Appendix B. Source and Accuracy of the Estimates

SAMPLE DESIGN App-59
Selection of sample areas App-59
Selection of the sample housing units from the 1980 census App-59
Selection of new construction housing units in permit issuing areas App-60
Housing Unit Coverage Study sample App-60
Housing units added since the 1980 census App-60
Telephone interviewing experiments App-60

ESTIMATION App-61

ACCURACY OF THE ESTIMATES App-62
Sampling errors App-62
Standard error table locator App-63
Standard errors of estimates of levels App-63
Standard errors of estimates of percentages App-63
Standard errors of ratios App-64
Standard errors of differences App-64
Standard errors of medians App-64
Nonsampling errors App-65
Possible effects of decentralized telephone interviewing on the data App-65
Possible effects of Computer Assisted Telephone interviewing (CATI) on the data App-65
Reconciliation experiment App-67
Reinterview program App-67
Processing errors App-68
Standard Error Table Locator App-70

SAMPLE DESIGN

The 1989 estimates contained in this report are based on data collected from July 1989 through December 1989 for the American 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 sample for this survey was spread over 394 sample areas (called primary sampling units) comprising 878 counties and independent cities with coverage in each of the 50 States and the District of Columbia.

Approximately 49,400 sample housing units were selected for interview for the 1989 AHS. Of this number, about 2,700 were found to be ineligible because they no longer existed or information relevant to the 1989 housing inventory could not be obtained for the unit. Of the approximately 46,700 units (both occupied and vacant) that were eligible for interview, about 1,900 were classified as "non-interviews" because either no one was home after repeated visits, the respondent refused to be interviewed, or the interviewer was unable to locate the unit.

Selection of sample areas. The United States was divided into areas made up of counties and independent cities referred to as primary sampling units (PSU's). Of these PSU's, 170 were known as self-representing since the sample from the PSU represented only that PSU. These 170 PSU's were in sample with certainty. The remaining PSU's were grouped into strata and were referred to as non-self-representing, since the sample of housing units from the sample PSU represented all PSU's, both sample and nonsample, in the stratum. These non-self-representing sample PSU's were selected in two steps.

First, the Current Population Survey (CPS) formed groups consisting of one or more PSU's. In groups consisting of more than one PSU, one PSU was selected to represent all PSU's in a CPS stratum. The second step involved selecting a subset of PSU's selected by the CPS. The PSU's selected for the CPS sample (some of which were self-representing for CPS and some of which were non-self-representing for CPS) were grouped again for the AHS. For groups consisting of only one PSU selected for the CPS, that PSU was also selected for the AHS. For groups consisting of more than one PSU selected for the CPS, one PSU was selected for the AHS.

Selection of the sample housing units from the 1980 census. The overall sampling rate used to select the sample of housing units from the 1980 census for the 1989 AHS was about 1 in 2,148. The within-PSU sampling rate was determined so that the overall probability of selection for each sample housing unit was the same (e.g., if the probability of selecting a non-self-representing PSU was 1 in 10, then the within-PSU sampling rate would be 1 in 214.8).

In areas where addresses were, for the most part, complete and where new construction is monitored by permits (these areas will be referred to as address enumeration districts [ED's]), a sample of housing units that received long-form questionnaires in the 1980 census was selected directly from a list of all such housing units based on certain housing and geographic information of the housing unit. A sample of living quarters that did not meet the definition of a housing unit (e.g., military barracks, college dorm) was selected independently from housing

units in address ED's. This sample of living quarters that were not housing units was used to identify units that converted to housing units since the census.

In areas where at least 4 percent of the addresses were incomplete or inadequate, or where new construction was not monitored by building permits (most rural areas), a sample of 1980 census units that received long-form questionnaires was selected in several steps (these areas will be referred to as area ED's). First, the areas were grouped and a sample of areas was chosen. Next, an area of land, known as a segment, was chosen within each sample area. Finally, a sample of housing units that received 1980 census long forms was selected within the segment.

Selection of new construction housing units in permit issuing areas. The sample of permit new construction was selected from building permits issued such that the units are expected to be completed after April 1, 1980. For certain areas and structure sizes, this included permits issued as early as March 1979, but, for the most part, includes permits issued since July 1979. Only nonmobile home new construction is covered by the building permit frame. Within each PSU, building permits were selected so that the sample would be representative in terms of geography and month of issue for permits. Clusters of approximately four housing units were created. Housing units in these clusters were subsampled at the rate of 1 in 4, yielding clusters of size 1.

Housing Unit Coverage Study sample. Housing units at addresses missed in the 1980 census or units that were at inadequately described addresses in the census address registers did not have a chance of being selected for the AHS sample. A special study, done as part of the 1980 census, called the Housing Unit Coverage Study identified such units. A sample of these units was included in the AHS sample.

Housing units added since the 1980 census. Housing units added to the inventory since the 1980 census were represented using two methods. One method identified within-structure additions. These are units in structures that had a chance of being in sample because they contained at least one unit enumerated in the 1980 census. This method was used for the Housing Unit Coverage Study sample as well. The other method identified whole structure additions. These are units in structures for which none of the units in the structure were enumerated in the 1980 census.

In area ED's, all within-structure additions in structures containing at least one sample unit were interviewed for AHS. In address ED's, all within-structure additions in 1- to 15-unit structures containing at least one sample unit were interviewed for AHS. In 16-or-more-unit structures in address ED's, only units falling on AHS sample lines were interviewed for AHS.

In address ED's, whole-structure additions were identified using area sampling methods. Under area sampling, all housing units within a land area are first listed and then a systematic sample is selected using a start with and take every so that a desired sample size is achieved based on the expected number of units within the segment. Land areas in sample for the Health Interview Survey in 1985 were used. Only Health Interview Survey areas that were in AHS PSU's or in Health Interview Survey PSU's adjacent to AHS PSU's were used. Also, only units that were not already assigned to the Health Interview Survey were eligible. These units were then matched to the 1980 census address registers. If the address matched to the census, the unit was ineligible. (Only the basic address, i.e., 801 Main Street, had to match. Apartment number, mobile home site number, etc., did not have to match). At the time of listing, eligible units were then screened further so that only units with no previous chance of coming into sample were picked up. (The screening eliminated units such as nonmobile home new construction, which is covered by building permits, and census misses.) This address ED coverage improvement operation was not updated in 1989.

In area ED's where new construction is not monitored by building permits, all land areas chosen for the sample in area ED's were used. An expected four units were chosen using area sampling methods within these land areas to identify whole-structure additions. This sample was screened at the time of listing using the same criteria as for address ED's. However, this sample was not matched to the census. One important difference to note is that new construction was not eliminated during the screening process.

In area ED's where new construction is monitored by building permits, only one-third of the land areas chosen for the sample in area ED's was used. An expected eight units were chosen using area sampling methods within these segments to identify whole structure additions. This sample was screened at the time of listing using the same criteria as for address ED's. Again, this sample was not matched to the census. Nonmobile home new construction was eliminated by the screening process because it is covered by the building permit frame. For the area ED coverage improvement, only certain panels were updated in 1989 to pick up whole-structure additions since the 1987 enumeration

Telephone interviewing experiments. A large-scale Computer Assisted Telephone Interviewing (CATI) experiment was conducted as part of the 1987 enumeration of AHS-National to investigate the effects of CATI interviewing on AHS-National (AHS-N) data. The 1987 sample was divided into six panels. Two of the six panels (panels 5 and 6) were randomly assigned to a maximum CATI treatment (about 16,000 cases). The other four panels (about 32,000 units) were assigned to a maximum decentralized (local) telephone interviewing treatment (i.e., the non-CATI treatment).

Units in the CATI treatment sample that were not eligible to be interviewed by CATI were screened out and sent to the field for a personal visit interview. These screened units included new construction added since 1985, the supplemental rural sample added in 1987, 1985 noninterviews, 1985 vacant units, 1985 units temporarily occupied by persons with usual residence elsewhere (URE's), households with eight or more members, multiunit mobile homes, special places, unit address/structure type inconsistencies, and units interviewed in 1985 indicating that they didn't have a telephone number at which they could be contacted. The remaining 10,400 units, which were units interviewed in 1985 and for which a telephone number was provided, were assigned to the Hagerstown Telephone Center to attempt CATI. Actually interviewed by CATI were 6,400 units, which is 61 percent of eligible cases and 40 percent of the CATI treatment sample. The eligible units not interviewed by CATI were recycled to the field for a personal visit or decentralized telephone interview.

Within the non-CATI treatment, about 40 percent of the units were actually interviewed by telephone. Those units not eligible for interview by telephone, as well as the eligible units that couldn't actually be interviewed by telephone, were assigned for personal visit interviews.

Preliminary analysis of the 1987 AHS-National CATI experiment indicated that, for at least some characteristics, CATI had a substantial effect on the data. Another CATI experiment was conducted in 1989 to further study the effect of CATI on AHS data. This time, panels 1 and 5 were assigned for maximum CATI and panels 4 and 6 for maximum decentralized telephone interviewing. In addition, about 4,000 cases were assigned for maximum CATI in 71 PSU's in panels 2 and 3. The intent was to target these CATI cases to PSU's having problems recruiting and maintaining interviewers. From a list prioritized by regional offices, specific PSU's were selected and designated for CATI. For the 1989 experiment, about 47 percent of the total AHS-N sample was initially designated for CATI. About 63 percent of these cases (about 14,500 cases) were actually eligible for CATI. Eligibility criteria analogous to those used for the 1987 experiment were used in 1989. About 60 percent of the eligible cases were completed by CATI, while the remaining cases were recycled to the field for decentralized telephone or personal visit interviews.

ESTIMATION

After assigning each unit a weight that reflected the correct probability of selection for the unit, the AHS weighting procedure consisted of two phases. In the first phase, a series of adjustments were made to account for units that could not be interviewed for a number of reasons. For each of these adjustments, a factor was computed and applied to the appropriate units. The factors were equal to the following ratio:

$$\frac{\text{Housing units to be kept after factor applied} + \text{Housing units to be dropped after factor applied}}{\text{Housing units to be kept after factor applied}}$$

The housing units that are to be kept after a factor is applied will have that factor applied to them. The first of these adjustments was done in permit segments only, to account for permits that could not be sampled and units that could not be located. These were represented by all other units in permit segments including both interviews and noninterviews (excluding unable-to-locate units).

The second of the adjustments was done for units in structures built before April 1, 1980. It was done to account for units that could not be located. The unlocatable units were represented by both interviews and noninterviews (excluding unable-to-locate units).

The last of these adjustments was done to account for units that could not be interviewed because either no one was home after repeated visits or the respondent refused to be interviewed. When 1985 or 1987 AHS or 1980 census data was available, this information was used to determine the noninterview adjustment cell. The cells included characteristics such as tenure, geography, units in structure, and number of rooms. When previous data were not available, adjustment factors were computed separately using more general characteristics such as type of area and type of housing unit (i.e., mobile home, nonmobile home).

The second phase involved a three-stage ratio estimation procedure to adjust for the sampling of non-self-representing PSU's, to account for known sampling deficiencies in new construction, and to bring the sample estimate of housing units into close agreement with estimates derived from independent sources for several key characteristics.

The first-stage of this procedure was employed to reduce the contribution to the variance due to the sampling of non-self-representing PSU's. The procedure takes into account the differences that existed at the time of the 1980 census between the housing units estimated from the non-self-representing sample PSU's and the actual 1980 census count of housing units from all non-self-representing strata. Factors accounting for these differences were computed separately for 15 place-of-residence/tenure cells for the Northeast and Midwest regions, 35 place-of-residence/ethnicity-race/tenure cells for the South region, and 25 place-of-residence/ethnicity/tenure cells for the West region. The first-stage ratio estimation factor was equal to the following ratio:

$$\frac{\text{Actual 1980 census housing units for all non-self-representing strata in a cell}}{\text{Number of 1980 housing units in the same cell estimated from the sample non-self-representing PSU's}}$$

The numerators of the ratios were calculated by summing the 1980 census housing unit counts for each cell across all non-self-representing strata. For each cell, the

denominators were calculated by weighting the 1980 census housing unit counts from each non-self-representing sample PSU by the inverse of the probability of selection for that PSU and summing the weighted counts across all non-self-representing sample PSU's.

The second-stage of the ratio estimation procedure was employed to adjust the AHS sample estimate of new construction (i.e., units built since the 1980 census) to account for known deficiencies in the AHS sample (see the section on nonsampling errors). For nonmobile homes, the sample estimates were controlled to independently derived estimates from the Survey of Construction (SOC). For mobile homes, the sample estimates are controlled to independently derived estimates from the Survey of Mobile Home Placement (SMHP). These estimates were considered to be the best estimates available for these types of units. Factors were computed separately for each region. The second-stage factor was equal to the following ratio:

$$\frac{\text{Independently derived estimate for a cell}}{\text{AHS sample estimate in that cell}}$$

The denominators of the above ratio were obtained by summing the existing weight on each record after the first-stage of ratio estimation over all records for each cell in each region.

The third stage of the ratio estimation procedure was employed to adjust the AHS sample estimate of housing units to independently derived current estimates for certain key characteristics. It is believed that these characteristics are highly correlated with other characteristics of interest for AHS. This stage of the procedure was actually done in two steps for occupied units. During the first step, the sample estimate of occupied housing units was controlled to an independently derived estimate for tenure/ethnicity (i.e., Hispanic head of household–non-Hispanic head of household)/household-status cells for each region. After applying the factor computed in this step to the interviewed occupied units, the new sample estimate of occupied housing units was controlled to an independently derived estimate for tenure/race (i.e., Black head of household–non-Black head of household)/household-status cells for each region. The sample estimate of vacant housing units was controlled to an independently derived estimate for four type-of-vacant cells for each region. All third-stage factors were calculated in a similar manner using the following ratio:

$$\frac{\text{Independently derived estimate of housing units in a cell}}{\text{AHS sample estimate of housing units in that cell}}$$

For occupied units, the numerators of the factors were derived from data based on the CPS and the 1980 census. The 1980 census count of housing units was adjusted for net undercoverage and overcoverage. The CPS was used to measure changes since the census and to derive the distribution for the third-stage occupied cells.

For vacant units, the numerators of the factors were derived based on the distribution of vacant units from the Housing Vacancy Survey (HVS), a quarterly vacancy survey conducted by the Bureau of the Census.

The denominators of the factors were obtained by summing the weights, with all previous factors applied, on all records in a cell. For the Hispanic/non-Hispanic and vacant cells, this was the weight after the second stage of the ratio estimation procedure. For the Black/non-Black cells, this was the weight after the Hispanic/non-Hispanic portion of the third stage of the ratio estimation procedure.

The second stage and third stage of the ratio estimation procedure were iterated to bring the AHS sample estimates into closer agreement with all independent estimates used. The numerators of the factors were the same ones used previously. The denominators of the factors in this iterative process were obtained by summing the existing weights on all records in a cell. For example, for the second stage of the ratio estimation procedure, the existing weight after the third stage of the ratio estimation procedure from the previous iteration was used. The final weight that resulted from all iterations was used to produce the tabulations in this report.

The overall estimation procedure reduced the sampling error substantially for most statistics below what would have been obtained by simply weighting the sample by the inverse of the probability of selection.

ACCURACY OF THE ESTIMATES

There are two types of possible errors associated with estimates based on data from sample surveys—sampling and nonsampling errors. A description of the sampling and nonsampling errors associated with the AHS national sample is given below.

Sampling errors. These errors result from the fact that the particular sample used for this survey is only one of a large number of possible samples that could have been selected using the same sample design. Even if all interviewing conditions were the same, estimates from each of the samples would differ from each other. The amount by which the estimates from all possible samples differ from one another is known as the sampling error. The standard error is commonly used to measure sampling error. It indicates how precisely an estimate from a particular sample measures the average result from all possible samples. In addition, the standard error also partially reflects the variation in the estimates due to some nonsampling errors, but it does not measure any systematic biases in the data. The accuracy of the estimates contained in this report depends on the sampling and nonsampling error, as measured by the estimated standard error, and biases and other nonsampling errors not measured by the standard error.

The sample estimate and the estimated standard error permit the construction of intervals such that the average result from all possible samples lies within the interval with a known level of confidence. For example, if all possible samples were selected and surveyed under the same general conditions, and the estimate and estimated standard error were computed for all the samples, then 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 from all possible samples.

For intervals computed using estimates and estimated standard errors from this report, the average result from all possible samples either is or is not contained within the interval. However, it can be said that there is only a 1 in 10 chance that the sample selected will yield a 90-percent confidence interval that does not contain the average result from all possible samples.

The figures presented in the standard error tables are approximations to the standard errors for the estimates in this report. These approximations were necessary in order to produce standard errors applicable to a wide range of characteristics at a reasonable cost. The standard error tables provide an indication of the order of magnitude of the standard errors rather than the actual standard errors for any specific characteristic.

There are various types of estimates that can be made using the data in this report. For example, one can make an estimate of the total number of housing units having a specific characteristic (known as an estimate of a level), a percentage of housing units having a specific characteristic, a ratio of two different characteristics, the difference between two estimates, or medians. Other types of estimates can be made, but these are the most commonly used. Procedures for computing estimated standard errors for these types of estimates are given below.

Standard error table locator. To help identify which standard error table to use for a specific type of estimate from this report, a Standard Error Table Locator is provided. The rows of this table identify the population groups on the boxhead of the tables in this report, and the columns indicate the types of housing characteristics. For example, for general characteristics of the national housing inventory, table 1a should be used for estimating standard errors of estimates of levels; table 1b should be used for estimating standard errors of estimated percentages of these housing units; for fuels and type of heating and cooling equipment in rural areas, table 6a should be used for estimating standard errors of estimates of levels and table 6b should be used for estimating standard errors of estimated percentages of these housing units.

Standard errors of estimates of levels. Tables 1a to 7a present estimated standard errors for estimates of national and regional housing characteristics for 1989. Linear interpolation should be used to determine estimated standard errors for estimates not specifically shown in tables 1a to 7a. The following is an illustration of the use of table 1a.

Table 3-9 of this report shows that in the United States there were 20,751,000 owner-occupied housing units with two persons in 1989. The Standard Error Table Locator shows that table 1a should be used for this type of characteristic. Interpolation in standard error table 1a shows that the estimated standard error of an estimate of this size is 203,000. The following procedure was used in interpolating.

The information in the table below was taken from standard error table 1a. The entry for x is the standard error sought.

Size of estimate (thousands)	Standard error (thousands)
20,000	200
20,751	x
22,500	209

By vertically interpolating between 200,000 and 209,000, "x" is determined to be 203,000.

$$200{,}000 + \frac{20{,}751{,}000 - 20{,}000{,}000}{22{,}500{,}000 - 20{,}000{,}000}(209{,}000 - 200{,}000) = 203{,}000$$

The 90-percent confidence interval for the estimated number of owner-occupied housing units with two persons is from 20,426,000 to 21,076,000. Thus, the average estimate from all possible samples of the these types of housing units will lie within an interval computed in this way for approximately 90 percent of all possible samples.

Standard errors of estimates of percentages. Estimated percentages from this report are computed using sample data for both the numerator and the denominator. The numerator is a subclass of the denominator. The reliability of an estimated percentage depends upon both the size of the percentage and the total upon which the percentage is based (i.e., the denominator). Estimated percentages are more reliable than the corresponding estimates of the numerators of the percentages, particularly if the estimated percentages are 50 percent or more. Tables 1b to 7b present estimated standard errors of national and regional estimated percentages of housing units for 1989. Two-way interpolation should be used for standard errors of estimated percentages not specifically shown in tables 1b to 7b.

Included in tables 1b to 7b are estimated standard errors for estimates of zero percent. These are considered to be overestimates of the true standard error and should be used primarily for the construction of confidence intervals for characteristics when an estimate of zero is obtained. The following is an illustration of the use of table 1b.

Table 3-9 shows that of the 20,751,000 owner-occupied housing units with two persons in the United States in 1989, 15,376,000 or 74.1 percent were in (P)MSA's. The

Standard Error Table Locator shows that table 1b should be used. Interpolation in standard error table 1b (i.e., interpolation on both the denominator and the percent) shows that the standard error of the above percent is 0.5. The following procedure was used in interpolating.

The information in the table below was taken from standard error table 1b. The entry for p is the standard error sought.

Denominator of percent (thousands)	Estimated percent		
	50	74.1	75
20,000	0.6	a	0.5
20,751		p	
22,500	0.5	b	0.5

1. First, interpolate horizontally between 0.6 and 0.5 to get the entry for cell "a." The entry for cell "a" is 0.5.

$$0.6 + \frac{74.1 - 50}{75 - 50}(0.5 - 0.6) = 0.5$$

2. Next, interpolate horizontally between 0.5 and 0.5 to get the entry for cell "b." The entry for cell "b" is 0.5.

$$0.5 + \frac{74.1 - 50}{75 - 50}(0.5 - 0.5) = 0.5$$

3. Finally, interpolate vertically between 0.5 and 0.5 to get the entry for cell "p." The entry for cell "p" is 0.5.

$$0.5 + \frac{20{,}751{,}000 - 20{,}000{,}000}{22{,}500{,}000 - 20{,}000{,}000}(0.5 - 0.5) = 0.5$$

Thus, the 90-percent confidence interval for this estimated percentage is between 73.3 and 74.9 percent.

Standard errors of ratios. For ratios of the form (100) (x/y), where x is not a subclass of y, the standard error tables for estimated percentages 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 the following:

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

where x = numerator of the ratio
y = denominator of the ratio
s_x = estimated standard error of the numerator
s_y = estimated standard error of the denominator

s_x and s_y are computed according to the method used for estimated standard errors of levels. The following is an illustration on how to compute the estimated standard error of a ratio.

Table 4-9 of this report shows that there were 9,357,000 renter-occupied housing units with two persons in the United States in 1989. The estimated standard error of this estimate is determined to be 146,000 using linear interpolation in standard error table 1a. The ratio of owner-occupied, two-person households to renter-occupied, two-person households is 2.22. The estimated standard error of this ratio is 4.10 and is calculated as follows:

$$100\left(\frac{20{,}751{,}000}{9{,}357{,}000}\right)\sqrt{\left(\frac{203{,}000}{20{,}751{,}000}\right)^2 + \left(\frac{146{,}000}{9{,}357{,}000}\right)^2} = 4.10$$

Standard errors of differences. The estimated standard errors shown in tables 1a to 7a are not directly applicable to the difference between two estimates. The estimated standard error of a difference can be computed by the following:

$$s_{x-y} = \sqrt{s_x^2 + s_y^2}$$

where s_x and s_y are the estimated standard errors for the two estimates x and y, respectively. They can be computed in the same manner as for estimated standard errors of levels. This formula is quite accurate for the difference between estimates of the same characteristics in two different areas or the difference between separate and uncorrelated characteristics in the same area. If a high positive correlation exists between the two characteristics, the formula will overestimate the true error. If there is a high negative correlation, the formula will underestimate the true standard error. The following illustration shows how to compute the estimated standard error of a difference.

Table 3-9 shows that in the United States there were 10,915,000 owner-occupied housing units with three persons in 1989. The estimated standard error on this estimate is 157,000. Recall that there were 20,751,000 owner-occupied housing units with two persons in the United States in 1989 with an estimated standard error of 203,000 housing units. The estimated difference between 1989 owner-occupied housing units with two persons and with three persons is 9,836,000, and the estimated standard error of this difference is 257,000 as computed by the following:

$$257{,}000 = \sqrt{(157{,}000)^2 + (203{,}000)^2}$$

The 90-percent confidence interval for the difference of 9,836,000 is from 9,425,000 to 10,247,000 and it can be concluded that the average estimate of this difference, derived from all possible samples, lies within an interval computed in this way for approximately 90 percent of all possible samples.

Standard errors of medians. For medians presented in certain tables in this report, the estimated standard error depends on the distribution of the characteristic and the

total number of housing units that constitute the distribution. A common method for approximating the reliability of the estimated median is to construct an interval about the estimated median such that the average median from all possible samples lies within the interval with a known level of confidence. The following procedure should be used to estimate the upper and lower limits of a 90-percent confidence interval of a median.

1. From the appropriate standard error table for estimated percentages, determine the estimated standard error of a 50-percent characteristic based on the total number of housing units from the distribution.

2. Add to and subtract from 50 percent 1.6 times the estimated standard error determined in step 1 to obtain the upper and lower percentage limits from which the confidence interval will be determined.

3. Determine the lower endpoint of the confidence interval by linearly interpolating within the category of the distribution that contains the lower percentage limit. The upper endpoint of the confidence interval is determined in the same manner using the upper percentage limit.

For about 90 out of 100 possible samples, the average median from all possible samples will lie within this 90-percent confidence interval. The following example illustrates how to compute a 90percent confidence interval for a median.

Table 3-9 of this report shows the median number of persons in owner-occupied housing units was 2.4 in 1989. The total number of housing units upon which the distribution is based is 59,916,000.

1. From table 1b, the standard error of a 50-percent characteristic based on 59,916,000 housing units is 0.4 percentage points.

2. To obtain a 90-percent confidence interval, add to and subtract from 50 percent 1.6 times the estimated standard error from step one giving upper and lower percentage limits of 49.4 and 50.6.

3. From table 3-9, the interval for owner-occupied housing units with two persons (for the purpose of calculating the median, the category of two persons is considered to be from 1.5 to 2.5 persons) contains the 49.4 percent derived in step 2. About 10,536,000 housing units or 17.6 percent fall below this interval, and 20,751,000 housing units or 34.6 percent fall within this interval.

By linear interpolation, the lower limit of the 90-percent confidence interval is found to be about 2.4.

$$1.5 + (2.5 - 1.5)\frac{(49.4 - 17.6)}{(34.6)} = 2.4$$

Similarly, the interval for owner-occupied housing units with two persons contains the 50.6 percent derived in step 2. About 10,536,000 housing units or 17.6 percent fall below this interval, and 20,751,000 housing units or 34.6 percent fall within this interval. The upper limit of the 90-percent confidence interval is found to be about 2.5.

$$1.5 + (2.5 - 1.5)\frac{(50.6 - 17.6)}{(34.6)} = 2.5$$

Thus the 90-percent confidence interval is from 2.4 to 2.5 persons.

Nonsampling errors. Nonsampling errors can be attributed to many sources. The respondent may be unable or unwilling to provide the correct response. The interviewers may be unable to find the unit or they may be unable to obtain information about all the cases. They may record the data incorrectly. Either the respondent or the interviewer may interpret the questions differently than they were intended. The collected data may be keyed incorrectly. The sample frames may be incomplete, introducing some coverage error. Processing of the data introduces errors due to rounding or adjusting for missing values. In addition to these errors, there are other errors of collection, response, processing, coverage, and estimation of missing data. Not all of these errors are unique to sample surveys since they can, and do, occur in complete censuses as well.

Possible effects of decentralized telephone interviewing on the data. The 1989 AHS-National interviews were conducted by decentralized telephone as much as possible, with the exception of cases assigned to the Computer Assisted Telephone Interviewing (CATI) facility. A large-scale decentralized telephone interviewing experiment was conducted in conjunction with the 1983 AHS-National sample in order to provide more definitive information about possible effects of decentralized telephone interviewing on AHS data. It was concluded that telephone interviewing has some effects on the data. The experimental data indicate that compared with personal-visit interviewing, telephone interviewing had the effect of increasing the item nonresponse rate for income items, although this effect does not appear to be causing any changes in the published estimates. There was some tendency to underreport problems with neighborhood quality as well, although this tendency was generally rather slight.

Possible effects of computer assisted telephone interviewing (CATI) on the data. (Note: the design of the CATI experiments is included in the sample design section of this appendix.)

Summary. There is strong evidence that there are differences in data collected by CATI and non-CATI treatments. It is not known for sure, however, which method produces better data. Preliminary analysis of a CATI experiment

conducted in 1987 indicated that CATI had a substantial effect on some characteristics of AHS-N data. Another experiment conducted in 1989 confirmed the results of the 1987 experiment. These findings affect all types of estimates and comparisons. In particular, change estimates across 1985, 1987, and 1989 are biased and longitudinal analysis is quite adversely affected. Further investigation is planned for 1991, but detailed plans have not been determined.

Analysis of results. The same method of analysis was used for both the 1987 and 1989 experiments. Data from the CATI (panels 1 and 5) and non-CATI (panels 4 and 6) treatment panels were weighted separately using the AHS-N estimation procedure described previously. Then estimates from the two treatments were produced in data tables for characteristics provided in Chapter 2 of this publication. Differences between estimates from the CATI and non-CATI treatments were tested using t-statistics.

The 1987 and 1989 analyses of the t-tests yielded similar results. The percents of significant differences observed at the 10-, 5-, and 1-percent significance levels were higher than what would be expected by chance (e.g., we expected that 10 percent of the tests would yield significant results by chance when tested at the s = .10 significance level).

Year	Proportion of significant tests (percent)		
	s = .10	s = .05	s = .01
1987	11.1	6.2	1.9
1989	11.7	6.8	2.3

Both the 1987 and 1989 experiments revealed that the owner, urban, and moderate physical problems subgroups exhibited high numbers of significant differences between CATI and non-CATI estimates, with the moderate physical problems subgroup displaying the highest incidence of significant differences. The 1989 analysis showed significant differences in the total occupied, suburbs (in MSA's), and moved in past year subgroups as well.

The analyses also indicated that CATI had an effect on certain characteristics of the subgroups. These items include lot size, water leakage, income, monthly housing costs as percent of income, housing ownership shared by person not living here, utilities paid separately, owners with a mortgage, and routine maintenance costs. The 1989 analysis indicated differences in heating equipment and other (additional) heating fuels also. CATI estimates tended to be lower than those for non-CATI for four items: water leakage, monthly housing costs as percent of income, other heating fuel, and owners with a mortgage. However, utilities paid separately, income and housing ownership shared by "person not living here" estimates were generally higher for CATI than those for non-CATI. The percent differences between estimates for the panels assigned to the CATI treatment and for the panels assigned to the non-CATI treatment for items in the total occupied subgroup ranged from about 6 percent to about 40 percent. More detailed information on which specific characteristics are affected or the extent of the effect can be obtained by writing to the Statistical Methods Division, Bureau of the Census, Washington, DC 20233.

If indeed CATI itself is the source of these differences, the magnitude of its effect is underestimated by our procedures, as not all cases assigned to CATI were actually interviewed using that methodology. We plan further investigation of the CATI methodology during the 1991 AHS-National interviewing. The nature of that investigation is not yet determined.

Results of a reinterview study conducted in 1989 on 17 items (mainly mortgage and water leakage items) found some weak evidence that CATI may produce more errors than non-CATI. CATI had a higher gross difference rate for 3 of the 17 items tested at the 10-percent significance level. Significant differences between CATI and non-CATI estimates for two of the three items (water leakage and presence of a mortgage) have been cited previously.

Conclusions The 1989 CATI experiment was designed and implemented to determine whether the results from the 1987 experiment, which showed that CATI had a substantial effect on some characteristics of AHS-N data, were valid or due to random variation. Since the 1989 results confirmed the findings from the 1987 study, there is strong evidence that there are differences in data collected in the two treatments. Although very little is known about which treatment provides better data, we speculate that CATI income estimates are probably better than non-CATI, but that most other estimates are probably worse. We base our speculation about income on our assumption that with the computer's assistance, CATI tends to ensure that all questions are asked. We believe that non-CATI estimates for several items are more accurate than CATI estimates because it seems unlikely that people would over- report things like water leaks.

Estimates from both CATI and non-CATI treatments were used to produce the data presented in the 1987 and 1989 publications. As a result, this will have the following impact on these data:

a. The 1987 and 1989 published estimates for the subgroups and items mentioned previously are different than what they would have been if a maximum decentralized telephone interviewing mode had been used for all six panels.

b. There are probably other subgroups and items that were similarly affected but were not detected or included in this preliminary analyses.

c. Cross-sectional comparisons for 1987 and 1989 that involve these items will probably also be affected.

d. Estimates of 1985 to 1987 and 1985 to 1989 change for these items will be biased.

e. The effect on estimates of 1987 to 1989 change for these items should be less since CATI was used in both years.

We encourage you to consider the effects mentioned previously when analyzing the 1987 and 1989 estimates and these change estimates.

Reconciliation Experiment. As part of the 1987 CATI experiment, a reconciliation study was conducted when the responses provided during the CATI interviews for any of the nine selected questions were different from the respective 1985 responses and beyond reasonable tolerance ranges.

Reconciliation questions were then asked immediately following the regular interview to determine whether there had been an actual change since 1985 or whether the 1985 or 1987 responses were wrong.

This reconciliation study indicated that respondents have reporting difficulties with items such as presence of basement, heating equipment, and heating fuel, based on the inconsistent responses provided between 1985 and 1987. These reporting difficulties are not necessarily due to the CATI mode of interviewing, but may reflect general reporting difficulties with select items. This is indicated by the fact that approximately an equal number of respondents stated that their 1985 responses were wrong, when all interviewing was conducted by personal visit, as did the number of respondents who stated that their 1987 responses were wrong. Caution should be taken when carrying out analyses using these data.

A reconciliation study was conducted again in 1989 with six of the nine questions from the 1987 study. The results were similar to those of the 1987 study. Results indicate reporting problems with the presence of a basement, and type of heating equipment items. However, more than half of the respondents indicated that the 1987 response rather than the 1989 response was wrong.

Reinterview Program. The 1989 AHS-N reinterview served as a check for interviewer evaluation and quality control. This check was made at a subsample of the original households to determine if the following was done during the original interview:

a. The correct unit was visited.

b. The correct information on "Tenure" was obtained.

c. The correct information on "Occupancy status" was obtained.

The 1989 reinterview program also served as the means to measure response variance for mortgage items. The results of this analysis were not available at the time of publication. In 1985, a reinterview program was conducted in an attempt to measure some of the nonsampling errors associated with the AHS estimates in addition to serving as an interviewer evaluation and quality control check. This study was conducted using a subsample of the original AHS households. These households were revisited and responses to select questions from the original questionnaire were obtained again. The original interview and the reinterview were assumed to be two independent readings and, thus, were the basis for the measurement of the response error associated with the AHS estimates. The 1985 AHS-N reinterview study was done for three groups of items. They are units in structure and description of structure, number and type of rooms, and appliances, including the age and fuel of the appliances. All items measured showed low levels of inconsistency except those listed in the table below. Included in the table are the levels of inconsistency.

Item	Level of inconsistency	
	Occupied units	Vacant units
Number of living rooms	Moderate	-
Number of dining rooms	Moderate	Moderate
Number of family rooms	Moderate	-
Number of "Other" types of rooms	Moderate	-
Age of refrigerator	-	Moderate
Age of garbage disposal	-	Moderate
Age of oven/cooking burner	-	Moderate
Age of dishwasher	-	Moderate
Age of washing machine	Moderate	-
Central air conditioning fuel	High	-
Cookstove or range with oven	Moderate to High	Moderate

Dashes in the table represent items for which there were not enough observations to compute reliable estimates or items that had low levels of inconsistency. Low levels of inconsistency indicate that the response error is insignificant relative to the standard error in this report. Moderate levels of inconsistency indicate that the response error is not insignificant compared to the standard error in this report. High levels of inconsistency indicate that the response error is very significant compared to the standard error in this report and caution should be used when examining estimates of these characteristics.

Cross-tabulations involving those items that are subject to high levels of inconsistency may also be subject to a large distortion as a consequence and thus, are considered to be less reliable than comparable cross-tabulations that do not involve these data. Since the reinterview programs only measured inconsistencies for a sample of the items on the AHS questionnaire, there may be other items with high levels of inconsistency.

Reinterview studies were also conducted in conjunction with AHS enumerations prior to 1985. These studies included items dealing with poor housing quality, attitudes about the neighborhood, certain housing costs, journey-to-work, and mobility data. The following table shows the items that had moderate or high levels of inconsistency.

While these questions were not included in either the 1985, 1987, or 1989 reinterview studies, questions from previous enumerations were not altered enough to lead one to believe that the level of inconsistent responses would change.

Item	Level of inconsistency
Open cracks or holes on inside of building	Moderate to High
Holes in floors	Moderate to High
Broken plaster or peeling paint on ceilings and walls	High
Mice or rats	Moderate
Working electric outlet in all rooms	High
Concealed wiring	High
Blown fuses/tripped circuit breakers	Moderate to High
Neighborhood conditions: street noise; roads in need of repair; crime; trash, litter, junk in streets or on properties; boarded up/abandoned structures; nonresidential activities; odors, smoke, gas	Moderate to High
Satisfactory neighborhood services: police protection; hospitals/health clinics; public transportation; shopping; elementary schools	Moderate to High
Electricity cost	High
Gas cost	High
Oil, coal, kerosene, wood or other fuel cost	Moderate to High
Fire/hazard insurance	Moderate to High
Real estate taxes	Moderate to High
Cost of real estate taxes	Moderate to High
Cost of water supply and sewage disposal	High
Cost of garbage collection	Moderate to High
Gross income	High
Type of vacant	Moderate to High
Prefer to live in same area or somewhere else	Moderate

A possible explanation for the results of the reinterview studies, as well as the surveys themselves, is that respondents may lack precise information. Also, since the results of the reinterview studies are derived from sample surveys, there is sampling error associated with these estimates of nonsampling error. The possibility of such errors should be taken into account when considering the results of the studies.

Coverage errors. AHS misses approximately 25 percent of the new mobile homes (i.e., those built after January 1, 1980). It is believed that most of the difference is due to poor coverage of new mobile home parks in address ED's.

The coverage of old construction housing units is only as good as the coverage of the 1980 census. The third stage of the ratio estimation procedure attempted to correct for these deficiencies.

Another area of the AHS sample where coverage deficiencies exist is the sampling of building permits to represent conventional (i.e., nonmobile home) new construction. Due to time constraints, only permits issued more than 6 months before interviewing began were eligible to be selected to represent conventional new construction. This is more of a problem for single-unit rather than multiunit structures. In fact, the time lag between issuance of a permit and completion of construction for multiunit structures is generally more than 6 months depending on the size of the structure. Also, new construction in special places such as colleges or military bases is not covered. This is a deficiency in both permit and nonpermit areas.

In identifying whole-structure additions in address and area ED's, units that were in sample were screened to see if they were eligible for interview. The screening operation involved asking a series of questions. Therefore, the quality of coverage in these areas is only as good as the quality of the responses to these questions. It is conceivable that eligible units were omitted and ineligible units were included because the respondents' answers to the screening questions were incorrect. In addition, the quality of the listing of addresses will also affect the coverage of whole-structure additions.

It is also believed that a coverage deficiency exists for units that were nonresidential at the time of the 1980 census, but have since converted to residential units. The magnitude of this deficiency is not known.

The second and third stages of ratio estimation correct these deficiencies for the total number of housing units only. Biases of subtotals will still exist.

Processing errors. Several types of errors are associated with the processing of the data. The first type of processing error that may be introduced is keying error. A quality assurance operation conducted in conjunction with the keying of the data helps to insure that less than 0.4 percent of the data fields keyed from the questionnaire will be in error.

Another type of processing error is imputation error. If certain fields on a questionnaire are blank, values are assigned by the computer. These are generally items for which 1980 census data is available, as well as items that had an item nonresponse rate of 1.0 percent or less in 1983. It is not known how close these imputed values are to the actual values.

A problem may also exist for items for which there are no imputations for items of nonresponse. Totals for these items and any subcategories of these items may be underestimated. Percent distributions may also be distorted.

Nonsampling error also occurs because of noninterview. The noninterview adjustments assume that interviewed units of similar size and geographic location (i.e.,[P]MSA status, urban/rural status) can adequately represent noninterviews. The extent to which this assumption does not hold true will determine the magnitude of the nonsampling error from these units. Finally, another type of processing error is rounding error. The data are processed using double precision to minimize the effect of the rounding

errors. However, the error may still be significant for small percentages and small medians when these figures are derived from relatively large bases. Thus, confidence intervals formed from the standard errors may be distorted. This should be taken into consideration when analyzing the results of this survey.

Standard Error Table Locator: Population Group by Type of Characteristic

(Tables "a" used for estimates; tables "b" used for percentages)

Population group[1]	Table number by characteristics group			
	General[2]	Fuel and type of heating/cooling equipment	Neighborhood[3]	Special[4]
United States:				
Total[5]	1a, 1b	5a, 5b	5a, 5b	6a, 6b
Year-round or seasonal vacants	4a, 4b	5a, 5b	5a, 5b	6a, 6b
Black	1a, 1b	5a, 5b	5a, 5b	6a, 6b
Hispanic	[6]2a, [6]2b	5a, 5b	5a, 5b	6a, 6b
Elderly	1a, 1b	5a, 5b	5a, 5b	6a, 6b
Urban	2a, 2b	5a, 5b	5a, 5b	6a, 6b
Rural	3a, 3b	6a, 6b	5a, 5b	6a, 6b
Mobile home	1a, 1b	6a, 6b	5a, 5b	6a, 6b
New construction	1a, 1b	5a, 5b	5a, 5b	6a, 6b
In (P)MSA's—Central Cities	2a, 2b	5a, 5b	5a, 5b	6a, 6b
In (P)MSA's—Suburbs	2a, 2b	5a, 5b	5a, 5b	6a, 6b
Outside (P)MSA's	4a, 4b	7a, 7b	7a, 7b	7a, 7b
Regions:				
Northeast	2a, 2b	5a, 5b	5a, 5b	6a, 6b
Midwest	1a, 1b	5a, 5b	5a, 5b	6a, 6b
South	3a, 3b	6a, 6b	5a, 5b	6a, 6b
West	1a, 1b	5a, 5b	5a, 5b	6a, 6b

[1]For multiple population groups (for example; Blacks in the Northeast or new construction in central cities) use the standard error table with the highest standard error for a given estimate.

[2]General includes all characteristics except fuels and heating/cooling equipment, neighborhood items, and special items.

[3]Neighborhood items include all characteristics in "neighborhood" tables except "mobile home in group."

[4]Special items include all characteristics pertaining to cooperatives or condominiums; no complete bathroom; less than 1,500 square feet of detached one-family or mobile homes; well serving 1 to 5 units; mobile homes in a group of seven or more; area within 300 feet includes open space, park, farm or ranch; and major street repairs needed.

[5]Total includes total housing units, year-round, occupied, owner, renter, physical problems, moved in past year, below poverty level.

[6]Use table 1 for the following Hispanic deficiency items: sagging roof; missing bricks, siding, and other outside material; broken windows; fuel other than electricity, gas, or oil; bars on windows of buildings within 300 feet; 1.51 or more persons per room; 400 to 699 square feet per person; water supply stoppage in last 3 months; no toilet working for at least 6 hours in last 3 months; sewage disposal—public sewer with breakdown lasting 6 hours or more in last 3 months; uncomfortably cold for 24 or more hours last winter; signs of rats in last 3 months; and broken plaster or peeling paint in interior.

Table 1a. **Standard Errors of Estimated Numbers of Housing Units**

(Numbers in thousands)

Size of estimate	Standard error				Size of estimate	Standard error			
	United States, elderly, new construction, or mobile home	Midwest region	West region	Black		United States, elderly, new construction, or mobile home	Midwest region	West region	Black
0	3	3	3	3	10,633	155	120	109	0
5	4	4	4	4	12,500	166	120	104	-
10	5	5	5	5	15,000	179	115	90	-
25	8	8	8	8	17,500	190	102	59	-
50	11	11	11	11	18,996	196	90	-	-
100	16	16	16	16	20,000	200	80	-	-
250	25	25	25	25	22,500	209	30	-	-
500	36	35	35	35	22,869	210	-	-	-
1,000	50	49	49	48	25,000	216	-	-	-
2,500	79	75	74	70	50,000	244	-	-	-
5,000	110	100	97	82	75,000	195	-	-	-
7,500	132	113	107	75	90,000	95	-	-	-
10,000	151	120	110	39	93,683	-	-	-	-

Table 1b. **Standard Errors of Estimated Percentages of Housing Units**

Base of percentage (thousands)	Estimated percentage							
	0 or 100	1 or 99	2 or 98	5 or 95	10 or 90	15 or 85	25 or 75	50
5	33.7	33.7	33.7	33.7	33.7	33.7	33.7	35.7
10	20.3	20.3	20.3	20.3	20.3	20.3	21.8	25.2
25	9.2	9.2	9.2	9.2	9.6	11.4	13.8	15.9
50	4.8	4.8	4.8	4.9	6.8	8.1	9.8	11.3
100	2.5	2.5	2.5	3.5	4.8	5.7	6.9	8.0
250	1.0	1.0	1.4	2.2	3.0	3.6	4.4	5.0
500	0.5	0.7	1.0	1.6	2.1	2.5	3.1	3.6
1,000	0.3	0.5	0.7	1.1	1.5	1.8	2.2	2.5
2,500	0.10	0.3	0.4	0.7	1.0	1.1	1.4	1.6
5,000	0.05	0.2	0.3	0.5	0.7	0.8	1.0	1.1
7,500	0.03	0.2	0.3	0.4	0.6	0.7	0.8	0.9
10,000	0.03	0.2	0.2	0.3	0.5	0.6	0.7	0.8
12,500	0.02	0.14	0.2	0.3	0.4	0.5	0.6	0.7
15,000	0.02	0.13	0.2	0.3	0.4	0.5	0.6	0.7
17,500	0.01	0.12	0.2	0.3	0.4	0.4	0.5	0.6
20,000	0.01	0.11	0.2	0.2	0.3	0.4	0.5	0.6
22,500	0.01	0.11	0.1	0.2	0.3	0.4	0.5	0.5
25,000	0.01	0.10	0.14	0.2	0.3	0.4	0.4	0.5
50,000	0.01	0.07	0.10	0.2	0.2	0.3	0.3	0.4
75,000	0.01	0.06	0.08	0.13	0.2	0.2	0.3	0.3
90,000	0.01	0.05	0.07	0.12	0.2	0.2	0.2	0.3
93,683	0.01	0.05	0.07	0.11	0.2	0.2	0.2	0.3

Table 2a. **Standard Errors of Estimated Numbers of Housing Units**

(Numbers in thousands)

Size of estimate	Standard error		Size of estimate	Standard error	
	Urban, central cities, MSA-suburb, or Hispanic (deficiency)[1]	Northeast region		Urban, central cities, MSA-suburb, or Hispanic (deficiency)[1]	Northeast region
0	2	2	5,000	104	92
5	3	3	7,500	125	102
10	5	5	10,000	142	105
25	8	8	15,000	169	88
50	11	11	19,389	187	0
100	15	15	20,000	189	-
250	24	24	25,000	204	-
500	34	33	50,000	230	-
1,000	47	46	75,000	184	-
2,500	74	70			

[1]Use standard errors as displayed in the table for estimates pertaining to the following Hispanic deficiency items: sagging roof; missing bricks, siding, and other outside material; broken windows; fuel other than electricity, gas, or oil; bars on windows of buildings within 300 feet; 1.51 or more persons per room; 400 to 699 square feet per person; water supply stoppage in last 3 months; no toilet working for at least 6 hours in last 3 months; sewage disposal—public sewer with breakdown lasting 6 hours or more in last 3 months; uncomfortably cold for 24 or more hours last winter; signs of rats in last 3 months; and broken plaster or peeling paint in interior.

Multiply standard errors by a factor of 0.92 for estimates pertaining to Hispanic general items.

Table 2b. **Standard Errors of Estimated Percentages of Housing Units**

Base of percentage (thousands)	Estimated percentage[1]							
	0 or 100	1 or 99	2 or 98	5 or 95	10 or 90	15 or 85	25 or 75	50
5	31.2	31.2	31.2	31.2	31.2	31.2	31.2	33.7
10	18.5	18.5	18.5	18.5	18.5	18.5	20.6	23.8
25	8.3	8.3	8.3	8.3	9.0	10.8	13.0	15.1
50	4.3	4.3	4.3	4.6	6.4	7.6	9.2	10.6
100	2.2	2.2	2.2	3.3	4.5	5.4	6.5	7.5
250	0.9	0.9	1.3	2.1	2.9	3.4	4.1	4.8
500	0.5	0.7	0.9	1.5	2.0	2.4	2.9	3.4
1,000	0.2	0.5	0.7	1.0	1.4	1.7	2.1	2.4
2,500	0.09	0.3	0.4	0.7	0.9	1.1	1.3	1.5
5,000	0.05	0.2	0.3	0.5	0.6	0.8	0.9	1.1
7,500	0.03	0.2	0.2	0.4	0.5	0.6	0.8	0.9
10,000	0.02	0.1	0.2	0.3	0.5	0.5	0.7	0.8
15,000	0.02	0.12	0.2	0.3	0.4	0.4	0.5	0.6
20,000	0.01	0.11	0.1	0.2	0.3	0.4	0.5	0.5
25,000	0.01	0.09	0.13	0.2	0.3	0.3	0.4	0.5
50,000	0.01	0.07	0.09	0.15	0.2	0.2	0.3	0.3
75,000	0.01	0.05	0.08	0.12	0.2	0.2	0.2	0.3

[1]Use standard errors as displayed in the table for estimates pertaining to the following Hispanic deficiency items: sagging roof; missing bricks, siding, and other outside material; broken windows; fuel other than electricity, gas, or oil; bars on windows of buildings within 300 feet; 1.51 or more persons per room; 400 to 699 square feet per person; water supply stoppage in last 3 months; no toilet working for at least 6 hours in last 3 months; sewage disposal—public sewer with breakdown lasting 6 hours or more in last 3 months; uncomfortably cold for 24 hours or more last winter; signs of rats in last 3 months; and broken plaster or peeling paint in interior.

Multiply standard errors by a factor of 0.92 for estimates pertaining to Hispanic general items.

TABLE 3a. **Standard Errors of Estimated Numbers of Housing Units**

(Numbers in thousands)

Size of estimate	Standard error		Size of estimate	Standard error	
	Rural	South region		Rural	South region
0	3	3	2,500	81	79
5	4	4	5,000	113	107
10	5	5	7,500	137	125
25	8	8	10,000	155	137
50	12	12	15,000	185	148
100	16	16	20,000	206	144
250	26	26	25,000	223	124
500	37	36	30,000	235	78
1,000	52	51	33,000	241	-

TABLE 3b. **Standard Errors of Estimated Percentages of Housing Units**

Base of percentage (thousands)	Estimated percentage							
	0 or 100	1 or 99	2 or 98	5 or 95	10 or 90	15 or 85	25 or 75	50
5	35.1	35.1	35.1	35.1	35.1	35.1	35.1	36.8
10	21.3	21.3	21.3	21.3	21.3	21.3	22.5	26.0
25	9.8	9.8	9.8	9.8	9.9	11.7	14.2	16.4
50	5.1	5.1	5.1	5.1	7.0	8.3	10.1	11.6
100	2.6	2.6	2.6	3.6	4.9	5.9	7.1	8.2
250	1.1	1.1	1.5	2.3	3.1	3.7	4.5	5.2
500	0.5	0.7	1.0	1.6	2.2	2.6	3.2	3.7
1,000	0.3	0.5	0.7	1.1	1.6	1.9	2.3	2.6
2,500	0.11	0.3	0.5	0.7	1.0	1.2	1.4	1.6
5,000	0.05	0.2	0.3	0.5	0.7	0.8	1.0	1.2
7,500	0.04	0.2	0.3	0.4	0.6	0.7	0.8	0.9
10,000	0.03	0.2	0.2	0.4	0.5	0.6	0.7	0.8
15,000	0.02	0.13	0.2	0.3	0.4	0.5	0.6	0.7
20,000	0.01	0.12	0.2	0.3	0.3	0.4	0.5	0.6
25,000	0.01	0.10	0.15	0.2	0.3	0.4	0.5	0.5
30,000	0.01	0.09	0.13	0.2	0.3	0.3	0.4	0.5
33,000	0.01	0.09	0.13	0.2	0.3	0.3	0.4	0.5

Table 4a. **Standard Errors of Estimated Numbers of Housing Units**

(Numbers in thousands)

Size of estimate	Standard error	Size of estimate	Standard error
0	3	5,000	204
5	4	7,500	288
10	5	10,000	372
25	8	12,500	456
50	12	15,000	539
100	17	17,500	623
250	28	20,000	706
500	41	22,500	790
1,000	62	25,000	873
2,500	118		

TABLE 4b. **Standard Errors of Estimated Percentages of Housing Units**

Base of percentage (thousands)	Estimated percentage							
	0 or 100	1 or 99	2 or 98	5 or 95	10 or 90	15 or 85	25 or 75	50
5	35.7	35.7	35.7	35.7	35.7	35.7	35.7	37.2
10	21.7	21.7	21.7	21.7	21.7	21.7	22.8	26.3
25	10.0	10.0	10.0	10.0	10.0	11.9	14.4	16.7
50	5.3	5.3	5.3	5.3	7.1	8.4	10.2	11.8
100	2.7	2.7	2.7	3.6	5.0	5.9	7.2	8.3
250	1.1	1.1	1.5	2.3	3.2	3.8	4.6	5.3
500	0.6	0.7	1.0	1.6	2.2	2.7	3.2	3.7
1,000	0.3	0.5	0.7	1.1	1.6	1.9	2.3	2.6
2,500	0.11	0.3	0.5	0.7	1.0	1.2	1.4	1.7
5,000	0.06	0.2	0.3	0.5	0.7	0.8	1.0	1.2
7,500	0.04	0.2	0.3	0.4	0.6	0.7	0.8	1.0
10,000	0.03	0.2	0.2	0.4	0.5	0.6	0.7	0.8
12,500	0.02	0.15	0.2	0.3	0.4	0.5	0.6	0.7
15,000	0.02	0.14	0.2	0.3	0.4	0.5	0.6	0.7
17,500	0.02	0.13	0.2	0.3	0.4	0.4	0.5	0.6
20,000	0.01	0.12	0.2	0.3	0.4	0.4	0.5	0.6
22,500	0.01	0.11	0.2	0.2	0.3	0.4	0.5	0.6
25,000	0.01	0.10	0.15	0.2	0.3	0.4	0.5	0.5

TABLE 5a. **Standard Errors of Estimated Numbers of Housing Units**

(Numbers in thousands)

Size of estimate	Standard error	Size of estimate	Standard error
0	3	10,000	168
5	4	12,500	185
10	6	15,000	199
25	9	17,500	211
50	13	20,000	222
100	18	22,500	232
250	28	25,000	240
500	40	50,000	271
1,000	56	75,000	217
2,500	87	90,000	105
5,000	122	93,683	0
7,500	147		

[1]Multiply standard errors by a factor of 0.93 for estimates pertaining to Hispanic items.

TABLE 5b. **Standard Errors of Estimated Percentages of Housing Units**

Base of percentage (thousands)	Estimated percentage							
	0 or 100	1 or 99	2 or 98	5 or 95	10 or 90	15 or 85	25 or 75	50
5	38.6	38.6	38.6	38.6	38.6	38.6	38.6	39.6
10	23.9	23.9	23.9	23.9	23.9	23.9	24.3	28.0
25	11.2	11.2	11.2	11.2	11.2	12.7	15.4	17.7
50	5.9	5.9	5.9	5.9	7.5	9.0	10.9	12.5
100	3.0	3.0	3.0	3.9	5.3	6.3	7.7	8.9
250	1.2	1.2	1.6	2.4	3.4	4.0	4.9	5.6
500	0.6	0.8	1.1	1.7	2.4	2.8	3.4	4.0
1,000	0.3	0.6	0.8	1.2	1.7	2.0	2.4	2.8
2,500	0.13	0.4	0.5	0.8	1.1	1.3	1.5	1.8
5,000	0.06	0.2	0.4	0.5	0.8	0.9	1.1	1.3
7,500	0.04	0.2	0.3	0.4	0.6	0.7	0.9	1.0
10,000	0.03	0.2	0.2	0.4	0.5	0.6	0.8	0.9
12,500	0.03	0.2	0.2	0.3	0.5	0.6	0.7	0.8
15,000	0.02	0.14	0.2	0.3	0.4	0.5	0.6	0.7
17,500	0.02	0.13	0.2	0.3	0.4	0.5	0.6	0.7
20,000	0.02	0.12	0.2	0.3	0.4	0.4	0.5	0.6
22,500	0.01	0.12	0.2	0.3	0.4	0.4	0.5	0.6
25,000	0.01	0.11	0.2	0.2	0.3	0.4	0.5	0.6
50,000	0.01	0.08	0.11	0.2	0.2	0.3	0.3	0.4
75,000	0.01	0.06	0.09	0.14	0.2	0.2	0.3	0.3
90,000	0.01	0.06	0.08	0.13	0.2	0.2	0.3	0.3
93,683	0.01	0.06	0.08	0.13	0.2	0.2	0.3	0.3

[1]Multiply standard errors by a factor of 0.93 for estimates pertaining to Hispanic items.

Table 6a. Standard Errors of Estimated Numbers of Housing Units

(Numbers in thousands)

Size of estimate	Standard error	Size of estimate	Standard error
0	5	7,500	192
5	6	10,000	218
10	7	12,500	240
25	12	15,000	259
50	16	17,500	276
100	23	20,000	290
250	36	22,500	302
500	52	25,000	313
1,000	73	30,000	330
2,500	114	35,000	342
5,000	159	40,000	350

[1]Multiply standard errors by a factor of 0.93 for estimates pertaining to Hispanic items.

TABLE 6b. Standard Errors of Estimated Percentages of Housing Units

Base of percentage (thousands)	Estimated percentage							
	0 or 100	1 or 99	2 or 98	5 or 95	10 or 90	15 or 85	25 or 75	50
5	51.6	51.6	51.6	51.6	51.6	51.6	51.6	51.7
10	34.8	34.8	34.8	34.8	34.8	34.8	34.8	36.5
25	17.6	17.6	17.6	17.6	17.6	17.6	20.0	23.1
50	9.6	9.6	9.6	9.6	9.8	11.7	14.1	16.3
100	5.1	5.1	5.1	5.1	6.9	8.3	10.0	11.6
250	2.1	2.1	2.1	3.2	4.4	5.2	6.3	7.3
500	1.1	1.1	1.4	2.3	3.1	3.7	4.5	5.2
1,000	0.5	0.7	1.0	1.6	2.2	2.6	3.2	3.7
2,500	0.2	0.5	0.6	1.0	1.4	1.7	2.0	2.3
5,000	0.11	0.3	0.5	0.7	1.0	1.2	1.4	1.6
7,500	0.07	0.3	0.4	0.6	0.8	1.0	1.2	1.3
10,000	0.05	0.2	0.3	0.5	0.7	0.8	1.0	1.2
12,500	0.04	0.2	0.3	0.5	0.6	0.7	0.9	1.0
15,000	0.04	0.2	0.3	0.4	0.6	0.7	0.8	0.9
17,500	0.03	0.2	0.2	0.4	0.5	0.6	0.8	0.9
20,000	0.03	0.2	0.2	0.4	0.5	0.6	0.7	0.8
22,500	0.02	0.2	0.2	0.3	0.5	0.6	0.7	0.8
25,000	0.02	0.15	0.2	0.3	0.4	0.5	0.6	0.7
30,000	0.02	0.13	0.2	0.3	0.4	0.5	0.6	0.7
35,000	0.02	0.12	0.2	0.3	0.4	0.4	0.5	0.6
40,000	0.01	0.11	0.2	0.3	0.3	0.4	0.5	0.6

[1]Multiply standard errors by a factor of 0.93 for estimates pertaining to Hispanic items.

TABLE 7a. Standard Errors of Estimated Numbers of Housing Units

(Numbers in thousands)

Size of estimate	Standard error	Size of estimate	Standard error
0	8	5,000	283
5	8	7,500	385
10	9	10,000	486
25	14	12,500	586
50	20	15,000	685
100	29	17,500	784
250	47	20,000	883
500	67	22,500	981
1,000	99	25,000	1080
2,500	174		

TABLE 7b. Standard Errors of Estimated Percentages of Housing Units

Base of percentage (thousands)	Estimated percentage							
	0 or 100	1 or 99	2 or 98	5 or 95	10 or 90	15 or 85	25 or 75	50
5	62.5	62.5	62.5	62.5	62.5	62.5	62.5	64.5
10	45.4	45.4	45.4	45.4	45.4	45.4	45.4	45.6
25	25.0	25.0	25.0	25.0	25.0	25.0	25.0	28.8
50	14.3	14.3	14.3	14.3	14.3	14.6	17.7	20.4
100	7.7	7.7	7.7	7.7	8.7	10.3	12.5	14.4
250	3.2	3.2	3.2	4.0	5.5	6.5	7.9	9.1
500	1.6	1.6	1.8	2.8	3.9	4.6	5.6	6.5
1,000	0.8	0.9	1.3	2.0	2.7	3.3	4.0	4.6
2,500	0.3	0.6	0.8	1.3	1.7	2.1	2.5	2.9
5,000	0.2	0.4	0.6	0.9	1.2	1.5	1.8	2.0
7,500	0.11	0.3	0.5	0.7	1.0	1.2	1.4	1.7
10,000	0.08	0.3	0.4	0.6	0.9	1.0	1.2	1.4
12,500	0.07	0.3	0.4	0.6	0.8	0.9	1.1	1.3
15,000	0.06	0.2	0.3	0.5	0.7	0.8	1.0	1.2
17,500	0.05	0.2	0.3	0.5	0.7	0.8	0.9	1.1
20,000	0.04	0.2	0.3	0.4	0.6	0.7	0.9	1.0
22,500	0.04	0.2	0.3	0.4	0.6	0.7	0.8	1.0
25,000	0.03	0.2	0.3	0.4	0.5	0.7	0.8	0.9