Source and Accuracy Statement for the
1989 Long Term Care Survey
(Prepared by the Bureau of the Census, Feb., 1992)

SOURCE OF DATA

The data for the 1989 Long Term Care Survey (LTC) were obtained during the period February 27, 1989 through July 14, 1989 by interviews from the Bureau of the Census. The 1989 LTC was designed to collect data on the health of the population 65 years and over. The survey contacted persons from the 1982 and/or 1984 LTC, and a sample of additional persons who had reached age 65 between April 1984 and April 1988. Each sample person who was defined as unimpaired in the 1984 LTC survey or newly selected for the 1989 survey was asked a series of screening questions designed to identify persons having certain disabilities or other health problems lasting three months or longer at the time of the 1989 survey. The identified persons and all other sample persons (namely those impaired or institutionalized in 1984) were designated for detailed 1989 interviews regarding their impairments, treatment, personal care, financial arrangements for care and treatment and other subjects. Those persons impaired or institutionalized in 1984 were only given a partial screener instead of a full screener to update their address and determine if they were still in scope (not deceased, in jail, or living overseas).

There were differences in the definition of persons in scope for the 1982, 1984, and 1989 LTC. For the 1982 LTC data was collected primarily on the health of the noninstitutionalized population and persons in institutions or deceased before April 1, 1982 were defined as out of scope. For the 1984 LTC data were collected on persons in institutions and deceased in addition to persons in the noninstitutional population. Persons in the noninstitutional population, persons in institutions before or on April 1, 1984 and persons in scope for the 1982 survey but deceased

before April 1, 1984 were defined as in scope of the 1984 LTC. For the 1989 LTC data were collected on the health of the institutional and noninstitutional population but data were not collected for deceased persons. However, the 1989 LTC also included a survey of caregivers. Persons in the noninstitutional population with disabilities and persons in institutions as of the date of their screener interview were defined as in scope for the 1989 LTC. Unpaid helpers or relatives who were paid or unpaid were defined as in scope for the 1989 caregiver survey.

THE SAMPLE

The 1989 LTC sample was primarily designed to make cross-sectional estimates. It consisted of 17,565 medicare enrollees who were selected to make 1989 cross-sectional estimates. Of these, 12,658 were previously selected for the 1984 survey. This included 651 persons who were institutionalized between April 1, 1984 and the time they were reached for the 1989 survey, 4990 medicare enrollees selected from the 14,819 medicare enrollees defined as unimpaired for the 1984 survey, an additional 3089 persons defined as unimpaired in the 1984 survey and 75 years and older on April 1, 1988 who were not part of the 4990 selected, and all 3,928 persons defined as impaired for the 1984 survey. The rest of the 1989 LTC sample included 4,907 medicare enrollees who turned 65 on or after April 1, 1984 but before or on April 1, 1988. These 4,907 medicare enrollees were selected through a series of four steps described below.

STEPS IN THE SELECTION OF THE 4,907 NEW MEDICARE ENROLLEES

(These steps are very similar to those for the 1982 and 1984 LTC sample, which are detailed in the October 4, 1984 memorandum from Jones to Walsh entitled "Source and Reliability

Statement for Long Term Care Survey" and the August 11, 1986 memorandum from Jones to Walsh entitled "Source and Reliability Statement for the 1984 Long Term Care Survey.")

1. Selection of Medicare Enrollees

The 4,907 medicare enrollees who turned 65 on or after April 1, 1984 but before or on April 1, 1988 (including institutionalized and noninstitutionalized persons of age 65 and over) were selected from files maintained by the Health Care Financing Administration. These files covered most of the desired population; undercoverage is discussed below. In most areas an initial 10 percent sample of medicare enrollees who turned 65 on or after April 1, 1984 but before or on April 1, 1988 and who were alive as of March 1, 1988 was selected from a March 1988 file of medicare enrollees. An initial 50 percent sample was selected in areas thought to require a sampling fraction greater than 10 percent. The initial sample described here was selected in order to reduce the amount of processing required during the rest of the sample selection.

2. Selection of A-sample PSUs

To reduce costs, the LTC sample was clustered in primary sampling units (PSUs) that consisted of individual counties or groups of adjoining counties. The PSUs for the 1989 sample were the same ones used for the 1984 sample. The sample of PSUs for the LTC was selected from the Census Bureau's A-sample of PSUs. The A-sample of PSUs was selected for use in a number of the Census Bureau's ongoing current surveys, including the Current Population Survey and the Health Interview Survey. Selection of the A-sample was as follows.

Prior to selecting the A-sample, the entire land area of the United States was divided into 1,924 PSUs. The PSUs were grouped into 376 strata based on criteria including geographic region, level of urbanization, percentage of population that is nonwhite, and per capita retail sales. One PSU was selected from each stratum. A PSU's selection probability was equal to its 1970 population as a proportion of the total 1970 population in its stratum.

One hundred fifty-six strata consist of but one PSU; these are called self-representing (SR) PSUs. Each of the remaining 220 strata contain a number of nonself-representing (NSR) PSUs. As mentioned above, one sample NSR PSU was selected from each of these 220 strata.

3. Selection of LTC PSUs

Before a subsample of A-sample PSUs was selected for the LTC, some A-sample PSUs were combined to form LTC PSUs. This was done so that as many LTC PSUs as possible consisted of whole counties or groups of whole counties, thus facilitating the use of Medicare files, which have counties as their most detailed level of geography. Most LTC PSUs consisted of just one A-sample PSU.

For the LTC, A-sample strata were collapsed to form 173 LTC strata. Thirty-nine of the strata consisted of only one A-sample PSU. Each had a 1978 Medicare population of 87,000 or greater. These are called LTC SR PSUs. The remaining LTC PSUs were grouped to form 134 strata according to their proportions of population 65 years and older and enrolled in Medicare.

Each of these 134 strata contained an estimated Medicare population of about 90,000. A stratum's Medicare population was estimated by inflating its A-sample PSUs' Medicare populations by the inverse of their probabilities of selection for the A-sample and summing those estimates. One LTC PSU was selected from each stratum with probability proportional to its A-sample stratum's estimated 1978 Medicare population. These sample PSUs are called LTC NSR PSUs.

4. Selection of the 4,907 Medicare enrollees from the LTC PSUs

Once the LTC sample PSUs were selected, the initial sample of Medicare enrollees was reduced by eliminating all those who lived outside the 173 LTC sample PSUs. The file was sorted into strata defined by the original reason for Medicare entitlement (age, disability). Both strata were subsampled at the same rate, ensuring proportional representation of persons by original reason for entitlement.

Within each sample PSU, the sampling rate was inversely proportional to the LTC PSUs' probability of selection. That is, the product of the within-PSU probability of selection and the probability of the PSUs being selected was equal to a constant. The constant is the overall probability of selection which is the same for each sample person.

THE SELECTION OF THE 8,079 UNIMPAIRED FROM THE 14,819 UNIMPAIRED

First a systematic sample of 4,990 medicare enrollees were selected from the 14,819 medicare enrollees defined as unimpaired in 1984. They were selected across all ages using a sampling interval of 2.9638. Then all other medicare enrollees defined as unimpaired in 1984 and 75 years

and older as of April 1, 1988 were selected to be in sample. This resulted in the selection of an additional 3,089 medicare enrollees defined as unimpaired in 1984.

ESTIMATION

1989 Cross-Sectional Estimation

The estimation procedure for 1989 cross-sectional estimates was essentially the same as the 1984 estimation procedure. However, a few changes were made involving the definition of age for 65 year olds and the calculation of the second stage factor. The estimation procedure involves several stages of weight adjustments. Each interviewed person received a screener weight and if the person qualified for a detailed interview he/she also received a detailed weight.

SCREENER WEIGHTS

Each interviewed person received two final screener weights equal to the product of the following components:

- base weight
- screener noninterview adjustment factor
- first-stage ratio estimate factor
- second-stage ratio estimate (original or new formula)
- (post-stratification) factor

The difference in the two final screener weights depends on which second stage factor is used.

The final screener weights should be used for analyses of data contained on the screener, when all persons with completed screeners are used as interviews. The base weight is the inverse of a person's probability of selection. Except for persons unimpaired in 1982 and/or 1984 and the 4,907 new medicare enrollees, persons were assigned an equal base weight of 730.5909. Persons who were unimpaired in 1982 received an additional weight of 97/44 to compensate for subsampling the 1982 unimpaired for the 1984 survey, and persons unimpaired in 1984 received an additional weight of 2.9638 to compensate for subsampling the 1984 unimpaired for the 1989 survey. Thus, persons in the 1989 survey who were unimpaired in 1982 and 1984 received additional weights of 97/44 and 2.9638. The 4,907 new medicare enrollees who aged in to the 1989 survey received a basic weight of 1505.836. Persons with a caregiver interview received a weight of 2.62 in addition to the weight already assigned to their associated impaired person.

Two broad classes of screener noninterviews were encountered during the 1989 LTC: Type C noninterviews and type A noninterviews. Type C noninterviews were persons who did not belong in the LTC universe and were therefore omitted with no adjustment. The sampling frame consisted of persons in the scope of the 1989 LTC on the date of sample selection (April 1, 1988).

To minimize the effect of panel attrition, especially due to death, on estimates of impairment status, age and eligibility status should be determined as of the approximate date of interview (which is when impairment status is determined) rather than the sample selection date of April 1, 1988. Thus, persons who are impaired on April 1, 1988 who die before their interview (treated as Type C's) are balanced by persons who are unimpaired on April 1, 1988 and become impaired

by the time of interview, to give an approximate representation of the population at the time of interview. Some error is introduced because most screening interviews take place before the majority of the detailed interviews. However, this lag time between the screener and detailed interview is less than in previous LTC surveys. To compensate for this partially, impaired persons who die between their screener interview and their detailed interview are treated as Type A noninterviews for the detailed questionnaire, so that other interviewed persons who also screened in are given higher weight because of them.

Type A noninterviews were persons who belonged in the LTC universe and thus were adjusted for in the weighting by giving higher weight to similar interviewed individuals. Specifically, type A noninterviews were, eligible persons who refused to be interviewed, could not be contacted or located, or could not respond and no proxy was available.

Each screened person's weight was adjusted to account for type A noninterviews. Noninterview adjustment was done by cells defined by age, original reason for Medicare entitlement, 1984 interview status group and LTC PSU. The 1984 interview status groups consisted of noninstitutionalized persons defined as impaired in 1984, noninstitutionalized persons defined as unimpaired and persons institutionalized as of April 1, 1984. Cells were often collapsed over a number of LTC PSUs to provide enough cases to yield reliable noninterview adjustment factors but institutional and noninstitutional persons were never combined.

A first-stage ratio estimate was applied to the weight of each person in an LTC NSR PSU to account for the LTC NSR PSUs' not having population distributions identical to those of the

strata from which they were selected. This adjustment provides lower variance of the estimates. Factors were applied according to census-region (Northeast, North Central, South, West), whether a person resided in a Standard Metropolitan Statistical Area (SMSA) and age.

Finally, each person received two alternative second-stage ratio estimate factors: an original second-stage factor and a new formulae second-stage factor. The original second-stage factor is defined and calculated the same way as in 1982 and 1984 and should be used when making comparisons with 1982 and 1984 estimates. The new second-stage factor was designed to make the best 1989 estimate, and to improve agreement with external data sources. It should be used when making the best estimate for 1989. The difference between the two factors depends on the way an institutionalized person is defined. For the new formula second-stage factor institutionalized persons are defined to better agree with the independent estimates being used for institutional persons. Both factors cause the weighted number of sample persons to agree with independent estimates of the residential population of the United States on April 1, 1989 by age, race, sex and institutional vs. noninstitutional status. These independent estimates are based on population controls from the decennial censuses and statistics on births, deaths, immigration, and emigration. The independent estimates used for this survey were derived from the 1980 Census of Population and were adjusted for population changes between Census Day (April 1, 1980) and April 1, 1989.

Because of the lag between the sample selection (April 1, 1988) and interviewing (screening started in February 1989, with detailed interviewing ending in July 1989), most of the detailed interviews with impaired persons took place in early 1989. Therefore, April 1, 1989 population

controls were used for the second-stage adjustment. --In the adjustments, persons who were 65 years of age on October 1, 1988 were included as 65 year-olds. Persons aged 66 years of age as of October 1, 1988 were included as 66-year-olds. All other persons were assigned their age as of April 1, 1989.

Using these dates in calculating the factors makes the weighted results as representative of the population aged 65 and older as possible. Note that in analyzing the data, persons should be assigned their age as of their individual date of interview.

1989 DETAILED WEIGHTS

Each person who screened in and completed the 1989 detailed interview received a 1989 detailed weight.

The 1989 detailed noninterview adjustment factor was computed using final screener weights and was defined by residence (residing in an SMSA or not) Census region, and type of impairment reported during the screening interview. The final 1989 detailed weight equals the product of the final screener weight and the 1989 detailed noninterview adjustment factor. The final weight is found in characters 400 through 411 of the LTC public use tape.

The 1989 detailed noninterviews are persons who refuse to be interviewed, are unable to respond, or cannot be located, or who moved outside the country after the screener was completed. All noninterviews are type A noninterviews, since all persons eligible for interviews were screened and therefore must have been members of the LTC universe.

Using the 1989 User Tape to make cross-Sectional Estimates

A cross-sectional estimate is an estimate based on a person's characteristic in the 1989 Long Term Care Survey such as the total number of impaired persons in 1989 requiring personal help eating, the percentage of impaired persons in 1989 who do not keep in touch with friends, or the total impaired persons ever admitted to a nursing home. In order to make estimates of the type described above, use the 1989 LTC users tape. In order to make the best 1989 estimate sum the weights in characters 400 through 411 on the tape for the persons with the characteristic of interest.

In order to make estimates compare to 1982 and/or 1984 estimates sum the weights in characters 388 through 399 of the tape for person with the characteristic of interest.

RELIABILITY OF ESTIMATES

Since the LTC estimates are based on a sample, they may differ somewhat from the figures that would have been obtained if a complete census had been taken using the same questionnaires, instructions, and enumerators. There are two types of errors possible in an estimate based on a sample survey: sampling and nonsampling. The standard errors primarily indicated the magnitude of the sampling error. They also partially measure the effect of some nonsamping errors in response and enumeration, but do not measure any systematic biases in the data. The full extent of nonsampling error is unknown. Consequently, particular care should be exercised in the interpretation of figures based on a relatively small number of cases or on small differences between estimates.

Nonsamping variability. Nonsampling errors can be attributed to many sources such as: inability to obtain information about all cases in the sample, definitional difficulties, differences in the interpretation of questions, inability or unwillingness of the respondents to provide correct information, inability to recall information, errors made in collection such as in recording or coding the data, errors made in processing the data, errors made in estimating values for missing data, and failure to represent all persons with the sample (undercoverage).

A error in creating the sampling frame for 1982 caused Railroad Retirement beneficiaries in 14 counties to have no chance of selection; Railroad Retirement beneficiaries aged 65 years through 65 years and 3 months at the time of sample selection were given no chance of selection in the remaining sample areas. This error also occurred in the sampling frame for the 4,916 new medicare enrollees. This omission resulted in undercoverage of less than .2 percent in 1984 cross-sectional and 1982 longitudinal estimates of the total population aged 65 and over. Bias could be greater for estimates for smaller subpopulations.

Persons 65 years and over who are not Medicare enrollees are missed by the sampling frame. This undercoverage is estimated to be no more than 3.7 percent of the population of interest. Undercoverage varies by age, race, and sex. Generally, undercoverage is greater for black persons less than 85 years old than for the corresponding nonBlack persons; and for either race, is greater for those less than 70 years than for those 70 years and over. Bias due to undercoverage is partially corrected for by post-stratification, described later.

The inability to obtain information for all cases in the sample has several causes. Of the 17,565 persons designated for screening interviews, 583 were found to be ineligible. Some of these persons were ineligible because they were ineligible for the 1982 survey and were deceased before April 1, 1989, others moved outside the country before April 1, 1989 and still others, who should have been treated as eligible noninterviews, were ineligible because they moved within the country beyond 100 miles of the regional office before April 1, 1989. Three hundred seventy three persons were eligible for interviews but interviews were not obtained because these persons could not be located, refused to be interviewed, could not be found home, or were unable to respond and a proxy was unavailable.

There were 4,962 persons designated for detailed interviews. This included 1,354 persons in institutions and persons not in institutions and three persons who were found to be ineligible for sample. There were 499 persons who were found to be eligible but interviews were not obtained for several reasons. Thirty-one of the 499 persons could not be located, 7 were not at home, 14 were temporarily absent and a proxy was not available, 161 refused to be interviewed, 14 were unable to respond and a proxy was unavailable, 199 were deceased on or after April 1, 1989, 59 were not interviewed for some other reason, 5 moved outside the country on or after April 1, 1989 and 9 moved within the country beyond the limit of 100 miles from the regional office.

Beside the above error, undercoverage in the LTC results from persons missed in the sampling frame. It is known that undercoverage varies by age, race, and sex. Ratio estimation to independent age-race-sex population controls, as described previously, partially corrects for the bias due to survey undercoverage. However, biases exist in the estimates to the extent that, in a

given age-race-sex group, missed persons have different characteristics than interviewed persons. Further, the independent population controls used have not been adjusted for undercoverage in the decennial census.

Sampling variability. The standard errors given in tables 1-4 are primarily measures of sampling variability, that is, of the variation that occurred by chance because a sample rather than the entire population was surveyed. The sample estimate and its standard error enable one to construct confidence intervals - ranges that would include the average results of all possible samples with a known probability. For example, if all possible samples were selected, each of these being surveyed under essentially the same general conditions and using the same sample design, and if an estimate and its 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 results of all possible samples.

The average estimate derived from all possible samples either is or is not contained in any particular computed interval. However, for a particular sample, one can say with a specified

confidence that the average estimate derived from all possible samples is included in the confidence interval.

Standard errors may also be used to perform hypothesis testing - a procedure for distinguishing between population parameters using sample estimates. The most common types of hypotheses are 1) the population parameters are identical and 2) they are different. An example of this would be comparing the number of persons impaired in their activities of daily living (ADL) to those limited in their instrumental activities of daily living (IADL). Tests may be performed at various levels of significance, where a level of significance is the probability of concluding that the parameters are different when, in fact, they are identical.

Note when using small estimates. Because of the large standard errors involved, there is little chance that summary measures would reveal useful information when computed on a base smaller than 115,000. However, smaller estimates are provided in some tables primarily to permit such combinations of the categories as serve each user's needs.

Table 1. Standard Errors for Cross-Sectional Estimates of Noninstitutionalized Persons

(Numbers in thousands)

Size of Estimate	Standard Error	Size of Estimate	Standard Error
25	6.2	1,000	36.0
50	8.7	2,000	45.5
100	12.3	3,000	48.4
250	19.2	4,000	45.9
500	26.6	5,000	36.9
750	31.9		

Use Table 1 to obtain standard errors for any 1989 cross-sectional estimate of noninstitutionalized persons (for example, the total number of 65+ noninstitutionalized persons who are incontinent in 1989).

Table 2. Standard Errors for Cross-Sectional Estimates of 1989 Institutionalized Persons

(Numbers in thousands)

Size of Estimate	Standard Error	Size of Estimate	Standard Error
25	5.2	1,000	35.8
50	7.2	2,000	46.3
100	10.1	3,000	48.2
250	15.1	4,000	45.6
500	19.1	5,000	36.7
750	20.4		

Use Table 2 to obtain standard errors for any 1989 cross-sectional estimate of institutionalized persons (for example, the number of 1989 institutionalized persons requiring personal help eating).

Table 3. Standard Errors of Estimated Percentages of Noninstitutionalized Persons

Base of estimated percentage (thousands)	Estimated Percentage				
	2 or 98	5 or 95	10 or 90	25 or 75	50
25	3.5	5.4	7.5	10.8	12.4
50	2.5	3.8	5.2	7.6	8.7
100	1.7	2.7	3.7	5.3	6.2
250	1.1	1.7	2.3	3.4	3.9
500	0.7	1.2	1.7	2.5	2.8
750	0.6	1.0	1.4	1.9	2.2
1000	0.5	0.9	1.2	1.7	1.9
2000	0.4	0.6	0.9	1.2	1.4
3000	0.3	0.5	0.6	1.0	1.2
4000	0.3	0.4	0.6	0.9	1.0
5000	0.2	0.4	0.5	0.7	0.9

Use table 3 to obtain standard errors for the percentage of 1989 noninstitutionalized persons with a particular characteristic (for example, the percentage of noninstitutionalized persons who are ADL and IADL).

Table 4. Standard Errors of Estimated Percentages of 1989 Institutionalized Persons

Base of estimated percentage (thousands)	Estimated Percentage				
	2 or 98	5 or 95	10 or 90	25 or 75	50
25	2.9	4.6	6.3	9.0	10.4
50	2.0	3.2	4.5	6.4	7.3
100	1.5	2.2	3.1	4.5	5.2
250	1.0	1.4	2.0	2.9	3.3
500	0.6	1.1	1.4	2.0	2.3
750	0.5	0.8	1.2	1.6	1.9
1000	0.4	0.7	1.0	1.4	1.6
2000	0.3	0.4	0.7	0.8	1.0
3000	0.3	0.4	0.5	0.8	1.0
4000	0.2	0.3	0.5	0.7	0.8
5000	0.2	0.3	0.4	0.6	0.7

Use table 4 to obtain standard errors for the percentage of 1989 institutionalized persons with a particular characteristic (for example, the percentage of 1989 institutionalized persons who require personal help getting out of bed).

Standard Error Tables and Their Use. A number of approximations were required to derive standard errors that would apply to a large number of estimates and that could be prepared at a moderate cost. Therefore, instead of providing an individual standard error for each estimate, generalized sets of standard errors are provided for various types of characteristics. As a result, the sets of standard errors provided give an indication of the order of magnitude of the standard error of an estimate rather than the precise standard error.

The figures presented in tables 1 through 4 are approximations to standard errors of various estimates for persons in the United States.

Standard errors for intermediate values not shown in the generalized tables of standard errors may be approximated by linear interpolation.

Two parameters, "*a*" and "*b*" are used to calculate the standard errors for each type of characteristics; they are presented in table 5. The values of *a* and *b* are determined by fitting curves of variances in terms of estimates, i.e., by treating the variance of an estimate as a function of the estimate itself. This procedure is useful because it tends to produce a smoothing effect on the variance estimates and, perhaps more importantly, because it enables the analyst to quickly compute variance estimates for any item of interest whereas it would clearly be impossible to publish variance estimates for every possible item of interest.

The *a* and *b* parameters were used to calculate the standard errors in tables 1-4. They also may be used to directly calculate the standard errors for estimated numbers and percentages, as explained in the following section.

Standard Errors of Estimated Numbers. The approximate standard error, S_x of an estimated number can be obtained in two ways. It may be obtained by using the standard error on the estimate obtained from table 1, 2,… or 4, alternatively, it may be approximated by using formula (1), from which the standard errors in table 1-4 were calculated.

$$s_x = \sqrt{ax^2 + bx} \qquad (1)$$

Here x is the size of the estimate and a and b are the parameters in table 5 associated with the particular type of characteristic.

Standard Errors of Estimated Percentages. The reliability of an estimated percentage, computed using sample data for both numerator and denominator, depends upon both the size of the percentage and the size of the total upon which this percentage is based. Estimated percentages tend to be relatively more reliable than the corresponding estimates of the numerators of the percentages, particularly if the percentages are 50 percent or more. When the numerator and denominator of the percentage are in different categories, use the "f" factor or parameters from table 5 indicated by the numerator. The approximate standard error, $S_{(x,p)}$, of an estimated percentage can be obtained by using the standard error on the estimate from tables 3 or 4, alternatively, it may be approximated by using formula (2) from which the standard errors in tables 3-4 were calculated. Use of this formula will give more accurate results than use of Tables 3-4.

$$s_{(x,p)} = \sqrt{(b/x)\cdot p(100-p)} \qquad (2)$$

Here x is the size of the subclass of persons which is the base of the percentage, p is the percentage ($0<p<100$) and b is the parameter in table 5 associated with the particular type of characteristic in the numerator of the percentage.

Illustration of the use of standard error tables. According to the 1989 LTC there were 2,790,005 persons requiring personal help bathing in 1989. Table 1 shows the standard error of an estimate of this size to be approximately 47,800. Alternatively, using the parameters in table 5,

$$s_x = \sqrt{(-0.00025455)(2{,}790{,}005)^2 + (1545)(2{,}790{,}005)} = 48{,}261$$

Table 5. The "*a*" and "*b*" Parameters for Computing Approximate Standard Errors of Estimated Numbers and Percentages

	Parameters	
Characteristics	*a*	*b*
Noninstitutionalized Persons	-.00025455	1545
Institutionalized Persons.	-.00070004	1079

The 68 percent confidence interval as shown by the data is from 2,741,744 to 2,838,266. Therefore, a conclusion, that the average estimate derived from all possible samples lies within a range computed in this way would be correct for roughly 68 percent of all possible samples. Similarly, we could conclude with 95 percent confidence that the average estimate derived from all possible samples lies within the interval from 2,693,483 to 2,886,526 using twice the standard error.

Of the 605,286 persons requiring help in eating, 15.2 percent were black. Table 3 shows the standard error of 15.2 percent on a base of 605,286 to be approximately 1.8 percent.

Alternatively, this standard error could have been derived by using the "*b*" parameter for noninstitutionalized persons (table 5) in formula (2)

Consequently, the 68 percent confidence interval as shown by this data is from 13.4 to 17 percent, and the 95 percent confidence interval is from 11.6 to 18.8 percent.

Standard error of a difference. For a difference between two sample estimates the standard error is approximately equal to

$$s_{x-y} = \sqrt{S_x^2 + S_y^2 - 2?S_x S_y}$$

where S_x and S_Y are the standard errors for the estimates x and y (from tables 1 through 4), respectively. The estimates can be numbers, percents, ratios, etc. The correlation coefficient r is not generally available and can be assumed to equal zero. Making this assumption will result in accurate estimates of standard errors for the difference between two estimates of the same characteristic in two different areas, or for the difference between two uncorrelated characteristics in the same area. If however, there is a high positive (negative) correlation between the two estimates, the formula will overstate (understate) the true standard error.

Illustration of the computation of the standard error of a difference in estimates. LTC estimates show the number of 65-74 year old persons requiring personal help in eating was 199,245 and the corresponding number of 75-79 year olds was 114,377. The estimated difference is 84,868. The standard error on the estimate 119,245 is 17,255 and the standard error on the estimate is 114,377 is 13,167 (both computed using formula (1)). The correlation coefficient in this case is known to equal zero. The standard error associated with the estimated difference of 84,868 is

$$\sqrt{471{,}104{,}914} = 21{,}705$$

This means that the 95 percent confidence interval around the 84,868 difference is from 42,322 to 127,410, i.e., 84,868 ± 1.96 *(21,705). A conclusion that the average estimate of the difference derived from all possible samples lies within a range computed in this way would be correct for roughly 95 percent of all possible samples. Thus, we can conclude with 95 percent confidence that the number of 65-74 year old persons requiring help eating is greater than the number of 75-79 year old persons requiring help eating.