 estimation procedure.—This report presents data on the housing characteristics of the 1978 housing inventory from the 1978 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1978.

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

RELIABILITY OF THE ESTIMATES

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

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

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing and the 1982 AHS-SMSA sample.

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

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

AHS-SMSA—The AHS-SMSA evaluation study was administered to a sample of units selected from the regular interviews containing recent movers who had previous residences within the SMSA. The purpose of the study was to evaluate the accuracy of responses regarding housing characteristics of recent mover's previous residence. The results of which are presented in the Census Bureau memorandum "Analysis of the Results from the AHS-SMSA Group AA-2 1982 Evaluation Study" (June 17, 1983).

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

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

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all housing units located inside these ED's would be represented in the sample. However, it has been estimated that the 1975 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. It should be noted that since these ED's were recanvassed for the 1978 and 1982 surveys, the number of missed housing units may be considerably less for 1982.

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

Rounding errors—For errors associated with processing, rounding of estimates introduces another source of error in the data, the severity of which depends on the statistics being measured. The effect of rounding is significant relative to the sampling error only for small percentages or small medians, when these figures are derived from relatively large bases (e.g., median number of persons per household). This means that confidence intervals formed from the standard errors given may be distorted, and this should be taken into account when considering the results of the survey.

Sampling errors for the AHS-SMSA sample—The particular sample used for this survey is one of a large number of possible samples of the same size that could have been selected using the same sample design. Even if the same questionnaires, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling

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

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

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

The figures presented in the tables that follow (page App-51) are approximations to the standard errors of various estimates shown in this report for this SMSA. In order to derive standard errors that would be applicable to a wide variety of items and also could be prepared at a moderate cost, a number of approximations were required. As a result, the tables of standard errors provide an indication of the order of magnitude of the standard errors rather than precise standard errors for any specific item. Standard errors applicable to estimates of characteristics of the 1978 housing inventory can be found in the AHS Series H-170 reports for 1978.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1982 housing inventory as well as estimates of characteristics of the 1978-1982 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 38,770 for the total SMSA, 13,970 for the central city of the SMSA, and 36,160 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 (page App-51) presents the standard errors of estimated percentages for the 1982 housing inventory as well as estimated percentages of the 1978-1982 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

Included in tables I and II are estimates of standard errors for estimates of zero and zero percent. These estimates of standard errors are considered as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is obtained.

For ratios, 100 (x/y), where x is not a subclass of y, table II underestimates the standard error of the ratio when there is little or no correlation between x and y. For this type of ratio, a better approximation of the standard error may be obtained by letting the standard error of the ratio be approximately equal to:

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

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

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1982 there were 232,200 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 3,930. The following interpolation procedure was used.

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

Size of estimate	Standard error
230,000	3,930
232,200	x
239,200	3,930

The entry of "x" is determined as follows by vertically interpolating between 3,930 and 3,930.

$$232{,}200 - 230{,}000 = 2{,}200$$
$$239{,}200 - 230{,}000 = 9{,}200$$

$$3{,}930 + \frac{2{,}200}{9{,}200}(3{,}930 - 3{,}930) = 3{,}930$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 228,270 to 236,130 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1982 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 225,910 to 238,490 housing units with 90 percent confidence; and that the average estimate lies within the interval from 224,340 to 240,060 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 232,200 owner-occupied housing units, 53,600, or 23.1 percent, had two bedrooms. Interpolation using table II of this appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 23.1 percent is approximately 1.0 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	23.1	25 or 75
200,000	0.8	a	1.1
232,200		p	
250,000	0.7	b	1.0

1. The entry for cell "a" is determined by horizontal interpolation between 0.8 and 1.1.

$$23.1-10.0 = 13.1$$
$$25.0-10.0 = 15.0$$

$$0.8 + \frac{13.1}{15.0}(1.1-0.8) = 1.1$$

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

$$23.1-10.0 = 13.1$$
$$25.0-10.0 = 15.0$$

$$0.7 + \frac{13.1}{15.0}(1.0-0.7) = 1.0$$

3. The entry for "p" is then determined by vertical interpolation between 1.1 and 1.0.

$$232{,}200-200{,}000 = 32{,}800$$
$$250{,}000-200{,}000 = 50{,}000$$

$$1.1 + \frac{32{,}800}{50{,}000}(1.0-1.1) = 1.0$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 22.1 to 24.1 percent; the 90-percent confidence interval is from 21.5 to 24.7 percent; and the 95-percent confidence interval is from 21.1 to 25.1 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. If there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1978 and 1982 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1978 and 1982 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1982 there were 119,900 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 66,300. Table I shows the standard error of 53,600 is approximately 2,490, and the standard error of 119,900 is approximately 3,400. Therefore, the standard error of the estimated difference of 66,300 is about 4,210.

$$4{,}210 = \sqrt{(2{,}490)^2 + (3{,}400)^2}$$

Consequently, the 68-percent confidence interval for the 66,300 difference is from 62,090 to 70,510 housing units. Therefore, a conclusion that the average estimate derived from all possible samples, of this difference, lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 59,560 to 73,040 housing units, and the 95-percent confidence interval is from 57,880 to 74,720 housing units. Thus, we can conclude with 95 percent confidence that the number of 1982 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied units with two bedrooms since the 95-percent confidence interval does not include zero or negative values.

Medians—For medians presented in certain tables, the sampling error depends on the size of the base and on the distribution upon which the median is based. An approximate method for measuring the reliability of the estimated median is to determine an interval about the estimated median so that there is a stated degree of confidence that the average median from all possible samples lies within the interval. The following procedure may be used to estimate confidence limits of a median based on sample data:

1. From table II, determine the standard error of a 50-percent characteristic on the base of the median.
2. Add to and subtract from 50 percent, the standard error determined in step 1.
3. Using the distribution of the characteristics, determine the confidence interval corresponding to the two points established in step 2. To find the lower endpoint of the confidence

interval, it is necessary to know into which interval of the distribution the lower percentage limit falls. Similarly, to find the upper endpoint to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median— Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.9. The base of the distribution from which this median was determined is 232,200 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 232,200 is approximately 1.2 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50-percent twice the standard error determined in step 1. This yields percentage limits of 47.6 and 52.4.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with three persons (for purposes of calculating the median, the category of three persons is considered to be from 2.5 to 3.5 persons) contains the 47.6 percent derived in step 2. About 100,500 housing units or 43.3 percent fall below this interval, and 43,000 housing units or 18.5 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(47.6-43.3)}{18.5} = 2.7$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.4 percent derived in step 2. About 100,500 housing units or 43.3 percent fall below this interval, and 43,000 housing units or 18.5 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(52.4-43.3)}{18.5} = 3.0$$

Thus, the 95-percent confidence interval ranges from 2.7 to 3.0 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1982 Housing Inventory and for Estimated Number of 1978-1982 Lost Units for the New Orleans, LA, SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	130	130	130	75,000	2,880	2,520	2,600
100	130	130	130	100,000	3,220	2,670	2,770
200	160	160	160	150,000	3,670	2,560	2,720
500	260	250	260	163,200	3,750	2,440	2,610
700	300	300	300	200,000	3,890	1,810	2,080
1,000	360	350	360	230,000	3,930	—	1,080
2,500	570	560	570	239,200	3,930	—	—
5,000	810	780	800	250,000	3,920	—	—
10,000	1,140	1,100	1,120	300,000	3,780	—	—
25,000	1,770	1,670	1,720	400,000	2,790	—	—
50,000	2,430	2,220	2,280	469,200	—	—	—

[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 and balance (not in central city) estimates. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero where the estimate for the total SMSA was derived from an independent estimate.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1982 Housing Inventory and for Estimated Percentages of 1978-1982 Lost Units for the New Orleans, LA, SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
200	39.7	39.7	39.7	39.7	39.7	40.6	75,000	0.2	0.4	0.9	1.3	1.8	2.1
500	20.9	20.9	20.9	20.9	22.2	25.7	100,000	0.13	0.4	0.8	1.1	1.6	1.8
700	15.9	15.9	15.9	15.9	18.8	21.7	150,000	0.09	0.3	0.6	0.9	1.3	1.5
1,000	11.7	11.7	11.7	11.7	15.7	18.2	200,000	0.07	0.3	0.6	0.8	1.1	1.3
2,500	5.0	5.0	5.0	6.9	9.9	11.5	250,000	0.05	0.2	0.5	0.7	1.0	1.1
5,000	2.6	2.6	3.5	4.9	7.0	8.1	300,000	0.04	0.2	0.5	0.6	0.9	1.0
10,000	1.3	1.3	2.5	3.4	5.0	5.7	400,000	0.03	0.2	0.4	0.5	0.8	0.9
25,000	0.5	0.7	1.6	2.2	3.1	3.6	500,000	0.03	0.2	0.4	0.5	0.7	0.8
50,000	0.3	0.5	1.1	1.5	2.2	2.6							

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

Paterson-Clifton-Passaic 1982

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1982 survey App-42
Selection of the 1978 AHS-SMSA sample App-43
Sample selection for the 1978 Coverage Improvement Program . App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1978-1982 additions to the housing inventory App-45
1970 Census of Population and Housing App-45
ESTIMATION App-45
1982 housing inventory App-45
1978-1982 lost housing units ... App-46
1978 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-47
Rounding errors App-47
Sampling errors for the AHS-SMSA sample App-47
Illustration of the use of the standard error tables App-48
Differences App-49
Illustration of the computation of the standard error of a difference App-49
Medians App-49
Illustration of the computation of the 95 percent confidence interval of a median ... App-50
Standard error tables App-51

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 12 SMSA's in this report series (H-170-82) are based on data collected from the 1982 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 12 SMSA's selected for interview during 1982 were interviewed previously in 1975 and 1978 (see the list of SMSA reports from the AHS in the introduction of this report).

Three of the larger SMSA's (Atlanta, GA; Philadelphia, PA-NJ; and San Francisco-Oakland, CA) in the 1982 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1975 and 1978 surveys, evenly divided between the central city and the balance of the respective SMSA. The 9 remaining SMSA's (Cincinnati, OH-KY-IN; Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Paterson-Clifton-Passaic, NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; and San Diego, CA) in the 1982 group were represented by a sample of 5,000 designated housing units in the 1975 and 1978 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1975 and 1978 surveys, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panels 1 through 3 were dropped from all SMSA's for the 1982 survey, excluding the central cities of the Paterson-Clifton-Passaic, NJ, SMSA. Due to additional budget limitations for the 1982 survey, it also became necessary to drop part of the sample of housing units from panels 4 through 12 in the Atlanta, GA; New Orleans, LA; Philadelphia, PA-NJ; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. The effect of this reduction was to maintain a sample size of 4,250 in all SMSA's. The interviewing was done during the period April 1982 through December 1982 with one-ninth of the sample housing units being visited each month.

In this SMSA, 4,132 housing units were eligible for interview. Of these sample housing units, 179 interviews were not obtained because, for occupied sample units, the occupants refused to be interviewed, were not at home after repeated visits, or were unavailable for some other reason; or, for vacant units, no informed respondent could be found after repeated visits. In addition to the 4,132 housing units eligible for interview, 133 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 1982 survey—The sample housing units designated to be interviewed in the 1982 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 1978 survey and remained in sample after the 1982 panel reduction.
2. All sample housing units that were type A noninterviews (i.e., units eligible to be interviewed) or type B noninterviews (i.e., units not eligible for interview at the time of the survey but which could become eligible in the future) in the 1978 survey and remained in sample after the 1982 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1982 AHS questionnaire, page App-18.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented the housing units built in permit-issuing areas since the 1978 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1978 survey.)
5. All sample housing units that were selected as part of the 1978 Coverage Improvement Program. (This sample represented most of the housing units which, until 1978, did not have a chance of selection.)

Selection of the 1978 AHS-SMSA sample—The sample for the SMSA's which, in 1970, were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). In 1970, the following four SMSA's were 100-percent permit-issuing: Paterson-Clifton-Passaic, NJ; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA. The remaining eight SMSA's contain a sample from the nonpermit universe.

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

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA during the 1970 Census of Population and Housing. This file contained records for occupied housing units, vacant housing units, and housing units in certain special places or group quarters. Sampling operations were done separately for the special place and group quarters records, and for the occupied and vacant housing unit records. Before the sample was selected from the occupied and vacant housing unit records, the occupied housing unit records were stratified by race of head (non-Black/Black) and the vacant records were stratified into four categories pertaining to the value or rent associated with the vacant housing units. The occupied housing unit records were further stratified so that each unit was assigned to 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household income	Tenure	
	Owner—Family size	Renter—Family size
	1 2 3 4 5+	1 2 3 4 5+
Under $3,000		
$3,000 to $5,999 . . .		
$6,000 to $9,999 . . .		
$10,000 to $14,999 .		
$15,000 and over . . .		

Thus, for this SMSA, the occupied housing unit records from the permit-issuing universe were assigned to 1 of 100 strata for either the central city or for the balance, and the vacant housing unit records were assigned to 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. However, whenever a record was selected to be in sample, the housing unit record adjacent to it on the file was also selected to be in sample, thereby insuring the necessary designated sample size.

Before the sample was selected from the group quarters and special place records, the records were stratified by census tract and census enumeration district (ED) within the central city and within the balance of the SMSA. A sample of special place records was then selected by a procedure that produced one-quarter of the desired sample size. However, at the time of the survey, the housing units at each of the special places were listed and subsampled at a rate which produced an expected four sample units, thereby insuring the necessary designated sample size.

The second frame from which this SMSA sample was selected was a list of new construction building permits issued since 1970 (i.e., the new construction universe). The sample selection from the list of new construction building permits was an independent operation within this SMSA. Prior to sample selection, the

list of permits was chronologically stratified by the date the permits were issued, and clusters of an expected four (usually adjacent) housing units were formed. These clusters were then sampled for inclusion at the overall sampling rate.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS sample was selected from a frame consisting of areas not under the jurisdiction of permit-issuing offices (i.e., the nonpermit universe). The first step in the sampling operation for the nonpermit universe was the selection (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

The sample ED's were then divided into segments; i.e., small land areas with well-defined boundaries having an expected size of four, or a multiple of four, housing units. At the time of the survey, those segments that did not have an expected size of four were further subdivided to produce an expected four sample housing units.

The next step was the selection of one of these segments within each sample ED. All housing units in existence at the time of interview in these selected segments were eligible for sample. Thus, housing units enumerated in the 1970 census as well as housing units built since the 1970 census are included.

Sample Selection for the 1978 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

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

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 SMSA. 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. 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 Rochester, NY, SMSA.

In the remaining 11 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. These procedures added an estimated 1,263 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 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 386 units to the coverage of the housing inventory of this SMSA.

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

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

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample unit were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and units within these structures were interviewed.

In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. This procedure added an estimated 16,107 units to the coverage of the housing inventory of this SMSA.

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

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 3,148 units to the coverage of the housing inventory of this SMSA.

1978-1982 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1978 survey, was selected to represent housing units built in permit-issuing areas since the 1978 survey. Sampling procedures were identical to those used in selecting the 1970-1978 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1977, to identify any housing units missed in the 1978 survey or any housing units added since the 1978 survey.

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 1982 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 1982 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1978 (i.e., 1978-1982 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1982 housing inventory—The AHS estimates of characteristics of the 1982 housing inventory were produced using a one-stage ratio estimation procedure for the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's. A two-stage ratio estimation procedure was used for the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. A three-stage ratio estimation procedure was used for the San Bernardino-Riverside-Ontario, CA, SMSA.

Prior to the implementation of the ratio estimation procedures, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

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

The following ratio estimation procedure was employed for all sample housing units from the permit-issuing universe. This factor was computed separately for all sample housing units within each permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

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

This ratio estimation procedure was introduced to correct the probabilities of selection for samples in each of the strata used in the sample selection of the permit-issuing universe. Prior to the AHS sample selection within each SMSA, housing units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. The same probability of selection was then applied to the remaining units to select the AHS sample. Since the number of housing units deleted from the AHS

universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the San Bernardino-Riverside-Ontario, CA, SMSA, the next ratio estimation procedure was employed for new construction mobile home units. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of mobile home units for each sector of the SMSA (central city and balance) by year built groups (1970-1974 and 1975-1982) to an independent estimate of mobile home units for the corresponding sector and year built group. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of mobile homes for the corresponding sector and year built group}}{\text{Sample estimate of mobile homes for the corresponding sector and year built group}}$$

The computed ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell.

The next ratio estimation procedure was applied in the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1982 housing inventory in each sector of the SMSA (central city and balance) to an independent estimate of the total housing units in each sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1982 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for the 12 SMSA's in 1982. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1982 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units using the existing sample weights were compared to the corresponding independent estimates for the central city, balance, and total for each of the 12 SMSA's, and the estimates which showed the most likely level of net growth since the 1980 census in both the central city and balance, as well as, the total SMSA were used in this ratio estimation procedure.

As a result of this analysis, the independent estimates were used in the Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco, CA, SMSA's. The sample estimates were used in the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's.

For the Columbus, OH and San Antonio, TX, SMSA's the independent estimates were used for the total SMSA only. The independent estimates were then proportioned between the central city and the balance of each SMSA according to the central city/balance distribution of the total housing units as given by the sample estimates. Due to the central city boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimates did not reflect the 1970 central city/balance SMSA definitions. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definitions were used.

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

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

1978-1982 lost housing units—The 1978-1982 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1978 housing inventory as was described in the 1978 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1978-1982 lost housing units existed, by definition, in the 1978 housing inventory, there was a 1978 housing inventory weight associated with each 1978-1982 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1978-1982 lost housing units.

1978 estimation procedure—This report presents data on the housing characteristics of the 1978 housing inventory from the 1978 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1978.

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

RELIABILITY OF THE ESTIMATES

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

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

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing and the 1982 AHS-SMSA sample.

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

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

AHS-SMSA—The AHS-SMSA evaluation study was administered to a sample of units selected from the regular interviews containing recent movers who had previous residences within the SMSA. The purpose of the study was to evaluate the accuracy of responses regarding housing characteristics of recent mover's previous residence. The results of which are presented in the Census Bureau memorandum "Analysis of the Results from the AHS-SMSA Group AA-2 1982 Evaluation Study" (June 17, 1983).

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

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

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all housing units located inside these ED's would be represented in the sample. However, it has been estimated that the 1975 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. It should be noted that since these ED's were recanvassed for the 1978 and 1982 surveys, the number of missed housing units may be considerably less for 1982.

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

Rounding errors—For errors associated with processing, rounding of estimates introduces another source of error in the data, the severity of which depends on the statistics being measured. The effect of rounding is significant relative to the sampling error only for small percentages or small medians, when these figures are derived from relatively large bases (e.g., median number of persons per household). This means that confidence intervals formed from the standard errors given may be distorted, and this should be taken into account when considering the results of the survey.

Sampling errors for the AHS-SMSA sample—The particular sample used for this survey is one of a large number of possible samples of the same size that could have been selected using the same sample design. Even if the same questionnaires, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling

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

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

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

The figures presented in the tables that follow (pages App-51 and App-52) are approximations to the standard errors of various estimates shown in this report for this SMSA. In order to derive standard errors that would be applicable to a wide variety of items and also could be prepared at a moderate cost, a number of approximations were required. As a result, the tables of standard errors provide an indication of the order of magnitude of the standard errors rather than precise standard errors for any specific item. Standard errors applicable to estimates of characteristics of the 1978 housing inventory can be found in the AHS Series H-170 reports for 1978.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1982 housing inventory as well as estimates of characteristics of the 1978-1982 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 15,900 for the total SMSA, 7,670 for the central city of the SMSA, and 13,930 for the balance of the SMSA.

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

Tables II and III (pages App-51 and App-52) presents the standard errors of estimated percentages for the 1982 housing inventory as well as estimated percentages of the 1978-1982 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in tables II and III.

Included in tables I, II, and III are estimates of standard errors for estimates of zero and zero percent. These estimates of standard errors are considered as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is obtained.

For ratios, 100 (x/y), where x is not a subclass of y, tables II and III underestimate the standard error of the ratio when there is little or no correlation between x and y. For this type of ratio, a better approximation of the standard error may be obtained by letting the standard error of the ratio be approximately equal to:

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

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

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1982 there were 295,500 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 4,550. The following interpolation procedure was used.

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

Size of estimate	Standard error
250,000	4,460
295,500	x
300,000	4,560

The entry of "x" is determined as follows by vertically interpolating between 4,460 and 4,560.

$$295{,}500 - 250{,}000 = 45{,}500$$
$$300{,}000 - 250{,}000 = 50{,}000$$

$$4{,}460 + \frac{45{,}500}{50{,}000}(4{,}560 - 4{,}460) = 4{,}550$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 290,950 to 300,050 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1982 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 288,220 to 302,780 housing units with 90 percent confidence; and that the average estimate lies within the interval from 286,400 to 304,600 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 295,500 owner-occupied housing units, 63,600, or 21.5 percent, had two bedrooms. Interpolation using table II of this appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 21.5 percent is approximately 0.8 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
250,000	0.7	a	1.0
295,500		p	
300,000	0.6	b	0.9

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 0.9.

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

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

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

$$295{,}500-250{,}000 = 45{,}500$$
$$300{,}000-250{,}000 = 50{,}000$$

$$0.9 + \frac{45{,}500}{50{,}000}\ (0.8-0.9) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 20.7 to 22.3 percent; the 90-percent confidence interval is from 20.2 to 22.8 percent; and the 95-percent confidence interval is from 19.9 to 23.1 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. If there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1978 and 1982 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1978 and 1982 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1982 there were 140,600 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 77,000. Table I shows the standard error of 63,600 is approximately 2,720, and the standard error of 140,600 is approximately 3,770. Therefore, the standard error of the estimated difference of 77,000 is about 4,650.

$$4{,}650 = \sqrt{(2{,}720)^2 + (3{,}770)^2}$$

Consequently, the 68-percent confidence interval for the 77,000 difference is from 72,350 to 81,650 housing units. Therefore, a conclusion that the average estimate derived from all possible samples, of this difference, lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 69,560 to 84,440 housing units, and the 95-percent confidence interval is from 67,700 to 86,300 housing units. Thus, we can conclude with 95 percent confidence that the number of 1982 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied units with two bedrooms since the 95-percent confidence interval does not include zero or negative values.

Medians—For medians presented in certain tables, the sampling error depends on the size of the base and on the distribution upon which the median is based. An approximate method for measuring the reliability of the estimated median is to determine an interval about the estimated median so that there is a stated degree of confidence that the average median from all possible samples lies within the interval. The following procedure may be used to estimate confidence limits of a median based on sample data:

1. From table II and III, determine the standard error of a 50-percent characteristic on the base of the median.
2. Add to and subtract from 50 percent, the standard error determined in step 1.
3. Using the distribution of the characteristics, determine the confidence interval corresponding to the two points established in step 2. To find the lower endpoint of the confidence

interval, it is necessary to know into which interval of the distribution the lower percentage limit falls. Similarly, to find the upper endpoint to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 3.0. The base of the distribution from which this median was determined is 295,500 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 295,500 is approximately 1.1 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50-percent twice the standard error determined in step 1. This yields percentage limits of 47.8 and 52.2.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with three persons (for purposes of calculating the median, the category of three persons is considered to be from 2.5 to 3.5 persons) contains the 47.8 percent derived in step 2. About 120,500 housing units or 40.8 percent fall below this interval, and 57,400 housing units or 19.4 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\frac{(47.8-40.8)}{19.4} = 2.9$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.2 percent derived in step 2. About 120,500 housing units or 40.8 percent fall below this interval, and 57,400 housing units or 19.4 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\frac{(52.2-40.8)}{19.4} = 3.1$$

Thus, the 95-percent confidence interval ranges from 2.9 to 3.1 persons.

TABLE I. **Standard Errors for Estimated Number of Housing Units in the 1982 Housing Inventory and for Estimated Number of 1978-1982 Lost Units for the Paterson-Clifton-Passaic, NJ, SMSA, for the Central Cities of the SMSA and for the Balance (Not in Central Cities) of the SMSA**

(68 chances out of 100)

Size of estimate	Standard error[1]			Size of estimate	Standard error[1]		
	SMSA	In central cities	Not in central cities		SMSA	In central cities	Not in central cities
0	130	100	140	75,000	2,940	2,250	2,980
100	130	100	140	98,000	3,290	2,380	3,300
200	160	140	170	100,000	3,320	—	3,330
500	260	220	260	150,000	3,870	—	3,800
700	300	260	310	200,000	4,240	—	4,030
1,000	360	310	370	250,000	4,460	—	4,080
2,500	570	500	590	300,000	4,560	—	3,940
5,000	810	700	830	377,200	4,500	—	3,320
10,000	1,130	970	1,170	400,000	4,420	—	—
25,000	1,770	1,490	1,820	475,200	3,990	—	—
50,000	2,450	1,980	2,500				

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

TABLE II. **Standard Errors for Estimated Percentages of Housing Units in the 1982 Housing Inventory and for Estimated Percentages of 1978-1982 Lost Housing Units for the Paterson-Clifton-Passaic, NJ, SMSA, and for the Balance (Not in Central Cities) of the SMSA**

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
200	41.2	41.2	41.2	41.2	41.2	41.9	75,000	0.2	0.4	0.9	1.3	1.9	2.2
500	21.9	21.9	21.9	21.9	22.9	26.5	100,000	0.14	0.4	0.8	1.1	1.6	1.9
700	16.7	16.7	16.7	16.7	19.4	22.4	150,000	0.09	0.3	0.7	0.9	1.3	1.5
1,000	12.3	12.3	12.3	12.3	16.2	18.7	200,000	0.07	0.3	0.6	0.8	1.1	1.3
2,500	5.3	5.3	5.3	7.1	10.3	11.8	250,000	0.06	0.2	0.5	0.7	1.0	1.2
5,000	2.7	2.7	3.6	5.0	7.3	8.4	300,000	0.05	0.2	0.5	0.6	0.9	1.1
10,000	1.4	1.4	2.6	3.6	5.1	5.9	400,000	0.04	0.2	0.4	0.6	0.8	0.9
25,000	0.6	0.7	1.6	2.2	3.2	3.7	500,000	0.03	0.2	0.4	0.5	0.7	0.8
50,000	0.3	0.5	1.2	1.6	2.3	2.6							

[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 1.1.

TABLE III. **Standard Errors for Estimated Percentages of Housing Units in the 1982 Housing Inventory and for Estimated Percentages of 1978-1982 Lost Housing Units for the Central Cities of the Paterson-Clifton-Passaic, NJ, SMSA**

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
200	33.2	33.2	33.2	33.2	33.2	35.2	10,000	1.0	1.0	2.2	3.0	4.3	5.0
500	16.6	16.6	16.6	16.6	19.3	22.3	25,000	0.4	0.6	1.4	1.9	2.7	3.2
700	12.4	12.4	12.4	12.4	16.3	18.8	50,000	0.2	0.4	1.0	1.3	1.9	2.2
1,000	9.0	9.0	9.0	9.5	13.6	15.8	75,000	0.13	0.4	0.8	1.1	1.6	1.8
2,500	3.8	3.8	4.3	6.0	8.6	10.0	100,000	0.10	0.3	0.7	0.9	1.4	1.6
5,000	1.9	1.9	3.1	4.2	6.1	7.0							

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

Philadelphia 1982

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1982 survey App-42
Selection of the 1978 AHS-SMSA sample App-43
Sample selection for the 1978 Coverage Improvement Program . App-44
Coverage improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1978-1982 additions to the housing inventory App-45
1970 Census of Population and Housing App-45
ESTIMATION App-45
1982 housing inventory App-45
1978-1982 lost housing units ... App-46
1978 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-47
Rounding errors App-47
Sampling errors for the AHS-SMSA sample App-47
Illustration of the use of the standard error tables App-48
Differences App-49
Illustration of the computation of the standard error of a difference App-49
Medians App-49
Illustration of the computation of the 95-percent confidence interval of a median ... App-50
Standard error tables App-51

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 12 SMSA's in this report series (H-170-82) are based on data collected from the 1982 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 12 SMSA's selected for interview during 1982 were interviewed previously in 1975 and 1978 (see the list of SMSA reports from the AHS in the introduction of this report).

Three of the larger SMSA's (Atlanta, GA; Philadelphia, PA-NJ; and San Francisco-Oakland, CA) in the 1982 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1975 and 1978 surveys, evenly divided between the central city and the balance of the respective SMSA. The 9 remaining SMSA's (Cincinnati, OH-KY-IN; Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Paterson-Clifton-Passaic, NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; and San Diego, CA) in the 1982 group were represented by a sample of 5,000 designated housing units in the 1975 and 1978 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1975 and 1978 surveys, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panels 1 through 3 were dropped from all SMSA's for the 1982 survey, excluding the central cities of the Paterson-Clifton-Passaic, NJ, SMSA. Due to additional budget limitations for the 1982 survey, it also became necessary to drop part of the sample of housing units from panels 4 through 12 in the Atlanta, GA; New Orleans, LA; Philadelphia, PA-NJ; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. The effect of this reduction was to maintain a sample size of 4,250 in all SMSA's. The interviewing was done during the period April 1982 through December 1982 with one-ninth of the sample housing units being visited each month.

In this SMSA, 3,989 housing units were eligible for interview. Of these sample housing units, 177 interviews were not obtained because, for occupied sample units, the occupants refused to be interviewed, were not at home after repeated visits, or were unavailable for some other reason; or, for vacant units, no informed respondent could be found after repeated visits. In addition to the 3,989 housing units eligible for interview, 133 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 1982 survey—The sample housing units designated to be interviewed in the 1982 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 1978 survey and remained in sample after the 1982 panel reduction.
2. All sample housing units that were type A noninterviews (i.e., units eligible to be interviewed) or type B noninterviews (i.e., units not eligible for interview at the time of the survey but which could become eligible in the future) in the 1978 survey and remained in sample after the 1982 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1982 AHS questionnaire, page App-18.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented the housing units built in permit-issuing areas since the 1978 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1978 survey.)
5. All sample housing units that were selected as part of the 1978 Coverage Improvement Program. (This sample represented most of the housing units which, until 1978, did not have a chance of selection.)

Selection of the 1978 AHS-SMSA sample—The sample for the SMSA's which, in 1970, were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). In 1970, the following four SMSA's were 100-percent permit-issuing: Paterson-Clifton-Passaic, NJ; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA. The remaining eight SMSA's contain a sample from the nonpermit universe.

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

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA during the 1970 Census of Population and Housing. This file contained records for occupied housing units, vacant housing units, and housing units in certain special places or group quarters. Sampling operations were done separately for the special place and group quarters records, and for the occupied and vacant housing unit records. Before the sample was selected from the occupied and vacant housing unit records, the occupied housing unit records were stratified by race of head (non-Black/Black) and the vacant records were stratified into four categories pertaining to the value or rent associated with the vacant housing units. The occupied housing unit records were further stratified so that each unit was assigned to 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household income	Tenure	
	Owner—Family size 1 2 3 4 5+	Renter—Family size 1 2 3 4 5+
Under $3,000		
$3,000 to $5,999 . . .		
$6,000 to $9,999 . . .		
$10,000 to $14,999 .		
$15,000 and over . . .		

Thus, for this SMSA, the occupied housing unit records from the permit-issuing universe were assigned to 1 of 100 strata for either the central city or for the balance, and the vacant housing unit records were assigned to 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. However, whenever a record was selected to be in sample, the housing unit record adjacent to it on the file was also selected to be in sample, thereby insuring the necessary designated sample size.

Before the sample was selected from the group quarters and special place records, the records were stratified by census tract and census enumeration district (ED) within the central city and within the balance of the SMSA. A sample of special place records was then selected by a procedure that produced one-quarter of the desired sample size. However, at the time of the survey, the housing units at each of the special places were listed and subsampled at a rate which produced an expected four sample units, thereby insuring the necessary designated sample size.

The second frame from which this SMSA sample was selected was a list of new construction building permits issued since 1970 (i.e., the new construction universe). The sample selection from the list of new construction building permits was an independent operation within this SMSA. Prior to sample selection, the

list of permits was chronologically stratified by the date the permits were issued, and clusters of an expected four (usually adjacent) housing units were formed. These clusters were then sampled for inclusion at the overall sampling rate.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS sample was selected from a frame consisting of areas not under the jurisdiction of permit-issuing offices (i.e., the nonpermit universe). The first step in the sampling operation for the nonpermit universe was the selection (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

The sample ED's were then divided into segments; i.e., small land areas with well-defined boundaries having an expected size of four, or a multiple of four, housing units. At the time of the survey, those segments that did not have an expected size of four were further subdivided to produce an expected four sample housing units.

The next step was the selection of one of these segments within each sample ED. All housing units in existence at the time of interview in these selected segments were eligible for sample. Thus, housing units enumerated in the 1970 census as well as housing units built since the 1970 census are included.

Sample Selection for the 1978 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

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

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 SMSA. 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. 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 Rochester, NY, SMSA.

In the remaining 11 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. These procedures added an estimated 14,180 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 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 4,219 units to the coverage of the housing inventory of this SMSA.

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

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

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample unit were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and units within these structures were interviewed.

In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. This procedure added an estimated 39,195 units to the coverage of the housing inventory of this SMSA.

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

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 6,121 units to the coverage of the housing inventory of this SMSA.

1978-1982 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1978 survey, was selected to represent housing units built in permit-issuing areas since the 1978 survey. Sampling procedures were identical to those used in selecting the 1970-1978 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1977, to identify any housing units missed in the 1978 survey or any housing units added since the 1978 survey.

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 1982 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 1982 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1978 (i.e., 1978-1982 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1982 housing inventory—The AHS estimates of characteristics of the 1982 housing inventory were produced using a one-stage ratio estimation procedure for the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's. A two-stage ratio estimation procedure was used for the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. A three-stage ratio estimation procedure was used for the San Bernardino-Riverside-Ontario, CA, SMSA.

Prior to the implementation of the ratio estimation procedures, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

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

The following ratio estimation procedure was employed for all sample housing units from the permit-issuing universe. This factor was computed separately for all sample housing units within each permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

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

This ratio estimation procedure was introduced to correct the probabilities of selection for samples in each of the strata used in the sample selection of the permit-issuing universe. Prior to the AHS sample selection within each SMSA, housing units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. The same probability of selection was then applied to the remaining units to select the AHS sample. Since the number of housing units deleted from the AHS

universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the San Bernardino-Riverside-Ontario, CA, SMSA, the next ratio estimation procedure was employed for new construction mobile home units. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of mobile home units for each sector of the SMSA (central city and balance) by year built groups (1970-1974 and 1975-1982) to an independent estimate of mobile home units for the corresponding sector and year built group. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of mobile homes for the corresponding sector and year built group}}{\text{Sample estimate of mobile homes for the corresponding sector and year built group}}$$

The computed ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell.

The next ratio estimation procedure was applied in the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1982 housing inventory in each sector of the SMSA (central city and balance) to an independent estimate of the total housing units in each sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1982 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for the 12 SMSA's in 1982. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1982 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units using the existing sample weights were compared to the corresponding independent estimates for the central city, balance, and total for each of the 12 SMSA's, and the estimates which showed the most likely level of net growth since the 1980 census in both the central city and balance, as well as, the total SMSA were used in this ratio estimation procedure.

As a result of this analysis, the independent estimates were used in the Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco, CA, SMSA's. The sample estimates were used in the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's.

For the Columbus, OH and San Antonio, TX, SMSA's the independent estimates were used for the total SMSA only. The independent estimates were then proportioned between the central city and the balance of each SMSA according to the central city/balance distribution of the total housing units as given by the sample estimates. Due to the central city boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimates did not reflect the 1970 central city/balance SMSA definitions. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definitions were used.

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

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

1978-1982 lost housing units—The 1978-1982 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1978 housing inventory as was described in the 1978 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas*. Since the 1978-1982 lost housing units existed, by definition, in the 1978 housing inventory, there was a 1978 housing inventory weight associated with each 1978-1982 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1978-1982 lost housing units.

1978 estimation procedure—This report presents data on the housing characteristics of the 1978 housing inventory from the 1978 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1978.

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

RELIABILITY OF THE ESTIMATES

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

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

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing and the 1982 AHS-SMSA sample.

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

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

AHS-SMSA—The AHS-SMSA evaluation study was administered to a sample of units selected from the regular interviews containing recent movers who had previous residences within the SMSA. The purpose of the study was to evaluate the accuracy of responses regarding housing characteristics of recent mover's previous residence. The results of which are presented in the Census Bureau memorandum "Analysis of the Results from the AHS-SMSA Group AA-2 1982 Evaluation Study." (June 17, 1983)

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

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

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all housing units located inside these ED's would be represented in the sample. However, it has been estimated that the 1975 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. It should be noted that since these ED's were recanvassed for the 1978 and 1982 surveys, the number of missed housing units may be considerably less for 1982.

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

Rounding errors—For errors associated with processing, rounding of estimates introduces another source of error in the data, the severity of which depends on the statistics being measured. The effect of rounding is significant relative to the sampling error only for small percentages or small medians, when these figures are derived from relatively large bases (e.g., median number of persons per household). This means that confidence intervals formed from the standard errors given may be distorted, and this should be taken into account when considering the results of the survey.

Sampling errors for the AHS-SMSA sample—The particular sample used for this survey is one of a large number of possible samples of the same size that could have been selected using the same sample design. Even if the same questionnaires, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling

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

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

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

The figures presented in the tables that follow (page App-51) are approximations to the standard errors of various estimates shown in this report for this SMSA. In order to derive standard errors that would be applicable to a wide variety of items and also could be prepared at a moderate cost, a number of approximations were required. As a result, the tables of standard errors provide an indication of the order of magnitude of the standard errors rather than precise standard errors for any specific item. Standard errors applicable to estimates of characteristics of the 1978 housing inventory can be found in the AHS Series H-170 reports for 1978.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1982 housing inventory as well as estimates of characteristics of the 1978-1982 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 69,340 for the total SMSA, 29,780 for the central city of the SMSA, and 63,410 for the balance of the SMSA.

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

Table II (page App-51) presents the standard errors of estimated percentages for the 1982 housing inventory as well as estimated percentages of the 1978-1982 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

Included in tables I and II are estimates of standard errors for estimates of zero and zero percent. These estimates of standard errors are considered as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is obtained.

For ratios, 100 (x/y), where x is not a subclass of y, table II underestimates the standard error of the ratio when there is little or no correlation between x and y. For this type of ratio, a better approximation of the standard error may be obtained by letting the standard error of the ratio be approximately equal to:

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

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

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1982 there were 1,134,400 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 13,920. The following interpolation procedure was used.

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

Size of estimate	Standard error
1,100,000	14,090
1,134,400	x
1,200,000	13,590

The entry of "x" is determined as follows by vertically interpolating between 14,090 and 13,590.

$$1{,}134{,}400 - 1{,}100{,}000 = 34{,}400$$
$$1{,}200{,}000 - 1{,}100{,}000 = 100{,}000$$

$$14{,}090 + \frac{34{,}400}{100{,}000}(13{,}590 - 14{,}090) = 13{,}920$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 1,120,480 to 1,148,320 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1982 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 1,112,130 to 1,156,670 housing units with 90 percent confidence; and that the average estimate lies within the interval from 1,106,560 to 1,162,240 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 1,134,400 owner-occupied housing units, 172,900, or 15.2 percent, had two bedrooms. Interpolation using table II of this appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 15.2 percent is approximately 0.7 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	10 or 90	15.2	25 or 75
1,100,000	0.6	a	0.9
1,134,400		p	
1,200,000	0.6	b	0.9

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

$$15.2-10.0 = 5.2$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{5.2}{15.0}\ (0.9-0.6) = 0.7$$

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

$$15.2-10.0 = 5.2$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{5.2}{15.0}\ (0.9-0.6) = 0.7$$

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

$$1{,}134{,}400-1{,}100{,}000 = 34{,}400$$
$$1{,}200{,}000-1{,}100{,}000 = 100{,}000$$

$$0.7 + \frac{34{,}400}{100{,}000}\ (0.7-0.7) = 0.7$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 14.5 to 15.9 percent; the 90-percent confidence interval is from 14.1 to 16.3 percent; and the 95-percent confidence interval is from 13.8 to 16.6 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. If there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1978 and 1982 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1978 and 1982 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1982 there were 625,600 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 452,700. Table I shows the standard error of 172,900 is approximately 8,560, and the standard error of 625,600 is approximately 13,820. Therefore, the standard error of the estimated difference of 452,700 is about 16,260:

$$16{,}260 = \sqrt{(8{,}560)^2 + (13{,}820)^2}$$

Consequently, the 68-percent confidence interval for the 452,700 difference is from 436,440 to 468,960 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 426,680 to 478,720 housing units, and the 95-percent confidence interval is from 420,180 to 485,220 housing units. Thus, we can conclude with 95 percent confidence that the number of 1982 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied units with two bedrooms since the 95-percent confidence interval does not include zero or negative values.

Medians—For medians presented in certain tables, the sampling error depends on the size of the base and on the distribution upon which the median is based. An approximate method for measuring the reliability of the estimated median is to determine an interval about the estimated median so that there is a stated degree of confidence that the average median from all possible samples lies within the interval. The following procedure may be used to estimate confidence limits of a median based on sample data:

1. From table II, determine the standard error of a 50-percent characteristic on the base of the median.
2. Add to and subtract from 50 percent, the standard error determined in step 1.
3. Using the distribution of the characteristics, determine the confidence interval corresponding to the two points established in step 2. To find the lower endpoint of the confidence

interval, it is necessary to know into which interval of the distribution the lower percentage limit falls. Similarly, to find the upper endpoint to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median— Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.9. The base of the distribution from which this median was determined is 1,134,400 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 1,134,400 is approximately 1.1 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50-percent twice the standard error determined in step 1. This yields percentage limits of 47.8 and 52.2.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with three persons (for purposes of calculating the median, the category of three persons is considered to be from 2.5 to 3.5 persons) contains the 47.8 percent derived in step 2. About 480,100 housing units or 42.3 percent fall below this interval, and 229,300 housing units or 20.2 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(47.8-42.3)}{20.2} = 2.8$$

Similarly, the interval for owner occupied housing units with three persons contains the 52.2 percent derived in step 2. About 480,100 housing units or 42.3 percent fall below this interval, and 229,300 housing units or 20.2 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(52.2-42.3)}{20.2} = 3.0$$

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

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1982 Housing Inventory and for Estimated Number of 1978-1982 Lost Units for the Philadelphia, PA-NJ, SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	470	500	450	400,000	12,100	9,120	10,730
100	470	500	450	500,000	13,030	8,240	11,100
200	470	500	450	600,000	13,700	6,180	11,080
500	490	500	480	688,300	14,120	–	10,750
700	570	590	560	700,000	14,160	–	10,680
1,000	690	700	670	800,000	14,420	–	9,860
2,500	1,090	1,110	1,060	900,000	14,500	–	8,480
5,000	1,530	1,570	1,500	1,000,000	14,390	–	6,190
10,000	2,170	2,210	2,120	1,092,300	14,120	–	–
25,000	3,410	3,460	3,330	1,100,000	14,090	–	–
50,000	4,790	4,800	4,650	1,200,000	13,590	–	–
75,000	5,820	5,760	5,630	1,300,000	12,870	–	–
100,000	6,680	6,510	6,420	1,400,000	11,890	–	–
150,000	8,050	7,630	7,660	1,500,000	10,560	–	–
200,000	9,160	8,390	8,610	1,750,000	3,770	–	–
250,000	10,070	8,890	9,360	1,780,700	–	–	–
300,000	10,850	9,170	9,940				

[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 balance (not in central city) and 1.2 for the central city estimates. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero where the estimate for the total SMSA was derived from an independent estimate.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1982 Housing Inventory and for Estimated Percentages of 1978-1982 Lost Housing Units for the Philadelphia, PA-NJ, SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
500	49.8	49.8	49.8	49.8	49.8	49.8	400,000	0.12	0.4	0.8	1.1	1.5	1.8
700	41.5	41.5	41.5	41.5	41.5	42.1	500,000	0.10	0.3	0.7	0.9	1.4	1.6
1,000	33.2	33.2	33.2	33.2	33.2	35.2	600,000	0.08	0.3	0.6	0.9	1.2	1.4
2,500	16.6	16.6	16.6	16.6	19.3	22.3	700,000	0.07	0.3	0.6	0.8	1.2	1.3
5,000	9.0	9.0	9.0	9.5	13.6	15.8	800,000	0.06	0.2	0.5	0.7	1.1	1.2
10,000	4.7	4.7	4.9	6.7	9.6	11.1	900,000	0.06	0.2	0.5	0.7	1.0	1.2
25,000	1.9	1.9	3.1	4.2	6.1	7.0	1,000,000	0.05	0.2	0.5	0.7	1.0	1.1
50,000	1.0	1.0	2.2	3.0	4.3	5.0	1,100,000	0.05	0.2	0.5	0.6	0.9	1.1
75,000	0.7	0.8	1.8	2.4	3.5	4.1	1,200,000	0.04	0.2	0.4	0.6	0.9	1.0
100,000	0.5	0.7	1.5	2.1	3.1	3.5	1,300,000	0.04	0.2	0.4	0.6	0.8	1.0
150,000	0.3	0.6	1.3	1.7	2.5	2.9	1,400,000	0.04	0.2	0.4	0.6	0.8	0.9
200,000	0.2	0.5	1.1	1.5	2.2	2.5	1,500,000	0.03	0.2	0.4	0.5	0.8	0.9
250,000	0.2	0.4	1.0	1.3	1.9	2.2	1,750,000	0.03	0.2	0.4	0.5	0.7	0.8
300,000	0.2	0.4	0.9	1.2	1.8	2.0	2,000,000	0.02	0.2	0.3	0.5	0.7	0.8

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

Riverside-San Bernardino-Ontario 1982

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1982 survey App-42
Selection of the 1978 AHS-SMSA sample App-43
Sample selection for the 1978 Coverage Improvement Program . App-44
Coverage Improvement for deficiency 1 ..... App-44
Coverage improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1978-1982 additions to the housing inventory App-45
1970 Census of Population and Housing App-45
ESTIMATION App-45
1982 housing inventory App-45
1978-1982 lost housing units ... App-46
1978 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-47
Rounding errors App-47
Sampling errors for the AHS-SMSA sample App-47
Illustration of the use of the standard error tables App-48
Differences App-49
Illustration of the computation of the standard error of a difference App-49
Medians App-49
Illustration of the computation of the 95-percent confidence interval of a median ... App-50
Standard error tables App-51

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 12 SMSA's in this report series (H-170-82) are based on data collected from the 1982 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 12 SMSA's selected for interview during 1982 were interviewed previously in 1975 and 1978 (see the list of SMSA reports from the AHS in the introduction of this report).

Three of the larger SMSA's (Atlanta, GA; Philadelphia, PA-NJ; and San Francisco-Oakland, CA) in the 1982 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1975 and 1978 surveys, evenly divided between the central city and the balance of the respective SMSA. The 9 remaining SMSA's (Cincinnati, OH-KY-IN; Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Paterson-Clifton-Passaic, NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; and San Diego, CA) in the 1982 group were represented by a sample of 5,000 designated housing units in the 1975 and 1978 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1975 and 1978 surveys, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panels 1 through 3 were dropped from all SMSA's for the 1982 survey, excluding the central cities of the Paterson-Clifton-Passaic, NJ, SMSA. Due to additional budget limitations for the 1982 survey, it also became necessary to drop part of the sample of housing units from panels 4 through 12 in the Atlanta, GA; New Orleans, LA; Philadelphia, PA-NJ; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. The effect of this reduction was to maintain a sample size of 4,250 in all SMSA's. The interviewing was done during the period April 1982 through December 1982 with one-ninth of the sample housing units being visited each month.

In this SMSA, 3,854 housing units were eligible for interview. Of these sample housing units, 189 interviews were not obtained because, for occupied sample units, the occupants refused to be interviewed, were not at home after repeated visits, or were unavailable for some other reason; or, for vacant units, no informed respondent could be found after repeated visits. In addition to the 3,854 housing units eligible for interview, 277 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 1982 survey—The sample housing units designated to be interviewed in the 1982 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 1978 survey and remained in sample after the 1982 panel reduction.
2. All sample housing units that were type A noninterviews (i.e., units eligible to be interviewed) or type B noninterviews (i.e., units not eligible for interview at the time of the survey but which could become eligible in the future) in the 1978 survey and remained in sample after the 1982 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1982 AHS questionnaire, page App-18.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented the housing units built in permit-issuing areas since the 1978 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1978 survey.)
5. All sample housing units that were selected as part of the 1978 Coverage Improvement Program. (This sample represented most of the housing units which, until 1978, did not have a chance of selection.)

Selection of the 1978 AHS-SMSA sample—The sample for the SMSA's which, in 1970, were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). In 1970, the following four SMSA's were 100-percent permit-issuing: Paterson-Clifton-Passaic, NJ; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA. The remaining eight SMSA's contain a sample from the nonpermit universe.

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

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA during the 1970 Census of Population and Housing. This file contained records for occupied housing units, vacant housing units, and housing units in certain special places or group quarters. Sampling operations were done separately for the special place and group quarters records, and for the occupied and vacant housing unit records. Before the sample was selected from the occupied and vacant housing unit records, the occupied housing unit records were stratified by race of head (non-Black/Black) and the vacant records were stratified into four categories pertaining to the value or rent associated with the vacant housing units. The occupied housing unit records were further stratified so that each unit was assigned to 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household income	Tenure	
	Owner—Family size	Renter—Family size
	1 2 3 4 5+	1 2 3 4 5+
Under $3,000		
$3,000 to $5,999 . . .		
$6,000 to $9,999 . . .		
$10,000 to $14,999 .		
$15,000 and over . . .		

Thus, for this SMSA, the occupied housing unit records from the permit-issuing universe were assigned to 1 of 100 strata for either the central city or for the balance, and the vacant housing unit records were assigned to 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. However, whenever a record was selected to be in sample, the housing unit record adjacent to it on the file was also selected to be in sample, thereby insuring the necessary designated sample size.

Before the sample was selected from the group quarters and special place records, the records were stratified by census tract and census enumeration district (ED) within the central city and within the balance of the SMSA. A sample of special place records was then selected by a procedure that produced one-quarter of the desired sample size. However, at the time of the survey, the housing units at each of the special places were listed and subsampled at a rate which produced an expected four sample units, thereby insuring the necessary designated sample size.

The second frame from which this SMSA sample was selected was a list of new construction building permits issued since 1970 (i.e., the new construction universe). The sample selection from the list of new construction building permits was an independent operation within this SMSA. Prior to sample selection, the

list of permits was chronologically stratified by the date the permits were issued, and clusters of an expected four (usually adjacent) housing units were formed. These clusters were then sampled for inclusion at the overall sampling rate.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS sample was selected from a frame consisting of areas not under the jurisdiction of permit-issuing offices (i.e., the nonpermit universe). The first step in the sampling operation for the nonpermit universe was the selection (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

The sample ED's were then divided into segments; i.e., small land areas with well-defined boundaries having an expected size of four, or a multiple of four, housing units. At the time of the survey, those segments that did not have an expected size of four were further subdivided to produce an expected four sample housing units.

The next step was the selection of one of these segments within each sample ED. All housing units in existence at the time of interview in these selected segments were eligible for sample. Thus, housing units enumerated in the 1970 census as well as housing units built since the 1970 census are included.

Sample Selection for the 1978 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

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

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 SMSA. 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. 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 Rochester, NY, SMSA.

In the remaining 11 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. These procedures added an estimated 1,247 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 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 19,046 units to the coverage of the housing inventory of this SMSA.

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

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

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample unit were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and units within these structures were interviewed.

In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. This procedure added an estimated 14,165 units to the coverage of the housing inventory of this SMSA.

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

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 1,039 units to the coverage of the housing inventory of this SMSA.

1978-1982 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1978 survey, was selected to represent housing units built in permit-issuing areas since the 1978 survey. Sampling procedures were identical to those used in selecting the 1970-1978 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1977, to identify any housing units missed in the 1978 survey or any housing units added since the 1978 survey.

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 1982 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 1982 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1978 (i.e., 1978-1982 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1982 housing inventory—The AHS estimates of characteristics of the 1982 housing inventory were produced using a one-stage ratio estimation procedure for the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's. A two-stage ratio estimation procedure was used for the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. A three-stage ratio estimation procedure was used for the San Bernardino-Riverside-Ontario, CA, SMSA.

Prior to the implementation of the ratio estimation procedures, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

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

The following ratio estimation procedure was employed for all sample housing units from the permit-issuing universe. This factor was computed separately for all sample housing units within each permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

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

This ratio estimation procedure was introduced to correct the probabilities of selection for samples in each of the strata used in the sample selection of the permit-issuing universe. Prior to the AHS sample selection within each SMSA, housing units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. The same probability of selection was then applied to the remaining units to select the AHS sample. Since the number of housing units deleted from the AHS

universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the San Bernardino-Riverside-Ontario, CA, SMSA, the next ratio estimation procedure was employed for new construction mobile home units. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of mobile home units for each sector of the SMSA (central city and balance) by year built groups (1970-1974 and 1975-1982) to an independent estimate of mobile home units for the corresponding sector and year built group. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of mobile homes for the corresponding sector and year built group}}{\text{Sample estimate of mobile homes for the corresponding sector and year built group}}$$

The computed ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell.

The next ratio estimation procedure was applied in the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1982 housing inventory in each sector of the SMSA (central city and balance) to an independent estimate of the total housing units in each sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1982 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for the 12 SMSA's in 1982. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1982 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units using the existing sample weights were compared to the corresponding independent estimates for the central city, balance, and total for each of the 12 SMSA's, and the estimates which showed the most likely level of net growth since the 1980 census in both the central city and balance, as well as, the total SMSA were used in this ratio estimation procedure.

As a result of this analysis, the independent estimates were used in the Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco, CA, SMSA's. The sample estimates were used in the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's.

For the Columbus, OH and San Antonio, TX, SMSA's the independent estimates were used for the total SMSA only. The independent estimates were then proportioned between the central city and the balance of each SMSA according to the central city/balance distribution of the total housing units as given by the sample estimates. Due to the central city boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimates did not reflect the 1970 central city/balance SMSA definitions. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definitions were used.

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

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

1978-1982 lost housing units—The 1978-1982 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1978 housing inventory as was described in the 1978 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas*. Since the 1978-1982 lost housing units existed, by definition, in the 1978 housing inventory, there was a 1978 housing inventory weight associated with each 1978-1982 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1978-1982 lost housing units.

1978 estimation procedure—This report presents data on the housing characteristics of the 1978 housing inventory from the 1978 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1978.

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

RELIABILITY OF THE ESTIMATES

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

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

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing and the 1982 AHS-SMSA sample.

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

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

AHS-SMSA—The AHS-SMSA evaluation study was administered to a sample of units selected from the regular interviews containing recent movers who had previous residences within the SMSA. The purpose of the study was to evaluate the accuracy of responses regarding housing characteristics of recent mover's previous residence. The results of which are presented in the Census Bureau memorandum "Analysis of the Results from the AHS-SMSA Group AA-2 1982 Evaluation Study" (June 17, 1983).

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

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

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all housing units located inside these ED's would be represented in the sample. However, it has been estimated that the 1975 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. It should be noted that since these ED's were recanvassed for the 1978 and 1982 surveys, the number of missed housing units may be considerably less for 1982.

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

Rounding errors—For errors associated with processing, rounding of estimates introduces another source of error in the data, the severity of which depends on the statistics being measured. The effect of rounding is significant relative to the sampling error only for small percentages or small medians, when these figures are derived from relatively large bases (e.g., median number of persons per household). This means that confidence intervals formed from the standard errors given may be distorted, and this should be taken into account when considering the results of the survey.

Sampling errors for the AHS-SMSA sample—The particular sample used for this survey is one of a large number of possible samples of the same size that could have been selected using the same sample design. Even if the same questionnaires, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling

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

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

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

The figures presented in the tables that follow (pages App-51 and App-52) are approximations to the standard errors of various estimates shown in this report for this SMSA. In order to derive standard errors that would be applicable to a wide variety of items and also could be prepared at a moderate cost, a number of approximations were required. As a result, the tables of standard errors provide an indication of the order of magnitude of the standard errors rather than precise standard errors for any specific item. Standard errors applicable to estimates of characteristics of the 1978 housing inventory can be found in the AHS Series H-170 reports for 1978.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1982 housing inventory as well as estimates of characteristics of the 1978-1982 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 29,900 for the total SMSA, 12,720 for the central city of the SMSA, and 27,060 for the balance of the SMSA.

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

Tables II and III (pages App-51 and App-52) presents the standard errors of estimated percentages for the 1982 housing inventory as well as estimated percentages of the 1978-1982 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in tables II and III.

Included in tables I, II, and III are estimates of standard errors for estimates of zero and zero percent. These estimates of standard errors are considered as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is obtained.

For ratios, 100 (x/y), where x is not a subclass of y, tables II and III underestimate the standard error of the ratio when there is little or no correlation between x and y. For this type of ratio, a better approximation of the standard error may be obtained by letting the standard error of the ratio be approximately equal to:

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

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

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1982 there were 392,300 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 5,880. 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	5,910
392,300	x
400,000	5,880

The entry of "x" is determined as follows by vertically interpolating between 5,910 and 5,880.

$$392{,}300 - 300{,}000 = 92{,}300$$
$$400{,}000 - 300{,}000 = 100{,}000$$

$$5{,}910 + \frac{92{,}300}{100{,}000}(5{,}880 - 5{,}910) = 5{,}880$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 386,420 to 398,180 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1982 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 382,890 to 401,710 housing units with 90 percent confidence; and that the average estimate lies within the interval from 380,540 to 404,060 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 392,300 owner-occupied housing units, 131,600, or 33.5 percent, had two bedrooms. Interpolation using table II of this appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 33.5 percent is approximately 1.1 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	33.5	50
300,000	1.2	a	1.4
392,300		p	
400,000	1.0	b	1.2

1. The entry for cell "a" is determined by horizontal interpolation between 1.2 and 1.4.

$$33.5-25.0 = 8.5$$
$$50.0-25.0 = 25.0$$

$$1.2 + \frac{8.5}{25.0}\ (1.4-1.2) = 1.3$$

2. The entry for cell "b" is determined by horizontal interpolation between 1.0 and 1.2.

$$33.5-25.0 = 8.5$$
$$50.0-25.0 = 25.0$$

$$1.0 + \frac{8.5}{25.0}\ (1.2-1.0) = 1.1$$

3. The entry for "p" is then determined by vertical interpolation between 1.3 and 1.1.

$$392{,}300-300{,}000 = 92{,}300$$
$$400{,}000-300{,}000 = 100{,}000$$

$$1.3 + \frac{92{,}300}{100{,}000}\ (1.1-1.3) = 1.1$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 32.4 to 34.6 percent; the 90-percent confidence interval is from 31.7 to 35.3 percent; and the 95-percent confidence interval is from 31.3 to 35.7 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. If there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1978 and 1982 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1978 and 1982 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1982 there were 174,100 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 42,500. Table I shows the standard error of 131,600 is approximately 4,660 and the standard error of 174,100 is approximately 5,160. Therefore, the standard error of the estimated difference of 42,500 is about 6,950:

$$6{,}950 = \sqrt{(4{,}660)^2 + (5{,}160)^2}$$

Consequently, the 68-percent confidence interval for the 42,500 difference is from 35,500 to 49,450 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 31,380 to 53,620 housing units, and the 95-percent confidence interval is from 28,600 to 56,400 housing units. Thus, we can conclude with 95 percent confidence that the number of 1982 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied units with two bedrooms since the 95-percent confidence interval does not include zero or negative values.

Medians—For medians presented in certain tables, the sampling error depends on the size of the base and on the distribution upon which the median is based. An approximate method for measuring the reliability of the estimated median is to determine an interval about the estimated median so that there is a stated degree of confidence that the average median from all possible samples lies within the interval. The following procedure may be used to estimate confidence limits of a median based on sample data:

1. From tables II and III, determine the standard error of a 50-percent characteristic on the base of the median.
2. Add to and subtract from 50 percent, the standard error determined in step 1.
3. Using the distribution of the characteristics, determine the confidence interval corresponding to the two points established in step 2. To find the lower endpoint of the confidence

interval, it is necessary to know into which interval of the distribution the lower percentage limit falls. Similarly, to find the upper endpoint to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median— Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.4. The base of the distribution from which this median was determined is 392,300 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 392,300 is approximately 1.2 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50-percent twice the standard error determined in step 1. This yields percentage limits of 47.6 and 52.4.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) contains the 47.6 percent derived in step 2. About 67,200 housing units or 17.1 percent fall below this interval, and 146,900 housing units or 37.4 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$1.5 + (2.5 - 1.5)\,\frac{(47.6 - 17.1)}{37.4} = 2.3$$

Similarly, the interval for owner-occupied housing units with two persons contains the 52.4 percent derived in step 2. About 67,200 housing units or 17.1 percent fall below this interval, and 146,900 housing units or 37.4 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$1.5 + (2.5 - 1.5)\,\frac{(52.4 - 17.1)}{37.4} = 2.4$$

Thus, the 95-percent confidence interval ranges from 2.3 to 2.4 persons. Although it appears that this confidence interval has the same estimate as the higher limit, it actually is a reflection on the rounding error associated with this median (see the paragraph on rounding errors in the nonsampling errors section of this appendix).

TABLE I. **Standard Errors for Estimated Number of Housing Units in the 1982 Housing Inventory and for Estimated Number of 1978-1982 Lost Units for the San Bernardino-Riverside-Ontario, CA, SMSA, for the Central Cities of the SMSA and for the Balance (Not in Central Cities) of the SMSA**

(68 chances out of 100)

Size of estimate	Standard error[1]			Size of estimate	Standard error[1]		
	SMSA	In central cities	Not in central cities		SMSA	In central cities	Not in central cities
0	210	160	220	100,000	4,200	2,280	4,280
100	210	160	220	147,000	4,890	–	4,910
200	210	180	220	150,000	4,920	–	4,940
500	320	280	330	200,000	5,410	–	5,320
700	380	340	400	250,000	5,730	–	5,500
1,000	450	400	470	300,000	5,910	–	5,480
2,500	720	630	750	400,000	5,880	–	4,850
5,000	1,010	890	1,050	500,000	5,320	–	2,940
10,000	1,430	1,230	1,480	541,800	4,890	–	–
25,000	2,230	1,840	2,310	600,000	4,000	–	–
50,000	3,090	2,320	3,190	688,800	–	–	–
75,000	3,710	2,440	3,810				

[1]For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2 for the total SMSA and balance (not in central cities) and 1.3 for the central cities estimates. Standard errors of estimates pertaining to total housing units for the central cities, balance, and total SMSA are assumed to be equal to zero where the estimate for the total SMSA was derived from an independent estimate.

TABLE II. **Standard Errors for Estimated Percentages of Housing Units in the 1982 Housing Inventory and for Estimated Percentages of 1978-1982 Lost Housing Units for the San Bernardino-Riverside-Ontario, CA, SMSA and for the Balance (Not in Central Cities) of the SMSA**

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
300	42.8	42.8	42.8	42.8	42.8	43.3	100,000	0.2	0.5	1.0	1.4	2.1	2.4
500	31.0	31.0	31.0	31.0	31.0	33.5	150,000	0.15	0.4	0.8	1.2	1.7	1.9
700	24.3	24.3	24.3	24.3	24.5	28.3	200,000	0.11	0.3	0.7	1.0	1.5	1.7
1,000	18.3	18.3	18.3	18.3	20.5	23.7	250,000	0.09	0.3	0.7	0.9	1.3	1.5
2,500	8.2	8.2	8.2	9.0	13.0	15.0	300,000	0.07	0.3	0.6	0.8	1.2	1.4
5,000	4.3	4.3	4.6	6.4	9.2	10.6	400,000	0.06	0.2	0.5	0.7	1.0	1.2
10,000	2.2	2.2	3.3	4.5	6.5	7.5	500,000	0.04	0.2	0.5	0.6	0.9	1.1
25,000	0.9	0.9	2.1	2.8	4.1	4.7	600,000	0.04	0.2	0.4	0.6	0.8	1.0
50,000	0.4	0.7	1.5	2.0	2.9	3.4	700,000	0.03	0.2	0.4	0.5	0.8	0.9
75,000	0.3	0.5	1.2	1.6	2.4	2.7							

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

TABLE III. **Standard Errors for Estimated Percentages of Housing Units in the 1982 Housing Inventory and for Estimated Percentages of 1978-1982 Lost Units for the Central Cities of the San Bernardino-Riverside-Ontario, CA, SMSA**

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
300	35.1	35.1	35.1	35.1	35.1	36.8	10,000	1.6	1.6	2.8	3.8	5.5	6.4
500	24.5	24.5	24.5	24.5	24.7	28.5	25,000	0.6	0.8	1.8	2.4	3.5	4.0
700	18.8	18.8	18.8	18.8	20.9	24.1	50,000	0.3	0.6	1.2	1.7	2.5	2.9
1,000	14.0	14.0	14.0	14.0	17.5	20.2	75,000	0.2	0.5	1.0	1.4	2.0	2.3
2,500	6.1	6.1	6.1	7.6	11.0	12.7	100,000	0.2	0.4	0.9	1.2	1.7	2.0
5,000	3.1	3.1	3.9	5.4	7.8	9.0	150,000	0.11	0.3	0.7	1.0	1.4	1.6

[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.

Rochester 1982

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN . . . App-42
Annual Housing Survey . . . App-42
Designation of sample housing units for the 1982 survey . . . App-42
Selection of the 1978 AHS-SMSA sample . . . App-43
Sample selection for the 1978 Coverage Improvement Program . App-44
Coverage improvement for deficiency 1 . . . App-44
Coverage improvement for deficiency 2 . . . App-44
Coverage improvement for deficiencies 3-6 . . . App-44
1978-1982 additions to the housing inventory . . . App-45
1970 Census of Population and Housing . . . App-45
ESTIMATION . . . App-45
1982 housing inventory . . . App-45
1978-1982 lost housing units . . . App-46
1978 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-47
Rounding errors . . . App-47
Sampling errors for the AHS-SMSA sample . . . App-47
Illustration of the use of the standard error tables . . . App-48
Differences . . . App-49
Illustration of the computation of the standard error of a difference . . . App-49
Medians . . . App-49
Illustration of the computation of the 95-percent confidence interval of a median . . . App-50
Standard error tables . . . App-51

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 12 SMSA's in this report series (H-170-82) are based on data collected from the 1982 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 12 SMSA's selected for interview during 1982 were interviewed previously in 1975 and 1978 (see the list of SMSA reports from the AHS in the introduction of this report).

Three of the larger SMSA's (Atlanta, GA; Philadelphia, PA-NJ; and San Francisco-Oakland, CA) in the 1982 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1975 and 1978 surveys, evenly divided between the central city and the balance of the respective SMSA. The 9 remaining SMSA's (Cincinnati, OH-KY-IN; Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Paterson-Clifton-Passaic, NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; and San Diego, CA) in the 1982 group were represented by a sample of 5,000 designated housing units in the 1975 and 1978 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1975 and 1978 surveys, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panels 1 through 3 were dropped from all SMSA's for the 1982 survey, excluding the central cities of the Paterson-Clifton-Passaic, NJ, SMSA. Due to additional budget limitations for the 1982 survey, it also became necessary to drop part of the sample of housing units from panels 4 through 12 in the Atlanta, GA; New Orleans, LA; Philadelphia, PA-NJ; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. The effect of this reduction was to maintain a sample size of 4,250 in all SMSA's. The interviewing was done during the period April 1982 through December 1982 with one-ninth of the sample housing units being visited each month.

In this SMSA, 3,884 housing units were eligible for interview. Of these sample housing units, 119 interviews were not obtained because, for occupied sample units, the occupants refused to be interviewed, were not at home after repeated visits, or were unavailable for some other reason; or, for vacant units, no informed respondent could be found after repeated visits. In addition to the 3,884 housing units eligible for interview, 326 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 1982 survey—The sample housing units designated to be interviewed in the 1982 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 1978 survey and remained in sample after the 1982 panel reduction.
2. All sample housing units that were type A noninterviews (i.e., units eligible to be interviewed) or type B noninterviews (i.e., units not eligible for interview at the time of the survey but which could become eligible in the future) in the 1978 survey and remained in sample after the 1982 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1982 AHS questionnaire, page App-18.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented the housing units built in permit-issuing areas since the 1978 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1978 survey.)
5. All sample housing units that were selected as part of the 1978 Coverage Improvement Program. (This sample represented most of the housing units which, until 1978, did not have a chance of selection.)

Selection of the 1978 AHS-SMSA sample—The sample for the SMSA's which, in 1970, were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). In 1970, the following four SMSA's were 100-percent permit-issuing: Paterson-Clifton-Passaic, NJ; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA. The remaining eight SMSA's contain a sample from the nonpermit universe.

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

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA during the 1970 Census of Population and Housing. This file contained records for occupied housing units, vacant housing units, and housing units in certain special places or group quarters. Sampling operations were done separately for the special place and group quarters records, and for the occupied and vacant housing unit records. Before the sample was selected from the occupied and vacant housing unit records, the occupied housing unit records were stratified by race of head (non-Black/Black) and the vacant records were stratified into four categories pertaining to the value or rent associated with the vacant housing units. The occupied housing unit records were further stratified so that each unit was assigned to 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household income	Tenure	
	Owner—Family size	Renter—Family size
	1 2 3 4 5+	1 2 3 4 5+
Under $3,000		
$3,000 to $5,999 . . .		
$6,000 to $9,999 . . .		
$10,000 to $14,999 .		
$15,000 and over . . .		

Thus, for this SMSA, the occupied housing unit records from the permit-issuing universe were assigned to 1 of 100 strata for either the central city or for the balance, and the vacant housing unit records were assigned to 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. However, whenever a record was selected to be in sample, the housing unit record adjacent to it on the file was also selected to be in sample, thereby insuring the necessary designated sample size.

Before the sample was selected from the group quarters and special place records, the records were stratified by census tract and census enumeration district (ED) within the central city and within the balance of the SMSA. A sample of special place records was then selected by a procedure that produced one-quarter of the desired sample size. However, at the time of the survey, the housing units at each of the special places were listed and subsampled at a rate which produced an expected four sample units, thereby insuring the necessary designated sample size.

The second frame from which this SMSA sample was selected was a list of new construction building permits issued since 1970 (i.e., the new construction universe). The sample selection from the list of new construction building permits was an independent operation within this SMSA. Prior to sample selection, the

list of permits was chronologically stratified by the date the permits were issued, and clusters of an expected four (usually adjacent) housing units were formed. These clusters were then sampled for inclusion at the overall sampling rate.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS sample was selected from a frame consisting of areas not under the jurisdiction of permit-issuing offices (i.e., the nonpermit universe). The first step in the sampling operation for the nonpermit universe was the selection (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

The sample ED's were then divided into segments; i.e., small land areas with well-defined boundaries having an expected size of four, or a multiple of four, housing units. At the time of the survey, those segments that did not have an expected size of four were further subdivided to produce an expected four sample housing units.

The next step was the selection of one of these segments within each sample ED. All housing units in existence at the time of interview in these selected segments were eligible for sample. Thus, housing units enumerated in the 1970 census as well as housing units built since the 1970 census are included.

Sample Selection for the 1978 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

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

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 SMSA. 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. 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 Rochester, NY, SMSA.

In the remaining 11 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. These procedures added an estimated 8,182 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 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 629 units to the coverage of the housing inventory of this SMSA.

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

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

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample unit were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and units within these structures were interviewed.

In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. This procedure added an estimated 5,433 units to the coverage of the housing inventory of this SMSA.

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

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 2,354 units to the coverage of the housing inventory of this SMSA.

1978-1982 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1978 survey, was selected to represent housing units built in permit-issuing areas since the 1978 survey. Sampling procedures were identical to those used in selecting the 1970-1978 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1977, to identify any housing units missed in the 1978 survey or any housing units added since the 1978 survey.

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 1982 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 1982 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1978 (i.e., 1978-1982 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1982 housing inventory—The AHS estimates of characteristics of the 1982 housing inventory were produced using a one-stage ratio estimation procedure for the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's. A two-stage ratio estimation procedure was used for the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. A three-stage ratio estimation procedure was used for the San Bernardino-Riverside-Ontario, CA, SMSA.

Prior to the implementation of the ratio estimation procedures, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

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

The following ratio estimation procedure was employed for all sample housing units from the permit-issuing universe. This factor was computed separately for all sample housing units within each permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

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

This ratio estimation procedure was introduced to correct the probabilities of selection for samples in each of the strata used in the sample selection of the permit-issuing universe. Prior to the AHS sample selection within each SMSA, housing units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. The same probability of selection was then applied to the remaining units to select the AHS sample. Since the number of housing units deleted from the AHS

universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the San Bernardino-Riverside-Ontario, CA, SMSA, the next ratio estimation procedure was employed for new construction mobile home units. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of mobile home units for each sector of the SMSA (central city and balance) by year built groups (1970-1974 and 1975-1982) to an independent estimate of mobile home units for the corresponding sector and year built group. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of mobile homes for the corresponding sector and year built group}}{\text{Sample estimate of mobile homes for the corresponding sector and year built group}}$$

The computed ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell.

The next ratio estimation procedure was applied in the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1982 housing inventory in each sector of the SMSA (central city and balance) to an independent estimate of the total housing units in each sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1982 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for the 12 SMSA's in 1982. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1982 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units using the existing sample weights were compared to the corresponding independent estimates for the central city, balance, and total for each of the 12 SMSA's, and the estimates which showed the most likely level of net growth since the 1980 census in both the central city and balance, as well as, the total SMSA were used in this ratio estimation procedure.

As a result of this analysis, the independent estimates were used in the Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco, CA, SMSA's. The sample estimates were used in the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's.

For the Columbus, OH and San Antonio, TX, SMSA's the independent estimates were used for the total SMSA only. The independent estimates were then proportioned between the central city and the balance of each SMSA according to the central city/balance distribution of the total housing units as given by the sample estimates. Due to the central city boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimates did not reflect the 1970 central city/balance SMSA definitions. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definitions were used.

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

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

1978-1982 lost housing units—The 1978-1982 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1978 housing inventory as was described in the 1978 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas*. Since the 1978-1982 lost housing units existed, by definition, in the 1978 housing inventory, there was a 1978 housing inventory weight associated with each 1978-1982 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1978-1982 lost housing units.

1978 estimation procedure—This report presents data on the housing characteristics of the 1978 housing inventory from the 1978 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1978.

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

RELIABILITY OF THE ESTIMATES

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

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

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing and the 1982 AHS-SMSA sample.

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

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

AHS-SMSA—The AHS-SMSA evaluation study was administered to a sample of units selected from the regular interviews containing recent movers who had previous residences within the SMSA. The purpose of the study was to evaluate the accuracy of responses regarding housing characteristics of recent mover's previous residence. The results of which are presented in the Census Bureau memorandum "Analysis of the Results from the AHS-SMSA Group AA-2 1982 Evaluation Study" (June 17, 1983).

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

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

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all housing units located inside these ED's would be represented in the sample. However, it has been estimated that the 1975 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. It should be noted that since these ED's were recanvassed for the 1978 and 1982 surveys, the number of missed housing units may be considerably less for 1982.

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

Rounding errors—For errors associated with processing, rounding of estimates introduces another source of error in the data, the severity of which depends on the statistics being measured. The effect of rounding is significant relative to the sampling error only for small percentages or small medians, when these figures are derived from relatively large bases (e.g., median number of persons per household). This means that confidence intervals formed from the standard errors given may be distorted, and this should be taken into account when considering the results of the survey.

Sampling errors for the AHS-SMSA sample—The particular sample used for this survey is one of a large number of possible samples of the same size that could have been selected using the same sample design. Even if the same questionnaires, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling

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

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

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

The figures presented in the tables that follow (page App-51) are approximations to the standard errors of various estimates shown in this report for this SMSA. In order to derive standard errors that would be applicable to a wide variety of items and also could be prepared at a moderate cost, a number of approximations were required. As a result, the tables of standard errors provide an indication of the order of magnitude of the standard errors rather than precise standard errors for any specific item. Standard errors applicable to estimates of characteristics of the 1978 housing inventory can be found in the AHS Series H-170 reports for 1978.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1982 housing inventory as well as estimates of characteristics of the 1978-1982 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 14,540 for the total SMSA, 6,710 for the central city of the SMSA, and 12,900 for the balance of the SMSA.

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

Table II (page App-51) presents the standard errors of estimated percentages for the 1982 housing inventory as well as estimated percentages of the 1978-1982 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

Included in tables I and II are estimates of standard errors for estimates of zero and zero percent. These estimates of standard errors are considered as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is obtained.

For ratios, 100 (x/y), where x is not a subclass of y, table II underestimates the standard error of the ratio when there is little or no correlation between x and y. For this type of ratio, a better approximation of the standard error may be obtained by letting the standard error of the ratio be approximately equal to:

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

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

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1982 there were 210,000 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 2,640. The following interpolation procedure was used.

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

Size of estimate	Standard error
200,000	2,690
210,000	x
232,900	2,520

The entry of "x" is determined as follows by vertically interpolating between 2,690 and 2,520.

$$210{,}000-200{,}000 = 10{,}000$$
$$232{,}900-200{,}000 = 32{,}900$$

$$2{,}690 + \frac{10{,}000}{32{,}900}\ (2{,}520-2{,}690) = 2{,}640$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 207,360 to 212,640 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1982 owner-occupied housing units lies within a range computed in this way would be correct for roughly

68 percent of all possible samples. Similarly, we could conclude that the average estimate derived from all possible samples, lies within the interval from 205,780 to 214,220 housing units with 90 percent confidence; and that the average estimate lies within the interval from 204,720 to 215,280 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 210,000 owner-occupied housing units, 37,900, or 18.0 percent, had two bedrooms. Interpolation using table II of this appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 18.0 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	18.0	25 or 75
200,000	0.6	a	0.9
210,000		p	
250,000	0.6	b	0.8

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

$$18.0-10.0 = 8.0$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{8.0}{15.0}(0.9-0.6) = 0.8$$

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

$$18.0-10.0 = 8.0$$
$$25.0-10.0 = 15.0$$

$$0.6 + \frac{8.0}{15.0}(0.8-0.6) = 0.7$$

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

$$210{,}000-200{,}000 = 10{,}000$$
$$250{,}000-200{,}000 = 50{,}000$$

$$0.8 + \frac{10{,}000}{50{,}000}(0.7-0.8) = 0.8$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 17.2 to 18.8 percent; the 90-percent confidence interval is from 16.7 to 19.3 percent; and the 95-percent confidence interval is from 16.4 to 19.6 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. If there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1978 and 1982 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1978 and 1982 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1982 there were 107,700 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 69,800. Table I shows the standard error of 37,900 is approximately 1,700, and the standard error of 107,700 is approximately 2,550. Therefore, the standard error of the estimated difference of 69,800 is about 3,060:

$$3{,}060 = \sqrt{(1{,}700)^2 + (2{,}550)^2}$$

Consequently, the 68-percent confidence interval for the 69,800 difference is from 66,740 to 72,860 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 confidence interval is from 64,900 to 74,700 housing units, and the 95-percent confidence interval is from 63,680 to 75,920 housing units. Thus, we can conclude with 95 percent confidence that the number of 1982 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied units with two bedrooms since the 95-percent confidence interval does not include zero or negative values.

Medians—For medians presented in certain tables, the sampling error depends on the size of the base and on the distribution upon which the median is based. An approximate method for measuring the reliability of the estimated median is to determine an interval about the estimated median so that there is a stated degree of confidence that the average median from all possible samples lies within the interval. The following procedure may be used to estimate confidence limits of a median based on sample data:

1. From table II, determine the standard error of a 50-percent characteristic on the base of the median.
2. Add to and subtract from 50 percent, the standard error determined in step 1.
3. Using the distribution of the characteristics, determine the confidence interval corresponding to the two points established in step 2. To find the lower endpoint of the confidence

interval, it is necessary to know into which interval of the distribution the lower percentage limit falls. Similarly, to find the upper endpoint to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.8. The base of the distribution from which this median was determined is 210,000 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 210,000 is approximately 1.1 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50-percent twice the standard error determined in step 1. This yields percentage limits of 47.8 and 52.2.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with three persons (for purposes of calculating the median, the category of three persons is considered to be from 2.5 to 3.5 persons) contains the 47.8 percent derived in step 2. About 90,600 housing units or 43.1 percent fall below this interval, and 41,000 housing units or 19.5 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\frac{(47.8-43.1)}{19.5} = 2.7$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.2 percent derived in step 2. About 90,600 housing units or 43.1 percent fall below this interval, and 41,000 housing units or 19.5 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\frac{(52.2-43.1)}{19.5} = 3.0$$

Thus, the 95-percent confidence interval ranges from 2.7 to 3.0 persons.

TABLE I. **Standard Errors for Estimated Number of Housing Units in the 1982 Housing Inventory and for Estimated Number of 1978-1982 Lost Units for the Rochester, NY, SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA**

(68 chances out of 100)

Size of estimate	Standard error[1]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	90	90	90	50,000	1,950	1,500	1,880
100	90	90	90	75,000	2,290	1,320	2,140
200	130	130	130	100,000	2,510	390	2,270
500	210	210	210	101,700	2,520	—	2,270
700	250	250	250	150,000	2,730	—	2,190
1,000	300	300	300	200,000	2,690	—	1,590
2,500	470	460	470	232,900	2,520	—	—
5,000	660	650	660	250,000	2,380	—	—
10,000	930	890	930	300,000	1,670	—	—
25,000	1,440	1,290	1,420	334,600	—	—	—

[1]For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.3. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero where the estimate for the total SMSA was derived from an independent estimate.

TABLE II. **Standard Errors for Estimated Percentages of Housing Units in the 1982 Housing Inventory and for Estimated Percentages of 1978-1982 Lost Housing Units for the Rochester, NY, SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA**

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
100	47.4	47.4	47.4	47.4	47.4	47.4	50,000	0.2	0.4	0.9	1.3	1.8	2.1
200	31.0	31.0	31.0	31.0	31.0	33.5	75,000	0.12	0.3	0.8	1.0	1.5	1.7
500	15.2	15.2	15.2	15.2	18.4	21.2	100,000	0.09	0.3	0.7	0.9	1.3	1.5
700	11.4	11.4	11.4	11.4	15.5	17.9	150,000	0.06	0.2	0.5	0.7	1.1	1.2
1,000	8.3	8.3	8.3	9.0	13.0	15.0	200,000	0.04	0.2	0.5	0.6	0.9	1.1
2,500	3.5	3.5	4.1	5.7	8.2	9.5	250,000	0.04	0.2	0.4	0.6	0.8	0.9
5,000	1.8	1.8	2.9	4.0	5.8	6.7	300,000	0.03	0.2	0.4	0.5	0.7	0.9
10,000	0.9	0.9	2.1	2.8	4.1	4.7	400,000	0.02	0.15	0.3	0.4	0.6	0.7
25,000	0.4	0.6	1.3	1.8	2.6	3.0							

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

San Antonio 1982

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1982 survey App-42
Selection of the 1978 AHS-SMSA sample App-43
Sample selection for the 1978 Coverage Improvement Program . App-44
Coverage Improvement for deficiency 1 ..... App-44
Coverage Improvement for deficiency 2 ..... App-44
Coverage Improvement for deficiencies 3-6 ..... App-44
1978-1982 additions to the housing inventory App-45
1970 Census of Population and Housing App-45
ESTIMATION App-45
1982 housing inventory App-45
1978-1982 lost housing units ... App-46
1978 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-47
Rounding errors App-47
Sampling errors for the AHS-SMSA sample App-47
Illustration of the use of the standard error tables App-48
Differences App-49
Illustration of the computation of the standard error of a difference App-49
Medians App-49
Illustration of the computation of the 95-percent confidence interval of a median ... App-50
Standard error tables App-51

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 12 SMSA's in this report series (H-170-82) are based on data collected from the 1982 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 12 SMSA's selected for interview during 1982 were interviewed previously in 1975 and 1978 (see the list of SMSA reports from the AHS in the introduction of this report).

Three of the larger SMSA's (Atlanta, GA; Philadelphia, PA-NJ; and San Francisco-Oakland, CA) in the 1982 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1975 and 1978 surveys, evenly divided between the central city and the balance of the respective SMSA. The 9 remaining SMSA's (Cincinnati, OH-KY-IN; Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Paterson-Clifton-Passaic, NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; and San Diego, CA) in the 1982 group were represented by a sample of 5,000 designated housing units in the 1975 and 1978 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1975 and 1978 surveys, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panels 1 through 3 were dropped from all SMSA's for the 1982 survey, excluding the central cities of the Paterson-Clifton-Passaic, NJ, SMSA. Due to additional budget limitations for the 1982 survey, it also became necessary to drop part of the sample of housing units from panels 4 through 12 in the Atlanta, GA; New Orleans, LA; Philadelphia, PA-NJ; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. The effect of this reduction was to maintain a sample size of 4,250 in all SMSA's. The interviewing was done during the period April 1982 through December 1982 with one-ninth of the sample housing units being visited each month.

In this SMSA, 3,927 housing units were eligible for interview. Of these sample housing 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 the 3,927 housing units eligible for interview, 358 units were visited but were not eligible for interview because they were condemned, unfit, demolished, converted to group quarters use, etc.

Designation of sample housing units for the 1982 survey—The sample housing units designated to be interviewed in the 1982 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 1978 survey and remained in sample after the 1982 panel reduction.
2. All sample housing units that were type A noninterviews (i.e., units eligible to be interviewed) or type B noninterviews (i.e., units not eligible for interview at the time of the survey but which could become eligible in the future) in the 1978 survey and remained in sample after the 1982 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1982 AHS questionnaire, page App-18.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented the housing units built in permit-issuing areas since the 1978 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1978 survey.)
5. All sample housing units that were selected as part of the 1978 Coverage Improvement Program. (This sample represented most of the housing units which, until 1978, did not have a chance of selection.)

Selection of the 1978 AHS-SMSA sample—The sample for the SMSA's which, in 1970, were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). In 1970, the following four SMSA's were 100-percent permit-issuing: Paterson-Clifton-Passaic, NJ; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA. The remaining eight SMSA's contain a sample from the nonpermit universe.

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

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA during the 1970 Census of Population and Housing. This file contained records for occupied housing units, vacant housing units, and housing units in certain special places or group quarters. Sampling operations were done separately for the special place and group quarters records, and for the occupied and vacant housing unit records. Before the sample was selected from the occupied and vacant housing unit records, the occupied housing unit records were stratified by race of head (non-Black/Black) and the vacant records were stratified into four categories pertaining to the value or rent associated with the vacant housing units. The occupied housing unit records were further stratified so that each unit was assigned to 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household income	Tenure	
	Owner—Family size 1 2 3 4 5+	Renter—Family size 1 2 3 4 5+
Under $3,000		
$3,000 to $5,999 . . .		
$6,000 to $9,999 . . .		
$10,000 to $14,999 .		
$15,000 and over . . .		

Thus, for this SMSA, the occupied housing unit records from the permit-issuing universe were assigned to 1 of 100 strata for either the central city or for the balance, and the vacant housing unit records were assigned to 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. However, whenever a record was selected to be in sample, the housing unit record adjacent to it on the file was also selected to be in sample, thereby insuring the necessary designated sample size.

Before the sample was selected from the group quarters and special place records, the records were stratified by census tract and census enumeration district (ED) within the central city and within the balance of the SMSA. A sample of special place records was then selected by a procedure that produced one-quarter of the desired sample size. However, at the time of the survey, the housing units at each of the special places were listed and subsampled at a rate which produced an expected four sample units, thereby insuring the necessary designated sample size.

The second frame from which this SMSA sample was selected was a list of new construction building permits issued since 1970 (i.e., the new construction universe). The sample selection from the list of new construction building permits was an independent operation within this SMSA. Prior to sample selection, the

list of permits was chronologically stratified by the date the permits were issued, and clusters of an expected four (usually adjacent) housing units were formed. These clusters were then sampled for inclusion at the overall sampling rate.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS sample was selected from a frame consisting of areas not under the jurisdiction of permit-issuing offices (i.e., the nonpermit universe). The first step in the sampling operation for the nonpermit universe was the selection (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

The sample ED's were then divided into segments; i.e., small land areas with well-defined boundaries having an expected size of four, or a multiple of four, housing units. At the time of the survey, those segments that did not have an expected size of four were further subdivided to produce an expected four sample housing units.

The next step was the selection of one of these segments within each sample ED. All housing units in existence at the time of interview in these selected segments were eligible for sample. Thus, housing units enumerated in the 1970 census as well as housing units built since the 1970 census are included.

Sample Selection for the 1978 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

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

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 SMSA. 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. 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 Rochester, NY, SMSA.

In the remaining 11 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. These procedures added an estimated 1,688 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 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 251 units to the coverage of the housing inventory of this SMSA.

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

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

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample unit were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and units within these structures were interviewed.

In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. This procedure added an estimated 1,248 units to the coverage of the housing inventory of this SMSA.

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

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure did not add any units to the coverage of the housing inventory of this SMSA.

1978-1982 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1978 survey, was selected to represent housing units built in permit-issuing areas since the 1978 survey. Sampling procedures were identical to those used in selecting the 1970-1978 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1977, to identify any housing units missed in the 1978 survey or any housing units added since the 1978 survey.

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 1982 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 1982 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1978 (i.e., 1978-1982 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1982 housing inventory—The AHS estimates of characteristics of the 1982 housing inventory were produced using a one-stage ratio estimation procedure for the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's. A two-stage ratio estimation procedure was used for the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. A three-stage ratio estimation procedure was used for the San Bernardino-Riverside-Ontario, CA, SMSA.

Prior to the implementation of the ratio estimation procedures, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

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

The following ratio estimation procedure was employed for all sample housing units from the permit-issuing universe. This factor was computed separately for all sample housing units within each permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

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

This ratio estimation procedure was introduced to correct the probabilities of selection for samples in each of the strata used in the sample selection of the permit-issuing universe. Prior to the AHS sample selection within each SMSA, housing units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. The same probability of selection was then applied to the remaining units to select the AHS sample. Since the number of housing units deleted from the AHS

universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the San Bernardino-Riverside-Ontario, CA, SMSA, the next ratio estimation procedure was employed for new construction mobile home units. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of mobile home units for each sector of the SMSA (central city and balance) by year built groups (1970-1974 and 1975-1982) to an independent estimate of mobile home units for the corresponding sector and year built group. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of mobile homes for the corresponding sector and year built group}}{\text{Sample estimate of mobile homes for the corresponding sector and year built group}}$$

The computed ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell.

The next ratio estimation procedure was applied in the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1982 housing inventory in each sector of the SMSA (central city and balance) to an independent estimate of the total housing units in each sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1982 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for the 12 SMSA's in 1982. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1982 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units using the existing sample weights were compared to the corresponding independent estimates for the central city, balance, and total for each of the 12 SMSA's, and the estimates which showed the most likely level of net growth since the 1980 census in both the central city and balance, as well as, the total SMSA were used in this ratio estimation procedure.

As a result of this analysis, the independent estimates were used in the Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco, CA, SMSA's. The sample estimates were used in the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's.

For the Columbus, OH and San Antonio, TX, SMSA's the independent estimates were used for the total SMSA only. The independent estimates were then proportioned between the central city and the balance of each SMSA according to the central city/balance distribution of the total housing units as given by the sample estimates. Due to the central city boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimates did not reflect the 1970 central city/balance SMSA definitions. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definitions were used.

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

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

1978-1982 lost housing units—The 1978-1982 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1978 housing inventory as was described in the 1978 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1978-1982 lost housing units existed, by definition, in the 1978 housing inventory, there was a 1978 housing inventory weight associated with each 1978-1982 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1978-1982 lost housing units.

1978 estimation procedure—This report presents data on the housing characteristics of the 1978 housing inventory from the 1978 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1978.

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

RELIABILITY OF THE ESTIMATES

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

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

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing and the 1982 AHS-SMSA sample.

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

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

AHS-SMSA—The AHS-SMSA evaluation study was administered to a sample of units selected from the regular interviews containing recent movers who had previous residences within the SMSA. The purpose of the study was to evaluate the accuracy of responses regarding housing characteristics of recent mover's previous residence. The results of which are presented in the Census Bureau memorandum "Analysis of the Results from the AHS-SMSA Group AA-2 1982 Evaluation Study." (June 17, 1983)

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

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

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all housing units located inside these ED's would be represented in the sample. However, it has been estimated that the 1975 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. It should be noted that since these ED's were recanvassed for the 1978 and 1982 surveys, the number of missed housing units may be considerably less for 1982.

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

Rounding errors—For errors associated with processing, rounding of estimates introduces another source of error in the data, the severity of which depends on the statistics being measured. The effect of rounding is significant relative to the sampling error only for small percentages or small medians, when these figures are derived from relatively large bases (e.g., median number of persons per household). This means that confidence intervals formed from the standard errors given may be distorted, and this should be taken into account when considering the results of the survey.

Sampling errors for the AHS-SMSA sample—The particular sample used for this survey is one of a large number of possible samples of the same size that could have been selected using the same sample design. Even if the same questionnaires, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling

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

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

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

The figures presented in the tables that follow (page App-51) are approximations to the standard errors of various estimates shown in this report for this SMSA. In order to derive standard errors that would be applicable to a wide variety of items and also could be prepared at a moderate cost, a number of approximations were required. As a result, the tables of standard errors provide an indication of the order of magnitude of the standard errors rather than precise standard errors for any specific item. Standard errors applicable to estimates of characteristics of the 1978 housing inventory can be found in the AHS Series H-170 reports for 1978.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1982 housing inventory as well as estimates of characteristics of the 1978-1982 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 39,120 for the total SMSA, 13,670 for the central city of the SMSA, and 36,660 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, particularly if the percentages are 50 percent or more.

Table II (page App-51) presents the standard errors of estimated percentages for the 1982 housing inventory as well as estimated percentages of the 1978-1982 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

Included in tables I and II are estimates of standard errors for estimates of zero and zero percent. These estimates of standard errors are considered as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is obtained.

For ratios, 100 (x/y), where x is not a subclass of y, table II underestimates the standard error of the ratio when there is little or no correlation between x and y. For this type of ratio, a better approximation of the standard error may be obtained by letting the standard error of the ratio be approximately equal to:

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

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

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1982 there were 218,200 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 3,160. 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	3,220
218,200	x
250,000	3,060

The entry of "x" is determined as follows by vertically interpolating between 3,220 and 3,060.

$$218{,}200 - 200{,}000 = 18{,}200$$
$$250{,}000 - 200{,}000 = 50{,}000$$

$$3{,}220 + \frac{18{,}200}{50{,}000}(3{,}060 - 3{,}220) = 3{,}160$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 215,040 to 221,360 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1982 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 213,140 to 223,260 housing units with 90 percent confidence; and that the average estimate lies within the interval from 211,880 to 224,520 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 218,200 owner-occupied housing units, 61,500, or 28.2 percent, had two bedrooms. Interpolation using table II of this appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 28.2 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	28.2	50
200,000	1.0	a	1.2
218,200		p	
250,000	0.9	b	1.0

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

$$28.2-25.0 = 3.2$$
$$50.0-25.0 = 25.0$$

$$1.0 + \frac{3.2}{25.0}(1.2-1.0) = 1.0$$

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

$$28.2-25.0 = 3.2$$
$$50.0-25.0 = 25.0$$

$$0.9 + \frac{3.2}{25.0}(1.0-0.9) = 0.9$$

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

$$218{,}200-200{,}000 = 18{,}200$$
$$250{,}000-200{,}000 = 50{,}000$$

$$1.0 + \frac{18{,}200}{50{,}000}(0.9-1.0) = 1.0$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 27.2 to 29.2 percent; the 90-percent confidence interval is from 26.6 to 29.8 percent; and the 95-percent confidence interval is from 26.2 to 30.2 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. If there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1978 and 1982 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1978 and 1982 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1982 there were 117,800 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 56,300. Table I shows the standard error of 61,500 is approximately 2,360 and the standard error of 117,800 is approximately 2,980. Therefore, the standard error of the estimated difference of 56,300 is about 3,800.

$$3{,}800 = \sqrt{(2{,}360)^2 + (2{,}980)^2}$$

Consequently, the 68-percent confidence interval for the 56,300 difference is from 52,500 to 60,100 housing units. Therefore, a conclusion that the average estimate derived from all possible samples, of this difference, lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 50,220 to 62,380 housing units, and the 95-percent confidence interval is from 48,700 to 63,900 housing units. Thus, we can conclude with 95 percent confidence that the number of 1982 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied units with two bedrooms since the 95-percent confidence interval does not include zero or negative values.

Medians—For medians presented in certain tables, the sampling error depends on the size of the base and on the distribution upon which the median is based. An approximate method for measuring the reliability of the estimated median is to determine an interval about the estimated median so that there is a stated degree of confidence that the average median from all possible samples lies within the interval. The following procedure may be used to estimate confidence limits of a median based on sample data:

1. From table II, determine the standard error of a 50-percent characteristic on the base of the median.
2. Add to and subtract from 50 percent, the standard error determined in step 1.
3. Using the distribution of the characteristics, determine the confidence interval corresponding to the two points established in step 2. To find the lower endpoint of the confidence

interval, it is necessary to know into which interval of the distribution the lower percentage limit falls. Similarly, to find the upper endpoint to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.8. The base of the distribution from which this median was determined is 218,200 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 218,200 is approximately 1.1 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50-percent twice the standard error determined in step 1. This yields percentage limits of 47.8 and 52.2.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with three persons (for purposes of calculating the median, the category of three persons is considered to be from 2.5 to 3.5 persons) contains the 47.8 percent derived in step 2. About 97,500 housing units or 44.7 percent fall below this interval, and 41,400 housing units or 19.0 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(47.8-44.7)}{19.0} = 2.7$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.2 percent derived in step 2. About 97,500 housing units or 44.7 percent fall below this interval, and 41,400 housing units or 19.0 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(52.2-44.7)}{19.0} = 2.9$$

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

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1982 Housing Inventory and for Estimated Number of 1978-1982 Lost Units for the San Antonio, TX, SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	110	100	110	50,000	2,180	1,990	1,780
100	110	100	110	75,000	2,570	2,290	1,790
200	150	140	140	100,000	2,840	2,450	1,470
500	230	220	230	126,100	3,040	2,500	—
700	280	260	270	150,000	3,150	2,460	—
1,000	330	310	320	200,000	3,220	2,040	—
2,500	520	490	510	250,000	3,060	590	—
5,000	740	700	710	253,600	3,040	—	—
10,000	1,030	970	980	300,000	2,620	—	—
25,000	1,600	1,490	1,450	379,600	—	—	—

[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 and the central city and 1.4 for the balance (not in central city) estimates. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero where the estimate for the total SMSA was derived from an independent estimate.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1982 Housing Inventory and for Estimated Percentages of 1978-1982 Lost Housing Units for the San Antonio, TX, SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
200	35.4	35.4	35.4	35.4	35.4	37.0	50,000	0.2	0.5	1.0	1.4	2.0	2.3
500	18.0	18.0	18.0	18.0	20.3	23.4	75,000	0.15	0.4	0.8	1.1	1.7	1.9
700	13.5	13.5	13.5	13.5	17.1	19.8	100,000	0.11	0.3	0.7	1.0	1.4	1.7
1,000	9.9	9.9	9.9	9.9	14.3	16.5	150,000	0.07	0.3	0.6	0.8	1.2	1.4
2,500	4.2	4.2	4.6	6.3	9.1	10.5	200,000	0.05	0.2	0.5	0.7	1.0	1.2
5,000	2.1	2.1	3.2	4.4	6.4	7.4	250,000	0.04	0.2	0.5	0.6	0.9	1.0
10,000	1.1	1.1	2.3	3.1	4.5	5.2	300,000	0.04	0.2	0.4	0.6	0.8	1.0
25,000	0.4	0.7	1.4	2.0	2.9	3.3	400,000	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 1.3 for the total SMSA, 1.2 for the central city, and 1.4 for the balance (not in central city) estimates.

San Diego 1982

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1982 survey App-42
Selection of the 1978 AHS-SMSA sample App-43
Sample selection for the 1978 Coverage Improvement Program . App-44
Coverage Improvement for deficiency 1 ..... App-44
Coverage Improvement for deficiency 2 ..... App-44
Coverage improvement for deficiencies 3-6 ..... App-44
1978-1982 additions to the housing inventory App-45
1970 Census of Population and Housing App-45
ESTIMATION App-45
1982 housing inventory App-45
1978-1982 lost housing units ... App-46
1978 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-47
Rounding errors App-47
Sampling errors for the AHS-SMSA sample App-47
Illustration of the use of the standard error tables App-48
Differences App-49
Illustration of the computation of the standard error of a difference App-49
Medians App-49
Illustration of the computation of the 95-percent confidence interval of a median ... App-50
Standard error tables App-51

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 12 SMSA's in this report series (H-170-82) are based on data collected from the 1982 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 12 SMSA's selected for interview during 1982 were interviewed previously in 1975 and 1978 (see the list of SMSA reports from the AHS in the introduction of this report).

Three of the larger SMSA's (Atlanta, GA; Philadelphia, PA-NJ; and San Francisco-Oakland, CA) in the 1982 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1975 and 1978 surveys, evenly divided between the central city and the balance of the respective SMSA. The 9 remaining SMSA's (Cincinnati, OH-KY-IN; Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Paterson-Clifton-Passaic, NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; and San Diego, CA) in the 1982 group were represented by a sample of 5,000 designated housing units in the 1975 and 1978 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1975 and 1978 surveys, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panels 1 through 3 were dropped from all SMSA's for the 1982 survey, excluding the central cities of the Paterson-Clifton-Passaic, NJ, SMSA. Due to additional budget limitations for the 1982 survey, it also became necessary to drop part of the sample of housing units from panels 4 through 12 in the Atlanta, GA; New Orleans, LA; Philadelphia, PA-NJ; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. The effect of this reduction was to maintain a sample size of 4,250 in all SMSA's. The interviewing was done during the period April 1982 through December 1982 with one-ninth of the sample housing units being visited each month.

In this SMSA, 4,003 housing units were eligible for interview. Of these sample housing units, 125 interviews were not obtained because, for occupied sample units, the occupants refused to be interviewed, were not at home after repeated visits, or were unavailable for some other reason; or, for vacant units, no informed respondent could be found after repeated visits. In addition to the 4,003 housing units eligible for interview, 212 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 1982 survey—The sample housing units designated to be interviewed in the 1982 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 1978 survey and remained in sample after the 1982 panel reduction.
2. All sample housing units that were type A noninterviews (i.e., units eligible to be interviewed) or type B noninterviews (i.e., units not eligible for interview at the time of the survey but which could become eligible in the future) in the 1978 survey and remained in sample after the 1982 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1982 AHS questionnaire, page App-18.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented the housing units built in permit-issuing areas since the 1978 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1978 survey.)
5. All sample housing units that were selected as part of the 1978 Coverage Improvement Program. (This sample represented most of the housing units which, until 1978, did not have a chance of selection.)

Selection of the 1978 AHS-SMSA sample—The sample for the SMSA's which, in 1970, were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). In 1970, the following four SMSA's were 100-percent permit-issuing: Paterson-Clifton-Passaic, NJ; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA. The remaining eight SMSA's contain a sample from the nonpermit universe.

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

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA during the 1970 Census of Population and Housing. This file contained records for occupied housing units, vacant housing units, and housing units in certain special places or group quarters. Sampling operations were done separately for the special place and group quarters records, and for the occupied and vacant housing unit records. Before the sample was selected from the occupied and vacant housing unit records, the occupied housing unit records were stratified by race of head (non-Black/Black) and the vacant records were stratified into four categories pertaining to the value or rent associated with the vacant housing units. The occupied housing unit records were further stratified so that each unit was assigned to 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household income	Tenure	
	Owner—Family size	Renter—Family size
	1 2 3 4 5+	1 2 3 4 5+
Under $3,000		
$3,000 to $5,999 . . .		
$6,000 to $9,999 . . .		
$10,000 to $14,999 .		
$15,000 and over . . .		

Thus, for this SMSA, the occupied housing unit records from the permit-issuing universe were assigned to 1 of 100 strata for either the central city or for the balance, and the vacant housing unit records were assigned to 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. However, whenever a record was selected to be in sample, the housing unit record adjacent to it on the file was also selected to be in sample, thereby insuring the necessary designated sample size.

Before the sample was selected from the group quarters and special place records, the records were stratified by census tract and census enumeration district (ED) within the central city and within the balance of the SMSA. A sample of special place records was then selected by a procedure that produced one-quarter of the desired sample size. However, at the time of the survey, the housing units at each of the special places were listed and subsampled at a rate which produced an expected four sample units, thereby insuring the necessary designated sample size.

The second frame from which this SMSA sample was selected was a list of new construction building permits issued since 1970 (i.e., the new construction universe). The sample selection from the list of new construction building permits was an independent operation within this SMSA. Prior to sample selection, the

list of permits was chronologically stratified by the date the permits were issued, and clusters of an expected four (usually adjacent) housing units were formed. These clusters were then sampled for inclusion at the overall sampling rate.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS sample was selected from a frame consisting of areas not under the jurisdiction of permit-issuing offices (i.e., the nonpermit universe). The first step in the sampling operation for the nonpermit universe was the selection (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

The sample ED's were then divided into segments; i.e., small land areas with well-defined boundaries having an expected size of four, or a multiple of four, housing units. At the time of the survey, those segments that did not have an expected size of four were further subdivided to produce an expected four sample housing units.

The next step was the selection of one of these segments within each sample ED. All housing units in existence at the time of interview in these selected segments were eligible for sample. Thus, housing units enumerated in the 1970 census as well as housing units built since the 1970 census are included.

Sample Selection for the 1978 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

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

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 SMSA. 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. 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 Rochester, NY, SMSA.

In the remaining 11 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. These procedures added an estimated 10,246 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 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 11,358 units to the coverage of the housing inventory of this SMSA.

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

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

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample unit were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and units within these structures were interviewed.

In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. This procedure added an estimated 12,508 units to the coverage of the housing inventory of this SMSA.

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

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 363 units to the coverage of the housing inventory of this SMSA.

1978-1982 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1978 survey, was selected to represent housing units built in permit-issuing areas since the 1978 survey. Sampling procedures were identical to those used in selecting the 1970-1978 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1977, to identify any housing units missed in the 1978 survey or any housing units added since the 1978 survey.

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 1982 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 1982 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1978 (i.e., 1978-1982 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1982 housing inventory—The AHS estimates of characteristics of the 1982 housing inventory were produced using a one-stage ratio estimation procedure for the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's. A two-stage ratio estimation procedure was used for the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. A three-stage ratio estimation procedure was used for the San Bernardino-Riverside-Ontario, CA, SMSA.

Prior to the implementation of the ratio estimation procedures, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

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

The following ratio estimation procedure was employed for all sample housing units from the permit-issuing universe. This factor was computed separately for all sample housing units within each permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

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

This ratio estimation procedure was introduced to correct the probabilities of selection for samples in each of the strata used in the sample selection of the permit-issuing universe. Prior to the AHS sample selection within each SMSA, housing units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. The same probability of selection was then applied to the remaining units to select the AHS sample. Since the number of housing units deleted from the AHS

universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the San Bernardino-Riverside-Ontario, CA, SMSA, the next ratio estimation procedure was employed for new construction mobile home units. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of mobile home units for each sector of the SMSA (central city and balance) by year built groups (1970-1974 and 1975-1982) to an independent estimate of mobile home units for the corresponding sector and year built group. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of mobile homes for the corresponding sector and year built group}}{\text{Sample estimate of mobile homes for the corresponding sector and year built group}}$$

The computed ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell.

The next ratio estimation procedure was applied in the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1982 housing inventory in each sector of the SMSA (central city and balance) to an independent estimate of the total housing units in each sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1982 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for the 12 SMSA's in 1982. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1982 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units using the existing sample weights were compared to the corresponding independent estimates for the central city, balance, and total for each of the 12 SMSA's, and the estimates which showed the most likely level of net growth since the 1980 census in both the central city and balance, as well as, the total SMSA were used in this ratio estimation procedure.

As a result of this analysis, the independent estimates were used in the Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco, CA, SMSA's. The sample estimates were used in the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's.

For the Columbus, OH and San Antonio, TX, SMSA's the independent estimates were used for the total SMSA only. The independent estimates were then proportioned between the central city and the balance of each SMSA according to the central city/balance distribution of the total housing units as given by the sample estimates. Due to the central city boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimates did not reflect the 1970 central city/balance SMSA definitions. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definitions were used.

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

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

1978-1982 lost housing units—The 1978-1982 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1978 housing inventory as was described in the 1978 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas*. Since the 1978-1982 lost housing units existed, by definition, in the 1978 housing inventory, there was a 1978 housing inventory weight associated with each 1978-1982 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1978-1982 lost housing units.

1978 estimation procedure—This report presents data on the housing characteristics of the 1978 housing inventory from the 1978 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1978.

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

RELIABILITY OF THE ESTIMATES

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

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

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing and the 1982 AHS-SMSA sample.

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

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

AHS-SMSA—The AHS-SMSA evaluation study was administered to a sample of units selected from the regular interviews containing recent movers who had previous residences within the SMSA. The purpose of the study was to evaluate the accuracy of responses regarding housing characteristics of recent mover's previous residence. The results of which are presented in the Census Bureau memorandum "Analysis of the Results from the AHS-SMSA Group AA-2 1982 Evaluation Study" (June 17, 1983).

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

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

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all housing units located inside these ED's would be represented in the sample. However, it has been estimated that the 1975 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. It should be noted that since these ED's were recanvassed for the 1978 and 1982 surveys, the number of missed housing units may be considerably less for 1982.

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

Rounding errors—For errors associated with processing, rounding of estimates introduces another source of error in the data, the severity of which depends on the statistics being measured. The effect of rounding is significant relative to the sampling error only for small percentages or small medians, when these figures are derived from relatively large bases (e.g., median number of persons per household). This means that confidence intervals formed from the standard errors given may be distorted, and this should be taken into account when considering the results of the survey.

Sampling errors for the AHS-SMSA sample—The particular sample used for this survey is one of a large number of possible samples of the same size that could have been selected using the same sample design. Even if the same questionnaires, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling

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

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

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

The figures presented in the tables that follow (page App-51) are approximations to the standard errors of various estimates shown in this report for this SMSA. In order to derive standard errors that would be applicable to a wide variety of items and also could be prepared at a moderate cost, a number of approximations were required. As a result, the tables of standard errors provide an indication of the order of magnitude of the standard errors rather than precise standard errors for any specific item. Standard errors applicable to estimates of characteristics of the 1978 housing inventory can be found in the AHS Series H-170 reports for 1978.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1982 housing inventory as well as estimates of characteristics of the 1978-1982 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 35,040 for the total SMSA, 26,250 for the central city of the SMSA, and 23,210 for the balance of the SMSA.

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

Table II (page App-51) presents the standard errors of estimated percentages for the 1982 housing inventory as well as estimated percentages of the 1978-1982 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in table II.

Included in tables I and II are estimates of standard errors for estimates of zero and zero percent. These estimates of standard errors are considered as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is obtained.

For ratios, 100 (x/y), where x is not a subclass of y, table II underestimates the standard error of the ratio when there is little or no correlation between x and y. For this type of ratio, a better approximation of the standard error may be obtained by letting the standard error of the ratio be approximately equal to:

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

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

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1982 there were 369,700 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 6,220. 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
353,700	6,220
369,700	x
389,600	6,220

The entry of "x" is determined as follows by vertically interpolating between 6,220 and 6,220.

$$369{,}700 - 353{,}700 = 16{,}000$$
$$389{,}600 - 353{,}700 = 35{,}900$$

$$6{,}220 + \frac{16{,}000}{35{,}900}\ (6{,}220 - 6{,}220) = 6{,}220$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 363,480 to 375,920 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1982 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 359,750 to 379,650 housing units with 90 percent confidence; and that the average estimate lies within the interval from 357,260 to 382,140 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 369,700 owner-occupied housing units, 110,300, or 29.8 percent, had two bedrooms. Interpolation using table II of this appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 29.8 percent is approximately 1.1 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	29.8	50
300,000	1.2	a	1.3
369,700		p	
400,000	1.0	b	1.2

1. The entry for cell "a" is determined by horizontal interpolation between 1.2 and 1.3.

$$29.8-25.0 = 4.8$$
$$50.0-25.0 = 25.0$$

$$1.2 + \frac{4.8}{25.0}\ (1.3-1.2) = 1.2$$

2. The entry for cell "b" is determined by horizontal interpolation between 1.0 and 1.2.

$$29.8-25.0 = 4.8$$
$$50.0-25.0 = 25.0$$

$$1.0 + \frac{4.8}{25.0}\ (1.2-1.0) = 1.0$$

3. The entry for "p" is then determined by vertical interpolation between 1.2 and 1.0.

$$369{,}700-300{,}000 = 69{,}700$$
$$400{,}000-300{,}000 = 100{,}000$$

$$1.2 + \frac{69{,}700}{100{,}000}\ (1.0-1.2) = 1.1$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 28.7 to 30.9 percent; the 90-percent confidence interval is from 28.0 to 31.6 percent; and the 95-percent confidence interval is from 27.6 to 32.0 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. If there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1978 and 1982 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1978 and 1982 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1982 there were 160,200 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 49,900. Table I shows the standard error of 110,300 is approximately 4,400 and the standard error of 160,200 is approximately 5,100. Therefore, the standard error of the estimated difference of 49,900 is about 6,740.

$$6{,}740 = \sqrt{(4{,}400)^2 + (5{,}100)^2}$$

Consequently, the 68-percent confidence interval for the 49,900 difference is from 43,160 to 56,640 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 39,120 to 60,680 housing units, and the 95-percent confidence interval is from 36,420 to 63,380 housing units. Thus, we can conclude with 95 percent confidence that the number of 1982 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied units with two bedrooms since the 95-percent confidence interval does not include zero or negative values.

Medians—For medians presented in certain tables, the sampling error depends on the size of the base and on the distribution upon which the median is based. An approximate method for measuring the reliability of the estimated median is to determine an interval about the estimated median so that there is a stated degree of confidence that the average median from all possible samples lies within the interval. The following procedure may be used to estimate confidence limits of a medi : based on sample data:

1. From table II, determine the standard error of a 50-percent characteristic on the base of the median.
2. Add to and subtract from 50 percent, the standard error determined in step 1.
3. Using the distribution of the characteristics, determine the confidence interval corresponding to the two points established in step 2. To find the lower endpoint of the confidence

interval, it is necessary to know into which interval of the distribution the lower percentage limit falls. Similarly, to find the upper endpoint to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median— Table A-1 of part A of this report shows the median number of persons for,owner-occupied housing units is 2.4. The base of the distribution from which this median was determined is 369,700 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 369,700 is approximately 1.2 percentage points.
2. To obtain a 95-percent confidence interval on the estimated median, initially add to and subtract from 50-percent twice the standard error determined in step 1. This yields percentage limits of 47.6 and 52.4.
3. From the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) contains the 47.6 percent derived in step 2. About 61,900 housing units or 16.7 percent fall below this interval, and 138,500 housing units or 37.5 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$1.5 + (2.5-1.5)\frac{(47.6-16.7)}{37.5} = 2.3$$

Similarly, the interval for owner-occupied housing units with two persons contains the 52.4 percent derived in step 2. About 61,900 housing units or 16.7 percent fall below this interval, and 138,500 housing units or 37.5 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$1.5 + (2.5-1.5)\frac{(52.4-16.7)}{37.5} = 2.5$$

Thus, the 95-percent confidence interval ranges from 2.3 to 2.5 persons.

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1982 Housing Inventory and for Estimated Number of 1978-1982 Lost Units for the San Diego, CA, SMSA, for the Central City of the SMSA and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]			Size of estimate	Standard error[1]		
	SMSA	In central city	Not in central city		SMSA	In central city	Not in central city
0	210	210	220	100,000	4,250	3,850	4,030
100	210	210	220	150,000	4,990	4,220	4,490
200	210	210	220	200,000	5,520	4,230	4,610
500	320	320	330	250,000	5,880	3,890	4,420
700	380	380	390	300,000	6,110	3,070	3,880
1,000	460	450	470	353,700	6,220	—	2,670
2,500	720	720	740	389,600	6,220	—	—
5,000	1,020	1,010	1,040	400,000	6,200	—	—
10,000	1,430	1,420	1,460	500,000	5,840	—	—
25,000	2,240	2,190	2,260	600,000	4,910	—	—
50,000	3,120	2,980	3,090	700,000	2,920	—	—
75,000	3,750	3,490	3,640	743,300	—	—	—

[1]For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.3 for the total SMSA, 1.2 for the central city, and 1.1 for the balance (not in central city) estimates. Standard errors of estimates pertaining to total housing units for the central city, balance, and total SMSA are assumed to be equal to zero where the estimate for the total SMSA was derived from an independent estimate.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1982 Housing Inventory and for Estimated Percentages of 1978-1982 Lost Housing Units for the San Diego, CA, SMSA, for the Central City and for the Balance (Not in Central City) of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
300	42.2	42.2	42.2	42.2	42.2	42.7	100,000	0.2	0.5	1.0	1.4	2.0	2.3
500	30.4	30.4	30.4	30.4	30.4	33.1	150,000	0.15	0.4	0.8	1.1	1.7	1.9
700	23.8	23.8	23.8	23.8	24.2	27.9	200,000	0.11	0.3	0.7	1.0	1.4	1.7
1,000	17.9	17.9	17.9	17.9	20.2	23.4	250,000	0.09	0.3	0.6	0.9	1.3	1.5
2,500	8.0	8.0	8.0	8.9	12.8	14.8	300,000	0.07	0.3	0.6	0.8	1.2	1.3
5,000	4.2	4.2	4.6	6.3	9.1	10.5	400,000	0.05	0.2	0.5	0.7	1.0	1.2
10,000	2.1	2.1	3.2	4.4	6.4	7.4	500,000	0.04	0.2	0.5	0.6	0.9	1.0
25,000	0.9	0.9	2.0	2.8	4.0	4.7	600,000	0.04	0.2	0.4	0.6	0.8	1.0
50,000	0.4	0.7	1.4	2.0	2.9	3.3	700,000	0.03	0.2	0.4	0.5	0.8	0.9
75,000	0.3	0.5	1.2	1.6	2.3	2.7	800,000	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 and the central city and 1.1 for the balance (not in central city) estimates.

San Francisco-Oakland 1982

Appendix B

Source and Reliability of the Estimates

SAMPLE DESIGN App-42
Annual Housing Survey App-42
Designation of sample housing units for the 1982 survey App-42
Selection of the 1978 AHS-SMSA sample App-43
Sample selection for the 1978 Coverage Improvement Program . App-44
Coverage Improvement for deficiency 1 ..... App-44
Coverage Improvement for deficiency 2 ..... App-44
Coverage Improvement for deficiencies 3-6 ..... App-45
1978-1982 additions to the housing inventory App-45
1970 Census of Population and Housing App-45
ESTIMATION App-45
1982 housing inventory App-45
1978-1982 lost housing units ... App-46
1978 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-47
Rounding errors App-47
Sampling errors for the AHS-SMSA sample App-48
Illustration of the use of the standard error tables App-49
Differences App-49
Illustration of the computation of the standard error of a difference App-50
Medians App-50
Illustration of the computation of the 95-percent confidence interval of a median ... App-50
Standard error tables App-51

SAMPLE DESIGN

Annual Housing Survey—The estimates for each of the 12 SMSA's in this report series (H-170-82) are based on data collected from the 1982 Annual Housing Survey (AHS) which was conducted by the Bureau of the Census acting as collection agent for the Department of Housing and Urban Development.

The SMSA's selected for the AHS are interviewed on a rotating basis. The group of 12 SMSA's selected for interview during 1982 were interviewed previously in 1975 and 1978 (see the list of SMSA reports from the AHS in the introduction of this report).

Three of the larger SMSA's (Atlanta, GA; Philadelphia, PA-NJ; and San Francisco-Oakland, CA) in the 1982 group of SMSA's were represented by a sample size of 15,000 designated housing units in the 1975 and 1978 surveys, evenly divided between the central city and the balance of the respective SMSA. The 9 remaining SMSA's (Cincinnati, OH-KY-IN; Columbus, OH; Kansas City MO-KS; New Orleans, LA; Paterson-Clifton-Passaic, NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; and San Diego, CA) in the 1982 group were represented by a sample of 5,000 designated housing units in the 1975 and 1978 surveys distributed proportionately between the central city and balance of the respective SMSA based on the distribution of total housing units in each sector.

In the 1975 and 1978 surveys, AHS sample units were divided among 12 panels with one-twelfth of the sample housing units being interviewed each month. Due to budget limitations, panels 1 through 3 were dropped from all SMSA's for the 1982 survey, excluding the central cities of the Paterson-Clifton-Passaic, NJ, SMSA. Due to additional budget limitations for the 1982 survey, it also became necessary to drop part of the sample of housing units from panels 4 through 12 in the Atlanta, GA; New Orleans, LA; Philadelphia, PA-NJ; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. The effect of this reduction was to maintain a sample size of 4,250 in all SMSA's. The interviewing was done during the period April 1982 through December 1982 with one-ninth of the sample housing units being visited each month.

In this SMSA, 4,075 housing units were eligible for interview. Of these sample housing units, 163 interviews were not obtained because, for occupied sample units, the occupants refused to be interviewed, were not at home after repeated visits, or were unavailable for some other reason; or, for vacant units, no informed respondent could be found after repeated visits. In addition to the 4,075 housing units eligible for interview, 176 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 1982 survey—The sample housing units designated to be interviewed in the 1982 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 1978 survey and remained in sample after the 1982 panel reduction.
2. All sample housing units that were type A noninterviews (i.e., units eligible to be interviewed) or type B noninterviews (i.e., units not eligible for interview at the time of the survey but which could become eligible in the future) in the 1978 survey and remained in sample after the 1982 panel reduction. (For a list of reasons for type A and type B noninterviews, see the facsimile of the 1982 AHS questionnaire, page App-18.)
3. All sample housing units that were selected from a listing of new residential construction building permits issued since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented the housing units built in permit-issuing areas since the 1978 survey.)
4. All sample housing units that were added to sample segments in the nonpermit universe since the 1978 survey and remained in sample after the 1982 panel reduction. (This sample represented additions to the housing inventory in nonpermit-issuing areas since the 1978 survey.)
5. All sample housing units that were selected as part of the 1978 Coverage Improvement Program. (This sample represented most of the housing units which, until 1978, did not have a chance of selection.)

Selection of the 1978 AHS-SMSA sample—The sample for the SMSA's which, in 1970, were 100-percent permit-issuing was selected from two sample frames—housing units enumerated in the 1970 Census of Population and Housing in areas under the jurisdiction of permit-issuing offices (the permit-issuing universe) and housing units constructed in permit-issuing areas since the 1970 census (the new construction universe). In addition, the sample for those SMSA's which were not 100-percent permit-issuing in 1970 included a sample selected from a third frame—those housing units located in areas not under the jurisdiction of permit-issuing offices (the nonpermit universe). In 1970, the following four SMSA's were 100-percent permit-issuing: Paterson-Clifton-Passaic, NJ; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA. The remaining eight SMSA's contain a sample from the nonpermit universe.

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

The major portion of the sample in each SMSA was selected from a file which represented the 20-percent sample of housing units enumerated in permit-issuing areas of the SMSA during the 1970 Census of Population and Housing. This file contained records for occupied housing units, vacant housing units, and housing units in certain special places or group quarters. Sampling operations were done separately for the special place and group quarters records, and for the occupied and vacant housing unit records. Before the sample was selected from the occupied and vacant housing unit records, the occupied housing unit records were stratified by race of head (non-Black/Black) and the vacant records were stratified into four categories pertaining to the value or rent associated with the vacant housing units. The occupied housing unit records were further stratified so that each unit was assigned to 1 of 50 strata according to its tenure, family size, and household income category as illustrated by the following table:

Household income	Tenure	
	Owner—Family size	Renter—Family size
	1 2 3 4 5+	1 2 3 4 5+
Under $3,000		
$3,000 to $5,999 . . .		
$6,000 to $9,999 . . .		
$10,000 to $14,999 .		
$15,000 and over . . .		

Thus, for this SMSA, the occupied housing unit records from the permit-issuing universe were assigned to 1 of 100 strata for either the central city or for the balance, and the vacant housing unit records were assigned to 1 of the 4 vacant strata for either the central city or for the balance of the SMSA. A sample selection procedure was then instituted that would produce one-half of the desired sample size. However, whenever a record was selected to be in sample, the housing unit record adjacent to it on the file was also selected to be in sample, thereby insuring the necessary designated sample size.

Before the sample was selected from the group quarters and special place records, the records were stratified by census tract and census enumeration district (ED) within the central city and within the balance of the SMSA. A sample of special place records was then selected by a procedure that produced one-quarter of the desired sample size. However, at the time of the survey, the housing units at each of the special places were listed and subsampled at a rate which produced an expected four sample units, thereby insuring the necessary designated sample size.

The second frame from which this SMSA sample was selected was a list of new construction building permits issued since 1970 (i.e., the new construction universe). The sample selection from the list of new construction building permits was an independent operation within this SMSA. Prior to sample selection, the

list of permits was chronologically stratified by the date the permits were issued, and clusters of an expected four (usually adjacent) housing units were formed. These clusters were then sampled for inclusion at the overall sampling rate.

For those SMSA's which were not 100-percent permit-issuing, the remainder of the AHS sample was selected from a frame consisting of areas not under the jurisdiction of permit-issuing offices (i.e., the nonpermit universe). The first step in the sampling operation for the nonpermit universe was the selection (using the overall sampling rate) of a sample of census enumeration districts within these areas. Prior to this sample selection, the ED's were stratified by census tract within the central city and within the balance of the SMSA. The probability of selection of an ED was proportionate to the following measure of size.

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

The sample ED's were then divided into segments; i.e., small land areas with well-defined boundaries having an expected size of four, or a multiple of four, housing units. At the time of the survey, those segments that did not have an expected size of four were further subdivided to produce an expected four sample housing units.

The next step was the selection of one of these segments within each sample ED. All housing units in existence at the time of interview in these selected segments were eligible for sample. Thus, housing units enumerated in the 1970 census as well as housing units built since the 1970 census are included.

Sample Selection for the 1978 Coverage Improvement Program—The Coverage Improvement Program was undertaken to correct certain deficiencies in the AHS-SMSA sample from the permit-issuing and new construction universes. The coverage deficiencies included the following units:

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

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 SMSA. 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. 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 Rochester, NY, SMSA.

In the remaining 11 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. These procedures added an estimated 9,486 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 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 659 units to the coverage of the housing inventory of this SMSA.

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

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

Initially, a subsample of AHS sample units was selected from the permit-issuing universe at a rate of 1 in 24. Then, succeeding structures in a defined path of travel to the right of the structure containing the sample unit were listed until eight structures (excluding the sample unit structure) were found that had been eligible to be selected for the AHS. Finally, the intervening structures that did not have a chance of selection in the AHS were identified and units within these structures were interviewed.

In cases where the interviewer workload would have been too great, a representative subsample of units within these structures was selected. This procedure added an estimated 16,877 units to the coverage of the housing inventory of this SMSA.

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

1. Units missed in the 1970 census.
2. Units converted to residential use since the 1970 census in structures that contained some residential units in 1970.

First, a subsample of AHS housing units in multiunit structures of less than 10 units was selected from the permit-issuing universe. Second, for the multiunit structure selected above, all housing units were listed and matched to the 1970 census. Any missed housing units were then assigned for interview. This procedure added an estimated 2,023 units to the coverage of the housing inventory of this SMSA.

1978-1982 additions to the housing inventory—In the permit-issuing universe, a sample of new construction building permits, issued since the 1978 survey, was selected to represent housing units built in permit-issuing areas since the 1978 survey. Sampling procedures were identical to those used in selecting the 1970-1978 new construction sample, which were described previously. In the nonpermit universe, sample segments were dependently recanvassed, using listing sheets from 1977, to identify any housing units missed in the 1978 survey or any housing units added since the 1978 survey.

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 1982 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 1982 housing inventory) and estimates pertaining to characteristics of housing units removed from the housing inventory since 1978 (i.e., 1978-1982 lost units). Each type of estimate employed separate, although similar, estimation procedures.

1982 housing inventory—The AHS estimates of characteristics of the 1982 housing inventory were produced using a one-stage ratio estimation procedure for the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's. A two-stage ratio estimation procedure was used for the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. A three-stage ratio estimation procedure was used for the San Bernardino-Riverside-Ontario, CA, SMSA.

Prior to the implementation of the ratio estimation procedures, the basic weight (i.e., the inverse of the probability of selection) for each interviewed sample housing unit was adjusted to account for the noninterviews previously mentioned. This noninterview adjustment was done separately for occupied and vacant housing units. The noninterview adjustment factor was equal to the following ratio:

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

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

The following ratio estimation procedure was employed for all sample housing units from the permit-issuing universe. This factor was computed separately for all sample housing units within each permit-issuing universe noninterview cell mentioned previously. The ratio estimation factor for each cell was equal to the following:

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

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

This ratio estimation procedure was introduced to correct the probabilities of selection for samples in each of the strata used in the sample selection of the permit-issuing universe. Prior to the AHS sample selection within each SMSA, housing units already selected for other Census Bureau surveys were deleted from the permit-issuing universe. The same probability of selection was then applied to the remaining units to select the AHS sample. Since the number of housing units deleted from the AHS

universe frame was not necessarily proportional among all strata, some variation in the actual probability of selection between strata were introduced during the AHS sample selection process.

For the San Bernardino-Riverside-Ontario, CA, SMSA, the next ratio estimation procedure was employed for new construction mobile home units. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of mobile home units for each sector of the SMSA (central city and balance) by year built groups (1970-1974 and 1975-1982) to an independent estimate of mobile home units for the corresponding sector and year built group. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of mobile homes for the corresponding sector and year built group}}{\text{Sample estimate of mobile homes for the corresponding sector and year built group}}$$

The computed ratio estimation factor was then applied to the existing weight for each sample housing unit in the corresponding cell.

The next ratio estimation procedure was applied in the Columbus, OH; Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Antonio, TX; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco-Oakland, CA, SMSA's. This procedure involved the ratio estimation of the AHS-SMSA weighted sample estimate of the August 1982 housing inventory in each sector of the SMSA (central city and balance) to an independent estimate of the total housing units in each sector. This ratio estimation factor equaled the following:

$$\frac{\text{Independent estimate of the August 1982 housing inventory for the corresponding sector of the SMSA}}{\text{AHS-SMSA sample estimate of the housing inventory for the corresponding sector of the SMSA}}$$

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

Independent estimates of total housing units were derived for the central city, balance, and the total SMSA for the 12 SMSA's in 1982. These estimates were derived by using 1980 census counts in conjunction with estimates of change in the housing inventory between the 1980 census and the 1982 survey. These estimates of change were based on estimates of new construction permit authorizations and post-census demolition permits.

The sample estimates of total housing units using the existing sample weights were compared to the corresponding independent estimates for the central city, balance, and total for each of the 12 SMSA's, and the estimates which showed the most likely level of net growth since the 1980 census in both the central city and balance, as well as, the total SMSA were used in this ratio estimation procedure.

As a result of this analysis, the independent estimates were used in the Kansas City, MO-KS; New Orleans, LA; Philadelphia, PA-NJ; Rochester, NY; San Bernardino-Riverside-Ontario, CA; San Diego, CA; and San Francisco, CA, SMSA's. The sample estimates were used in the Atlanta, GA; Cincinnati, OH-KY-IN; and Paterson-Clifton-Passaic, NJ, SMSA's.

For the Columbus, OH and San Antonio, TX, SMSA's the independent estimates were used for the total SMSA only. The independent estimates were then proportioned between the central city and the balance of each SMSA according to the central city/balance distribution of the total housing units as given by the sample estimates. Due to the central city boundary definitional changes and central city permit office annexations since 1970 in these SMSA's, the independent estimates did not reflect the 1970 central city/balance SMSA definitions. Since the design of the AHS-SMSA sample is based on the 1970 SMSA definition, independent estimates which best reflected the 1970 central city/balance SMSA definitions were used.

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

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

1978-1982 lost housing units—The 1978-1982 lost housing units (housing units removed from the inventory) estimates employed the one-stage ratio estimation procedure used to produce the AHS-SMSA estimates of the 1978 housing inventory as was described in the 1978 Current Housing Report, Series H-170, *Housing Characteristics for Selected Metropolitan Areas.* Since the 1978-1982 lost housing units existed, by definition, in the 1978 housing inventory, there was a 1978 housing inventory weight associated with each 1978-1982 lost unit. This weight was used to tabulate the estimates of the characteristics of the 1978-1982 lost housing units.

1978 estimation procedure—This report presents data on the housing characteristics of the 1978 housing inventory from the 1978 Annual Housing Survey SMSA sample. The AHS-SMSA estimation procedure employed a one-stage ratio estimation process. A detailed description of this ratio estimation procedure can be found in the AHS Series H-170 reports for 1978.

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

RELIABILITY OF THE ESTIMATES

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

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

Obtaining a measurement of the total nonsampling error associated with the estimates from a survey is very difficult, considering the number of possible sources of error. However, an attempt was made to measure some of the nonsampling errors associated with the estimates for the 1970 Census of Population and Housing and the 1982 AHS-SMSA sample.

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

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

AHS-SMSA—The AHS-SMSA evaluation study was administered to a sample of units selected from the regular interviews containing recent movers who had previous residences within the SMSA. The purpose of the study was to evaluate the accuracy of responses regarding housing characteristics of recent mover's previous residence. The results of which are presented in the Census Bureau memorandum "Analysis of the Results from the AHS-SMSA Group AA-2 1982 Evaluation Study." (June 17, 1983)

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

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

Deficiencies also exist in ED's where area sampling methods are used. It had been assumed that all housing units located inside these ED's would be represented in the sample. However, it has been estimated that the 1975 AHS sample missed as much as 2 percent of all housing units in these ED's because they were not listed during the canvassing. It should be noted that since these ED's were recanvassed for the 1978 and 1982 surveys, the number of missed housing units may be considerably less for 1982.

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

Rounding errors—For errors associated with processing, rounding of estimates introduces another source of error in the data, the severity of which depends on the statistics being measured. The effect of rounding is significant relative to the sampling error only for small percentages or small medians, when these figures are derived from relatively large bases (e.g., median number of persons per household). This means that confidence intervals formed from the standard errors given may be distorted, and this should be taken into account when considering the results of the survey.

Sampling errors for the AHS-SMSA sample—The particular sample used for this survey is one of a large number of possible samples of the same size that could have been selected using the same sample design. Even if the same questionnaires, instructions, and interviewers were used, estimates from each of the different samples would differ from each other. The sampling error of a survey estimate provides a measure of the variation among the estimates from all possible samples and thus, is a measure of the precision with which an estimate from a sample approximates the average result of all possible samples.

One common measure of the sampling error is the standard error. As calculated for this report, the standard error reflects the variation in the estimates due to sampling and nonsampling

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

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

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

The figures presented in the tables that follow (pages App-51 and App-52) are approximations to the standard errors of various estimates shown in this report for this SMSA. In order to derive standard errors that would be applicable to a wide variety of items and also could be prepared at a moderate cost, a number of approximations were required. As a result, the tables of standard errors provide an indication of the order of magnitude of the standard errors rather than precise standard errors for any specific item. Standard errors applicable to estimates of characteristics of the 1978 housing inventory can be found in the AHS Series H-170 reports for 1978.

Table I (page App-51) presents the standard errors applicable to estimates of characteristics of the 1982 housing inventory as well as estimates of characteristics of the 1978-1982 lost housing units (housing units removed from the inventory). Linear interpolation should be used to determine the standard errors for estimates not specifically shown in this table. The standard errors on the AHS estimates of the population in housing units shown in tables A-1, B-1, and C-1 of part A of this report are 50,800 for the total SMSA, 25,500 for the central city of the SMSA, and 39,000 for the balance of the SMSA.

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

Tables II and III (pages App-51 and App-52) presents the standard errors of estimated percentages for the 1982 housing inventory as well as estimated percentages of the 1978-1982 lost housing units (housing units removed from the inventory). Two-way interpolation should be used to determine standard errors for estimated percentages not specifically shown in tables II and III.

Included in tables I, II, and III are estimates of standard errors for estimates of zero and zero percent. These estimates of standard errors are considered as overestimates of the true standard errors and should be used primarily for construction of confidence intervals for characteristics when an estimate of zero is obtained.

For ratios, 100 (x/y), where x is not a subclass of y, tables II and III underestimate the standard error of the ratio when there is little or no correlation between x and y. For this type of ratio, a better approximation of the standard error may be obtained by letting the standard error of the ratio be approximately equal to:

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

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

Illustration of the use of the standard error tables—Table A-1 of part A of this report shows that in 1982 there were 682,400 owner-occupied housing units in this SMSA. Interpolation using table I of this appendix shows that the standard error of an estimate of this size is approximately 11,480. 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
600,000	11,410
682,400	x
700,000	11,490

The entry of "x" is determined as follows by vertically interpolating between 11,410 and 11,490.

$$682{,}400 - 600{,}000 = 82{,}400$$
$$700{,}000 - 600{,}000 = 100{,}000$$

$$11{,}410 + \frac{82{,}400}{100{,}000}(11{,}490 - 11{,}410) = 11{,}480$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 670,920 to 693,880 housing units. Therefore, a conclusion that the average estimate, derived from all possible samples, of 1982 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 664,030 to 700,770 housing units with 90 percent confidence; and that the average estimate lies within the interval from 659,440 to 705,360 housing units with 95 percent confidence.

Table A-1 of part A also shows that of the 682,400 owner-occupied housing units, 197,400, or 28.9 percent, had two bedrooms. Interpolation using table II of this appendix (i.e., interpolation on both the base and percent) shows that the standard error of the 28.9 percent is approximately 1.1 percentage points. The following interpolation procedure was used.

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

Base of percentage	Estimated percentage		
	25 or 75	28.9	50
600,000	1.2	a	1.3
682,400		p	
700,000	1.1	b	1.2

1. The entry for cell "a" is determined by horizontal interpolation between 1.2 and 1.3.

$$28.9-25.0 = 3.9$$
$$50.0-25.0 = 25.0$$

$$1.2 + \frac{3.9}{25.0}\ (1.3-1.2) = 1.2$$

2. The entry for cell "b" is determined by horizontal interpolation between 1.1 and 1.2.

$$28.9-25.0 = 3.9$$
$$50.0-25.0 = 25.0$$

$$1.1 + \frac{3.9}{25.0}\ (1.2-1.1) = 1.1$$

3. The entry for "p" is then determined by vertical interpolation between 1.2 and 1.1.

$$682{,}400-600{,}000 = 82{,}400$$
$$700{,}000-600{,}000 = 100{,}000$$

$$1.2 + \frac{82{,}400}{100{,}000}\ (1.1-1.2) = 1.1$$

Consequently, the 68-percent confidence interval, as shown by these data, is from 27.8 to 30.0 percent; the 90-percent confidence interval is from 27.1 to 30.7 percent; and the 95-percent confidence interval is from 26.7 to 31.1 percent.

Differences—The standard errors shown are not directly applicable to differences between two sample estimates. The standard error of a difference between estimates is approximately equal to the square root of the sum of the squares of the standard error of each estimate considered separately. This formula is quite accurate for the difference between estimates of the same characteristics in two different SMSA's or the difference between separate and uncorrelated characteristics in the same SMSA. If there is a high positive correlation between the two characteristics, the formula will overestimate the true standard error; but if there is a high negative correlation, the formula will underestimate the true standard error. Due to the overlap of the 1978 and 1982 AHS-SMSA samples a positive correlation should be expected when making comparisons between 1978 and 1982 characteristics.

Illustration of the computation of the standard error of a difference—Table A-1 of part A of this report shows that in 1982 there were 304,400 owner-occupied housing units with three bedrooms in this SMSA. Thus, the apparent difference, as shown by these data, between owner-occupied housing units with two bedrooms and owner-occupied housing units with three bedrooms is 107,000. Table I shows the standard error of 197,400 is approximately 8,080, and the standard error of 304,400 is approximately 9,560. Therefore, the standard error of the estimated difference of 107,000 is about 12,520.

$$12{,}520 = \sqrt{(8{,}080)^2 + (9{,}560)^2}$$

Consequently, the 68-percent confidence interval for the 107,000 difference is from 94,480 to 119,520 housing units. Therefore, a conclusion that the average estimate derived from all possible samples, of this difference, lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, the 90-percent confidence interval is from 86,970 to 127,030 housing units, and the 95-percent confidence interval is from 81,960 to 132,040 housing units. Thus, we can conclude with 95 percent confidence that the number of 1982 owner-occupied housing units with three bedrooms is greater than the number of owner-occupied units with two bedrooms since the 95-percent confidence interval does not include zero or negative values.

Medians—For medians presented in certain tables, the sampling error depends on the size of the base and on the distribution upon which the median is based. An approximate method for measuring the reliability of the estimated median is to determine an interval about the estimated median so that there is a stated degree of confidence that the average median from all possible samples lies within the interval. The following procedure may be used to estimate confidence limits of a median based on sample data:

1. From tables II and III, determine the standard error of a 50-percent characteristic on the base of the median.
2. Add to and subtract from 50 percent, the standard error determined in step 1.
3. Using the distribution of the characteristics, determine the confidence interval corresponding to the two points established in step 2. To find the lower endpoint of the confidence

interval, it is necessary to know into which interval of the distribution the lower percentage limit falls. Similarly, to find the upper endpoint to the confidence interval, it is necessary to know into which interval of the distribution the upper percentage limit falls. Note that these two distribution intervals could be different, although this will not happen very often.

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

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

Illustration of the computation of the 95-percent confidence interval of a median—Table A-1 of part A of this report shows the median number of persons for owner-occupied housing units is 2.4. The base of the distribution from which this median was determined is 682,400 housing units.

1. Interpolation using table II shows that the standard error of 50 percent on a base of 682,400 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 the distribution for "persons" in table A-1 of part A, the interval for owner-occupied housing units with two persons (for purposes of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) contains the 47.6 percent derived in step 2. About 111,700 housing units or 16.4 percent fall below this interval, and 240,300 housing units or 35.2 percent fall within this interval. By linear interpolation, the lower limit of the 95-percent confidence interval is found to be about:

$$1.5 + (2.5-1.5)\ \frac{(47.6-16.4)}{35.2} = 2.4$$

Similarly, the interval for owner-occupied housing units with three persons contains the 52.4 percent derived in step 2. About 352,000 housing units or 51.6 percent fall below this interval, and 119,100 housing units or 17.5 percent fall within this interval. The upper limit of the 95-percent confidence interval is found to be about:

$$2.5 + (3.5-2.5)\ \frac{(52.4-51.6)}{17.5} = 2.5$$

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

TABLE I. Standard Errors for Estimated Number of Housing Units in the 1982 Housing Inventory and for Estimated Number of 1978-1982 Lost Units for the San Francisco-Oakland, CA, SMSA, for the Central Cities of the SMSA and for the Balance (Not in Central Cities) of the SMSA

(68 chances out of 100)

Size of estimate	Standard error[1]			Size of estimate	Standard error[1]		
	SMSA	In central cities	Not in central cities		SMSA	In central cities	Not in central cities
0	390	420	330	250,000	8,890	7,060	7,760
100	390	420	330	300,000	9,520	6,800	8,160
200	390	420	330	400,000	10,460	5,050	8,590
500	440	460	410	470,600	10,930	—	8,630
700	520	540	480	500,000	11,080	—	8,580
1,000	620	650	580	600,000	11,410	—	8,130
2,500	980	1,030	910	700,000	11,490	—	7,130
5,000	1,390	1,450	1,290	800,000	11,320	—	5,320
10,000	1,960	2,040	1,820	894,800	10,930	—	—
25,000	3,080	3,170	2,850	900,000	10,900	—	—
50,000	4,320	4,360	3,970	1,000,000	10,180	—	—
75,000	5,240	5,170	4,790	1,100,000	9,100	—	—
100,000	5,990	5,780	5,450	1,200,000	7,500	—	—
150,000	7,190	6,590	6,460	1,300,000	4,910	—	—
200,000	8,130	6,990	7,200	1,365,400	—	—	—

[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 balance (not in central cities) and 1.1 for the central cities estimates. Standard errors of estimates pertaining to total housing units for the central cities, balance, and total SMSA are assumed to be equal to zero where the estimate for the total SMSA was derived from an independent estimate.

TABLE II. Standard Errors for Estimated Percentages of Housing Units in the 1982 Housing Inventory and for Estimated Percentages of 1978-1982 Lost Units for the San Francisco-Oakland, CA, SMSA, and for the Central City of the SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
300	58.6	58.6	58.6	58.6	58.6	59.5	250,000	0.2	0.4	0.9	1.2	1.8	2.1
500	45.9	45.9	45.9	45.9	45.9	46.1	300,000	0.14	0.4	0.8	1.1	1.6	1.9
700	37.8	37.8	37.8	37.8	37.8	38.9	400,000	0.11	0.3	0.7	1.0	1.4	1.6
1,000	29.8	29.8	29.8	29.8	29.8	32.6	500,000	0.08	0.3	0.6	0.9	1.3	1.5
2,500	14.5	14.5	14.5	14.5	17.8	20.6	600,000	0.07	0.3	0.6	0.8	1.2	1.3
5,000	7.8	7.8	7.8	8.7	12.6	14.6	700,000	0.06	0.2	0.5	0.7	1.1	1.2
10,000	4.1	4.1	4.5	6.2	8.9	10.3	800,000	0.05	0.2	0.5	0.7	1.0	1.2
25,000	1.7	1.7	2.8	3.9	5.6	6.5	900,000	0.05	0.2	0.5	0.7	0.9	1.1
50,000	0.8	0.9	2.0	2.8	4.0	4.6	1,000,000	0.04	0.2	0.4	0.6	0.9	1.0
75,000	0.6	0.7	1.6	2.3	3.3	3.8	1,100,000	0.04	0.2	0.4	0.6	0.9	1.0
100,000	0.4	0.6	1.4	2.0	2.8	3.3	1,200,000	0.04	0.2	0.4	0.6	0.8	0.9
150,000	0.3	0.5	1.2	1.6	2.3	2.7	1,300,000	0.03	0.2	0.4	0.5	0.8	0.9
200,000	0.2	0.5	1.0	1.4	2.0	2.3	1,400,000	0.03	0.2	0.4	0.5	0.8	0.9

[1]Standard errors are presented to the nearest one-tenth of one percentage point except when the standard error is less than fifteen-hundredths of one percentage point; in those cases, the standard error is shown to the nearest one-hundredth of one percentage point. For estimates pertaining to new construction, the standard errors shown in the table should be multiplied by a factor of 1.2 for the total SMSA and 1.1 for the central cities estimates.

TABLE III. Standard Errors for Estimated Percentages of Housing Units in the 1982 Housing Inventory and for Estimated Percentages of 1978-1982 Lost Units for the Balance (Not in Central Cities) of the San Francisco-Oakland, CA, SMSA

(68 chances out of 100)

Base of percentage	Estimated percentage[1]						Base of percentage	Estimated percentage[1]					
	0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50		0 or 100	1 or 99	5 or 95	10 or 90	25 or 75	50
300	52.7	52.7	52.7	52.7	52.7	52.8	150,000	0.2	0.5	1.0	1.4	2.0	2.4
500	40.0	40.0	40.0	40.0	40.0	40.9	200,000	0.2	0.4	0.9	1.2	1.8	2.0
700	32.3	32.3	32.3	32.3	32.3	34.5	250,000	0.13	0.4	0.8	1.1	1.6	1.8
1,000	25.0	25.0	25.0	25.0	25.0	28.9	300,000	0.11	0.3	0.7	1.0	1.4	1.7
2,500	11.8	11.8	11.8	11.8	15.8	18.3	400,000	0.08	0.3	0.6	0.9	1.3	1.4
5,000	6.3	6.3	6.3	7.8	11.2	12.9	500,000	0.07	0.3	0.6	0.8	1.1	1.3
10,000	3.2	3.2	4.0	5.5	7.9	9.1	600,000	0.06	0.2	0.5	0.7	1.0	1.2
25,000	1.3	1.3	2.5	3.5	5.0	5.8	700,000	0.05	0.2	0.5	0.7	0.9	1.1
50,000	0.7	0.8	1.8	2.5	3.5	4.1	800,000	0.04	0.2	0.4	0.6	0.9	1.0
75,000	0.4	0.7	1.5	2.0	2.9	3.3	900,000	0.04	0.2	0.4	0.6	0.8	1.0
100,000	0.3	0.6	1.3	1.7	2.5	2.9							

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