NET COVERAGE ERROR IN THE 1997 CENSUS OF AGRICULTURE

Glenn S. Wolfgang, National Agricultural Statistics Service
Washington, D.C. 20250

KEY WORDS: Census completeness, dual system estimation, classification error, mail list error

1 Introduction

Coverage evaluation has helped assess the accuracy and completeness of the census data through many changes over fifty years. The U. S. Commerce Department's Bureau of the Census has conducted each census of agriculture since the first in 1840. For the first time, the 1997 Census of Agriculture will be conducted by the National Agricultural Statistics Service of the U.S. Department of Agriculture. This paper describes the objectives, sample design, estimators, data releases, and improvements in the 1997 Census of Agriculture Coverage Evaluation Program. It updates and extends census of agriculture coverage evaluation descriptions in Lewis, Wolfgang, and Vacca (1994).

2 The Purpose of Coverage Evaluation

Coverage error is the failure to correctly collect and count whole-farm data. (Let the word farm represent anything that satisfies the definition of an agricultural operation.) As in other data collection efforts, errors of many types affect the results. Many errors, such as sampling errors and most of the response errors observed in census data, are assumed to be random error that seldom biases national or even county level data. However, coverage error consistently results in bias at all levels of data.

There is an overcount and an undercount effect in both mail list and farm status classification errors. Specifically, undercount arises when farms are never contacted in data collection because they are not on the mail list (NML) or when data are collected for a farm that becomes incorrectly classified as a nonfarm, i.e., incorrectly classified undercount (ICU). Overcount arises when a farm is represented more than once on the mail list, and the duplication (DUP) is not corrected in census processing or when data are collected for a nonfarm that becomes incorrectly classified as a farm, i.e., incorrectly classified overcount (ICO). Historically, the greatest effect comes from records not on the mail list when census forms are mailed out. Other coverage errors mitigate that effect somewhat and remain important, if not overpowering. The relative impact of the different coverage errors varies considerably by state.

3 Selected 1992 Census of Agriculture Coverage Evaluation Results

A total of 2,160,158 farms were estimated during the 1997 Census of Agriculture coverage evaluation. There was a net undercount of farms of approximately 11.1% (±0.7%). This includes an estimated 252,646 (±13,865) farms not on the mail list, 61,966 (±5,884) farms incorrectly classified as nonfarms, 50,400 (±5,223) nonfarms incorrectly classified as farms, and 23,505 (±2,513) duplicated farms. (Note: The numbers in parentheses represent 1.645 standard error above and below the estimates.) Over the years net coverage error has fluctuated between 3.4% in 1978 and around 11% in 1974 and 1992. A 16% net error in the Northeast region of the U.S. contrasts to 6.6% in the Midwest. The 4.8% overcount in the West contrasts with the rate (about 3%) elsewhere. The impact of NML undercount on the farm count (U.S.) was much higher (11.7%) than for land in farms (1.9%) or for value of products sold (1.1%).

Many other estimates from the 1992 Census of Agriculture coverage evaluation were made available in print (1992 Census of Agriculture, Volumes 1 and 2), on CD-ROM, and on the Internet. NML error in farm count totals and in subsets based on characteristics such as farm size or type of operation were reported at state, regional, and national levels. The impact of NML error on major crop acreages or livestock inventories were reported at regional and national levels. Overcount and net (undercount minus overcount) errors were reported for farm counts at regional and national levels.

4 1997 Census of Agriculture Coverage Evaluation Program

The 1997 Census of Agriculture Coverage Evaluation Program is conducted using data from several National Agricultural Statistics Service (NASS) survey sources as well as census questionnaire responses. Two area frame surveys, the 1997 June Agriculture Survey (JAS) and the Fall Area Survey (FAS), measure the NML error. The 1997 Classification Error Survey (CES) evaluates the DUP, ICU, and ICO components of coverage error. The JAS was used as in previous census evaluations when the census was conducted by the Census Bureau. The FAS is a new NASS survey. The CES, like the census itself, is being conducted by NASS for the first time.

4.1 Census of Agriculture Data Collection

The census of agriculture is the major comprehensive source of data for agricultural production and operator characteristics in the United States and its counties and states. Censuses are conducted on a five year cycle for years ending in 2 and 7. A census mail list is compiled from names and addresses of agricultural operations gathered from many administrative sources. Report forms for the 1997 Census of Agriculture will be mailed to farm and ranch operators in late December 1997 to collect data for the 1997 calendar year. Those not responding by February 2, 1998 will be contacted by mail follow-ups or computer assisted telephone interviews (CATI). Under a contractual arrangement with the Bureau of the Census, the questionnaires will be received by the Data Processing Division in Jeffersonville, Indiana. They will be checked for completeness and accuracy. The data will be keyed and edited with computer assistance. Forms will be forwarded to state statistical offices of NASS, where data will be reviewed, processed, and tabulated by the end of 1998. Census results will be used to revise 1997 NASS crop and livestock estimates. Census tabulations will be published in 1999.

4.2 Not-on-the-Mail-List Study

The 1997 Census of Agriculture Coverage Evaluation Program will produce state-level estimates of data for farms not on the census mail list. Estimates will not be made for Alaska or Hawaii. NML estimates will be made using a dual-system estimation model (Wolter, 1986), where the JAS and FAS together are a source of farm enumeration independent of the census. That independence arises from how differently the JAS and FAS frames are constructed. The census mail list, a list frame, is a cost-efficient way to locate agricultural operations. The JAS and FAS are based on an area frame, a more comprehensive way to find farms.

The area frame is developed by dividing the land within a state into six to eight land-use strata, such as intensively cultivated land, urban areas, agricultural urban areas, and rangeland. The land-use strata are outlined on county highway maps using permanent and easily recognizable land features. Substrata are defined based on kinds of agriculture. Substrata are then further divided into segments of approximately equal size using aerial photographs. On the average, a segment contains portions of about three farm operations. Since the land area within each segment is completely enumerated, the segment and not the farm is the basic unit of analysis for the area frame. Cotter and Nealon (1987) provide more details on the area frame design.

Each year, during the first two weeks of June, field enumerators from state statistical offices canvass segments selected for the JAS to complete the maps of farm areas (tracts) in those segments and collect, by personal interview, current farm data such as planted acreage of crops and names and addresses for each tract, the ultimate reporting unit. Only one farm operation is associated with a tract, although a farm may have tracts in more than one segment. The area sample that results provides the basis for several subsequent NASS surveys including the September, December, and March Agricultural Surveys.

The FAS sample segments in most states are different from those selected for the JAS. Canvassing and survey collection are done by mid-December. The FAS sample is newly designed in 1997 to collect information on livestock and on pesticide use. The area frame sample resulting from adding JAS and FAS names together is projected to be about the same size as that used in the 1992 Census of Agriculture Coverage Evaluation.

JAS and FAS names, addresses, and other identifier information for all sample area segment tracts that had any indication of agricultural activity will be available to the coverage evaluation by January 1998. The names and addresses on these records will be matched with those on the census mail list using AUTOMATCH, record linkage computer software. Each JAS and FAS record will be assigned an initial match status—either matched or nonmatched to a record on the census mail list. Each JAS and FAS sample name and address not found on the mail list (potential NML) will be included in CATI data collection (or mailed out) using the same instrument designed for the census follow-up of nonrespondents. Potential NML responses will be sent through the same computer edits as other census responses to avoid introducing processing bias.

During subsequent processing, newly collected census acres and other data will be compared to JAS or FAS responses to ensure that each name and address initially matched by the computer was a valid match. The census database will be searched to find potential matches to the nonmatched cases using the names, addresses, and data from the potential NML response. Potential NML cases must be verified as farms before they are counted as errors.

4.3 Classification Error Survey

To measure the three errors (ICU, ICO, and DUP) in census processing, the CES sample (targeting about 24,500 responses) will be selected from the census mail list before the census forms are mailed out using a systematic sample design stratified by state. The proportion of total CES sample allocated to each state is approximately the same as that state's proportion of tracts in the combined JAS and FAS area frames, with a minimum sample in each state to ensure adequate estimates at that level. Operations that are extensively reviewed in census processing (e.g., those with expected sales over $500,000, multi-units, Indian reservations, and institutional farms) are not eligible for selection.

When a census questionnaire for a CES sample case is returned, a CES CATI interview or, if the telephone number is not usable, a CES questionnaire mailing can proceed. Analysts in NASS state statistical offices review the CES responses to classify each record as either a farm or nonfarm and compare that farm status to the census farm status to identify cases that have been incorrectly classified. All census records are searched with the help of automated matching to find duplicated farm records in the census.

CES responses provide crop acreages, quantities harvested, and livestock inventory data used to estimate the impact of ICU errors for those items in census tabulations. Corresponding census data are used to estimate the impact of ICO and DUP errors.

5 Estimators of Coverage Error

The true total farm count (T) in a state is defined as the census tabulated number of farms (C) plus the net of all coverage errors introduced above (*Net*). To correct for the errors, undercounted farms (*NML* and *ICU*) are added to census counts, while overcounted farms (*ICO* and *DUP*) are subtracted. The universe total is then:

$$T = C + Net = C + NML + ICU - ICO - DUP \qquad (5.1)$$

The total for some agricultural commodity or farm subgroup x is defined similarly:

$$T_x = C_x + NML_x + ICU_x - ICO_x - DUP_x \qquad (5.2)$$

The estimation of the various components of T and T_x are discussed in the following sections.

5.1 Estimation of Farms Not on the Mail List

An NML estimate is derived using a coverage error model based on dual-system estimation theory. Its properties and derivation are discussed by Wolter (1986). The model can be extended to provide NML estimates for other census data as well. The major dual system model assumption is that the census and area frame are independent of one another—i.e., the probability of a farm being on the census mail list is independent of the probability of a farm being enumerated by the area frame. Other assumptions are listed by Wolter, including several that translate to assuming no errors due to processing (e.g., in matching area frame to census records), nonresponse imputation, response mistakes, and even other kinds of coverage error (i.e., list duplication).

After each record representing a name and address in the area frame sample has been determined to match or not match a record on the census mail list and to be a true farm, it can be assigned a value:

y_{Mi} = 1, if the area frame tract I was matched to a farm on the census mail list, or 0, otherwise;

y_{Qi} = 1, if the area frame tract I was not matched to a farm on the census mail list, or 0, otherwise.

These observed values are expanded to farm universe estimates in a weighted segment estimator, which uses the expansion factor and a farm weight appropriate to the tract. The expansion factor, e_i, is the inverse of the sampling proportion for the segment in which tract I is found. While the mechanics of drawing the area frame sample entail several levels of stratification—segments within substrata within land-use strata—there is simply one e_i value for each segment. The farm weight, a_i, is the proportion of tract acreage in the whole farm.

$$\hat{M} = \sum_{i=1}^{n} e_i \, a_i \, y_{Mi} \quad \text{and} \quad \hat{Q} = \sum_{i=1}^{n} e_i \, a_i \, y_{Qi} \qquad (5.3)$$

These estimates fit in two contingency table cells of the dual-system estimation model:

		In the area frame? Yes	No	
On the Census Mail List?	Yes	$\hat{M}$		$\hat{C}$
	No	$\hat{Q}$		$\hat{NML}$

$\hat{M}$ = expanded estimate of matched farms; farms represented both on the census mail list and in the area frame

$\hat{Q}$ = expanded number of nonmatched farms; farms represented in the area frame but not on the census mail list;

$\hat{C}$ = number of farms on the census mail list adjusted for classification error and list duplication;

$\hat{NML}$ = number of farms not on the mail list.

Independence in the model and, therefore, this contingency table means that

$$\hat{M} \,/\, \hat{Q} = \hat{C} \,/\, \hat{NML} \qquad (5.4)$$

$$\text{so, } \hat{NML} = \hat{Q}\,\hat{C} \,/\, \hat{M} \qquad (5.5)$$

To extend the estimator beyond the simple farm count, q_{xi}, $\hat{Q}_x$, and $\hat{NML}_x$ are introduced for some quantity x. The x denotes a farm commodity or subgrouping variable such as acres in the entire farm, acres harvested for a particular crop, farm counts within size or demographic subgroups, crop quantities produced, or livestock inventories.

$$\hat{Q}_x = \sum_i e_i\, a_i\, y_{Qi}\, q_{xi} \qquad (5.6)$$

$$N\hat{M}L_x = \hat{Q}_x\, \hat{C} \,/\, \hat{M} \qquad (5.7)$$

where,

q_{xi} = observed quantity for data item x on farm I;

$\hat{Q}_x$ = expanded farm data quantity x in the area frame associated with farms not on the mail list; and

$N\hat{M}L_x$ = undercount in farm data quantity x associated (only) with farms not on the mail list.

$\hat{Q}$ may be viewed as a special case of $\hat{Q}_x$, where quantity x is simply the entire farm count. $N\hat{M}L_x$ is used below in the general sense that includes $N\hat{M}L$, the farm count.

5.2 Poststratification of the JAS Records

Coverage evaluations of past censuses of agriculture have shown that small farms are less likely to be on the mail list than large farms. Poststratification is used in NML estimation to reduce the effects of such heterogeneous capture probabilities. The JAS records are assigned to one of two poststrata based on the total value of products sold (TVP) and using $2,500 as the cut-off. The poststrata are collapsed if fewer than ten records (including matched) are in either poststratum. The poststratified estimator is the sum of $N\hat{M}L_x$ applied within the h^{th} poststratum:

$$N\hat{M}L_x = \sum_{h=1}^{2} \hat{Q}_{xh}\, \hat{C}_h \,/\, \hat{M}_h \qquad (5.8)$$

5.3 Estimation of Incorrectly Classified and Duplicated Farms

1992 CES estimates of the number of incorrectly classified farms and the estimate of the number of duplicated farms were calculated at the region level for farm counts, including counts of farms in categories (e.g., the number of incorrectly classified farms smaller than 50 acres in size). 1997 CES errors may be reported at the state level for all quantities that provide NML estimates.

Several possible estimates of the impact on quantity x data due to farms incorrectly classified or duplicated in the state may be calculated. $IC\hat{U}_x$, $IC\hat{O}_x$, or $D\hat{U}P_x$ may be computed separately or combined into estimates of overcount, $\hat{O}_x = IC\hat{O}_x + D\hat{U}P_x$, or of general CES error, $\hat{E}_x = IC\hat{U}_x + IC\hat{O}_x + D\hat{U}P_x$. Let $C\hat{E}_x$ represent any of these.

$$C\hat{E}_x = \sum_{i=1}^{n} e\, y_{CEi}\, q_{xi} \qquad (5.9)$$

where,

e = the state expansion factor; $e = N/n$

N = the number of records on the mail list (excluding undeliverable mailings)

n = the number of CES responses in the state

y_{CEi} = 1 if farm I has ICU error and $C\hat{E}_x$ is $IC\hat{U}_x$ or $\hat{E}_x$

-1 if farm I has ICO error and $C\hat{E}_x$ is $IC\hat{O}_x$, $\hat{O}_x$, or $\hat{E}_x$

$-(m_i-1)/m_i$ if farm I has DUP error and $C\hat{E}_x$ is $D\hat{U}P_x$, $\hat{O}_x$, or $\hat{E}_x$

0 otherwise;

m_i = the number of times (including the sampled record) farm I was counted in census tabulations

q_{xi} = the value reported for quantity x on farm I.

The duplication factor, $(m_i-1)/m_i$, compensates for the multiple opportunities, m_i, the CES sampling provides to observe duplication error. Kish (1965) describes this problem and solution.

5.4 Estimators for Net and Other Combinations of Errors and Census Data

Adding NML and all CES errors, with a negative sign for overcount, produces the net error. NML plus ICU is U, the undercount error. ICO plus DUP is O, the overcount error. Total estimated farm data generally means tabulated census plus net error. But it has been useful to sometimes report estimated farm data including only some of the error components, e.g., census plus undercount. All of these variations derive by dropping terms in formula (5.2).

In censuses of agriculture prior to 1997, CES data were not computed at state levels, so net error and other estimates incorporating CES estimates are only available at region or national levels. Relevant individual state NML estimates are summed to the regional level and then adjusted by the rate of CES errors before they are combined into regional or national net error estimates.

5.5 Estimators for Percent Errors

Percent errors (expressed as proportions) may be reported for NML, undercount, overcount, or net error. The denominator is a sum of census data with some errors. An example from the 1992 Census of Agriculture coverage evaluation showing how state level NML errors are combined with region level CES errors is the proportion undercount:.

$$P\hat{U}_{xg} = \frac{N\hat{M}L_{xg} + IC\hat{U}_{xg}}{N\hat{M}L_{xg} + IC\hat{U}_{xg} + \hat{C}_{xg}} \qquad (5.10)$$

where g designates the region comprised of states (k) and poststrata (h) within states.

5.6 Estimators for Averages

Averages may be reported for NML, undercount, and net errors, as well as for census and adjusted totals. Average land in farms not on the mail list is an example. The numerator is the sum of NML error in total farm acres, and the denominator is the NML farm count.

$$AN\hat{M}L_x = N\hat{M}L_x \,/\, N\hat{M}L \qquad (5.11)$$

6 Variance Estimators

Variance estimation is the basic calculation of variance for a simple or stratified random sample (Cochran, 1977; Kott, 1988) for variables that are observed directly in CES or NML results, i.e.,

$\hat{y}_j = \{\hat{M},\ \hat{Q},\ \hat{Q}_x,\ \hat{C},\ \hat{C}_x,\ \hat{E},\ \hat{E}_x,\ \hat{O},\ \hat{O}_x,\ IC\hat{U}, \text{ or } IC\hat{U}_x\}$. Each element of $\hat{y}_j$ actually represents a collection of values observed at poststratum, state, or regional levels.

Taylor series methods described by Wolter (1985) were used to derive the variance estimators for error and census components combined into a ratio, $\hat{R}$. Shortcut derivations are often helpful. But estimators like (5.10) that involve many terms derive more easily using the calculus in the basic definition:

$$V\hat{a}r(\hat{R}) \approx \sum_{jj=1}^{m}\sum_{j=1}^{m} \frac{\partial\hat{R}}{\partial\hat{y}_j}\frac{\partial\hat{R}}{\partial\hat{y}_{jj}} C\hat{o}v(\hat{y}_j\hat{y}_{jj}) \qquad (6.1)$$

where $\hat{y}_j$ is any observed variable that appears in $\hat{R}$, and $\hat{y}_{jj}$ is any $\hat{y}_j$, even when identical.

To simplify the calculus, the ratio may be expressed so that terms in the numerator are all additive and likewise for the denominator. For example, (5.10) becomes:

$$P\hat{U}_{xg} = \frac{\sum_{kh\in g} \Pi_{M-1}\hat{Q}_{xkh}\hat{C}_{kh} + \Pi_M IC\hat{U}_{xg}}{\sum_{kh\in g} \Pi_{M-1}\hat{Q}_{xkh}\hat{C}_{kh} + \Pi_M IC\hat{U}_{xg} + \Pi_M \hat{C}_{xg}} \qquad (6.2)$$

where,

Π_M = the product of all $\hat{M}_{kh}$ in region g;

Π_{M-1} = the product of all $\hat{M}_{kh}$ in region g except the particular $\hat{M}_{kh}$ corresponding to the $\hat{Q}_{xkh}\hat{C}_{kh}$ in that term within the sum.

Details of such derivations may be found in Wolfgang (1996). For the 1997 Census of Agriculture coverage evaluation, estimating CES errors at the state level simplifies formulae such as this.

Assumptions that particular covariances equal zero also simplify these derivations. $\hat{Q}_x$ and $\hat{M}$ arise from area frame data and are independent of census data and the census-sampled CES data, so the covariances between variables from independent sources are assumed to be zero. Covariances between census and CES variables were computed for 1992 but should be assumed zero.

6.1 Variance of the Not-on-the-Mail-List Estimates

The variance estimator for the poststratified version of $N\hat{M}L_x$ builds upon variance estimators for poststrata $\hat{Q}_{xh}$, and $\hat{M}_h$, which are stratified variance estimators suited to the JAS sample design. For example, the variance of nonmatches in poststratum h is:

$$V\hat{a}r(\hat{Q}_{xh}) = \frac{n_h}{n_h - 1}\sum_{i=1}^{n_h}\left(e_i a_i y_{Qhi} q_{xi} - \frac{1}{n_h}\sum_{i=1}^{n_h} e_i a_i y_{Qhi} q_{xi}\right)^2 \qquad (6.3)$$

The variance of $\hat{M}_h$ and the covariance of $\hat{Q}_{xh}$ with $\hat{M}_h$ are similar. Details of the derivation for the poststratified $N\hat{M}L_x$ are in Lewis (1993) or Wolfgang (1996).

6.2 The CES Variance Estimators

The CES sample was drawn using systematic sampling. Variance estimators of incorrectly classified and duplicated cases computed as if drawn from a simple random sample are conservative if there is no periodic trend coinciding with the sampling interval in the sequence of the list from which the sample was selected.

The variance for errors due to farms found by the CES to be miscounted is:

$$V\hat{a}r(C\hat{E}_x) = e^2\frac{n}{n-1}\sum_{i=1}^{n}\left(y_{CEi}q_{xi} - \frac{1}{n}\sum_{i=1}^{n} y_{CEi}q_{xi}\right)^2 \qquad (6.4)$$

6.3 Variance Estimators for Net and Other Combined Errors

A general form of the variance estimators for net error and other combined estimates, including percent error estimates, was used. It provided structure to help verify and manage terms and make it easier to modularize computation programs. That general form follows the pattern of (6.1):

$$V\hat{a}r(\hat{R}) \approx f_0\sum_{jj=1}^{m}\sum_{j=1}^{m} f_j f_{jj} C\hat{o}v(\hat{y}_j\hat{y}_{jj}) \qquad (6.5)$$

where $\hat{R}$ is an estimate of undercount, adjusted totals, or error proportion for the nation or a region, summed, as applicable, over regions g, states k, and poststrata h. In the general variance estimator, all $\hat{y}_j$ are included—any poststratum, state, or regional level estimate of a $\hat{y}_j$ that was used in computing even one of the reported estimates. There were hundreds of such terms. The multipliers (e.g., f_j or f_{jj}) were extracted from the derivatives of the individual $\hat{y}_j$ or $\hat{y}_{jj}$, with common terms removed to f_0. The multipliers are often 0 or 1, but sometimes are ratios of terms in the formula for $\hat{R}$.

7 Improvements in the Coverage Evaluation Program for 1997

Various developments will allow improvements in the Coverage Evaluation Program for the 1997 Census of Agriculture. Nearly every state in the nation has a NASS state statistical office where agricultural specialists with experience and understanding of the farms in that area collect and process data simultaneously with other states. As a result:

1. Error coding will be done more accurately and efficiently by specialists who understand agriculture.
2. NML error estimates will be available earlier—by the end of 1998 rather than in 1999.
3. Net error estimates will be released at the same time, rather than nearly two years later.
4. Coverage error estimates will be better used in NASS annual commodity estimate revisions—a program very important to agriculture economists.

A larger CES sample size than in 1992, with a guaranteed minimum sample for each state, helps make possible that:

5. Coverage error impact estimates will be available for more commodity variables.
6. Net error estimates will be released at the state level, rather than only at region level.

Technological advances help make the following additional improvements possible:

7. 1997 Census of Agriculture undercount is expected to be reduced, due to new census procedures to better screen out nonfarms without dropping true farms from the mail list.
8. Blaise CATI will provide a more efficient data-collection and processing tool—especially with its interactive edit capability.
9. AUTOMATCH, an automated record linkage program, will be applied more often and more flexibly by the Research Division of NASS to make sure sample farms were not missed or duplicated among census cases.
10. Computer Assisted Processing is being redesigned, not only to better fit the state office configuration, but to make the processing more efficient.

8 Acknowledgments

The success of the 1992 and 1997 Coverage Evaluation Programs are due to the team efforts of many past Bureau of the Census and present NASS workers. The author wishes to thank reviewers for their helpful comments.

9 References

1992 Census of Agriculture, Volume 1, Summary and State Data (in 51 parts), Commerce/Census, Washington, D.C.

1992 Census of Agriculture, Volume 2, Part 2, Coverage Evaluation, Commerce/Census, Washington, D.C.

Cochran, W. G. 1977. Sampling Techniques, 3rd Edition. John Wiley & Sons, New York.

Cotter, J. and Nealon., J. P. 1987. Area Frame Design for Agricultural Surveys. USDA/NASS, Washington, D.C.

Kish, L. 1965. Survey Sampling. John Wiley & Sons, New York.

Kott, P. S. 1988. Estimating Variance for the June Enumerative Survey. STB-88-06, USDA/NASS, Washington, D.C.

Lewis, P. J. 1993. Estimating Farms not on the Census Mail List. Interoffice Memo #92EAG-A-MC-28-A-1.

Lewis, P. J., Wolfgang, G. S., and Vacca, A. E. 1994. 1992 Census of Agriculture Coverage Evaluation Estimation. Proceedings of the Section on Survey Research Methods, ASA, pp. 1019-1024.

Wolfgang, G. S. 1996. Estimating Net Census Coverage Error. Interoffice Memo #92EAG-A-MC-28-A-5.

Wolter, K. M. 1985. Introduction to Variance Estimation. Springer-Verlag, New York.

Wolter, K. M. 1986. Some Coverage Error Models for Census Data. JASA. 81:338-346.