USING ETHNOGRAPHIC DATA TO EVALUATE DUAL SYSTEM ESTIMATES

Sally W. Thurston and Alan M. Zaslavsky, Harvard University
Sally W. Thurston, Department of Statistics, 1 Oxford Street, Cambridge, MA 02138

Key Words: triple system estimation, movers, census undercount

1 Introduction

In 1990, the U.S. Census Bureau used a 'capture-recapture' or dual system estimation (DSE) methodology to estimate total population including those missed by the census. The two 'systems' are the census and a Post Enumeration Survey (PES) (Hogan and Wolter 1988). One of the assumptions underlying use of the DSE to estimate population size is that within each poststratum (defined by some set of geographic and demographic variables), being in the census is independent of being in the PES. When these events are not independent, there is a 'correlation bias' which typically leads to underestimation of the number of people who are in neither the census nor the PES. Reasons for the possible failure of this assumption of independence have been discussed (Hogan and Wolter 1988). One method of checking this assumption, or indeed of estimating the statistical dependency between the census and PES, makes use of a third source of names and addresses – an alternative list (Marks, Seltzer and Krotki 1974, chapter 7D; Zaslavsky and Wolfgang 1993). By using a third independent source of names and addresses, the 2×2 table underlying the DSE can be expanded into a $2 \times 2 \times 2$ table in which only one of the 8 cells is unknown. Estimates of the unknown cell and of total population may also be calculated under suitable assumptions. Zaslavsky and Wolfgang (1993) discuss a number of methods to estimate this cell. In this paper we focus on two of these estimates, 'DSE:$(E \bigcup P) \times A$' and 'DSE:$(E \Delta P) \times A$'.

One such alternative list may be formed by combining several administrative lists. A list consisting of portions of lists from the Employment Security, Internal Revenue Service, Selective Service, Veteran's Administration, and driver's licence records was used in the 1988 Administrative List Supplement program conducted by the Census as part of the PES test in St. Louis, Missouri (Zaslavsky and Wolfgang 1993). For further discussion of the use of administrative lists, see also Alvey and Scheuren (1982), Citro and Cohen (1985, chapter 4), and Thurston and Zaslavsky (1994).

Alternative lists may also be compiled by ethnographers (Vigil 1988; Brownrigg and de la Puente 1992, de la Puente 1993, Martin and de la Puente 1993). Ethnographers work intensely in an area, and by getting to know individuals in the neighborhood, compile lists of names which may be more complete than the census or PES address list (Hamid 1992, Lerch 1992, Wingerd 1992a, Vigil 1988). In the 1990 evaluation programs using ethnographers, the ethnographers had connections to the people in the area, either by having worked with members of the community, or in living nearby. The ethnographers typically collected data from May through July.

One of the challenges posed by triple system estimation is proper cross-classification of cases by inclusion/exclusion in each of three sources. Improper classification may bias the subsequent population estimates. In addition, movers and non-movers may have different coverage rates in each of the sources. Consequently, calculations based on considering movers separately from non-movers are likely to be more accurate than estimates in which movers are either dropped from the triple system estimates, or are combined with non-movers. In general, movers may either be over- or undercounted at a different rate than non-movers (Citro and Cohen 1985, chapter 5) and it is often harder to match movers than non-movers with census records (Schafer 1991).

In this paper, we discuss methods of estimating the number of movers and non-movers, cross-classified by inclusion in census, PES, and alternative list (administrative or ethnographer's lists). We discuss how these estimates can be used to give total population estimates, and the relative merits of each estimate. These methods are applied to two sites in which data were collected by ethnographers. One site was a rural site in North Carolina, while the other site was an urban site in Florda. We also discuss ways to estimate the statistical dependency of the census and PES, and apply these to the two sites.

2 Triple System Estimation using Ethnographers, with movers

We follow the notation of Zaslavsky and Wolfgang (1993), in which the number of people in a given cell is denoted by x_{epa}, where $e = 1$ for people in the census (in the PES block or elsewhere) or 0 otherwise, and p and a are likewise 1 for people in the PES or alternative list respectively, or 0 otherwise. Poststratification is implicit here, so all relationships are assumed to be within a single poststratum (see Diffendal (1988) for details about poststratification used in the PES). In order to distinguish between non-movers, people who move into PES blocks between census and PES days ('in-movers'), and people who move out of PES blocks between census and PES days ('out-movers'), when needed we add a fourth subscript, n, i, or o for non-movers, in-movers, and out-movers respectively.

Zaslavsky and Wolfgang propose a number of estimators using administrative list data. We restrict consideration to the 'DSE:$(E\bigcup P)\times A$' and 'DSE:$(E\Delta P)\times A$' estimates because they are based on explicit assumptions about comparability of coverage rates in different subpopulations. Both of these estimators are DSEs in which the census and PES are treated as a single source. The 'DSE:$(E\bigcup P)\times A$' estimate is based on the assumption that the event of being in neither the census nor PES is independent of the event of being in the alternative list. This gives an estimate of the unknown cell as

$\hat{x}_{000}=x_{001}\times(x_{110}+x_{100}+x_{010})/(x_{111}+x_{101}+x_{011})$.

The 'ratio DSE:$(E\Delta P)\times A$' estimate is based on the same assumption applied to the subpopulation of people who are in either the census or PES, but not both. The rationale is that people captured in both census and PES are "easy to count" and therefore least comparable to those omitted in both. This gives the estimate

$\hat{x}_{000}=x_{001}\times(x_{100}+x_{010})/(x_{101}+x_{011})$.

Once this cell is estimated, the correlation bias between the census and the PES can be calculated, as can coverage rate and total population size estimates. Note that people omitted from both the census and the PES are more likely to be omitted from the alternative list than those included in the census and/or the PES. Thus both estimates of x_{000} are likely to be underestimates.

In making these estimates, we consider the sample of interest either to be PES-A (those residing in the sample blocks on Census day, i.e. the non-movers plus the out-movers), or PES-B (those residing in the sample blocks at PES time, i.e. the non-movers plus the in-movers). In principle PES-A and PES-B are both samples of the same population, and coverage rates for either are estimators of population coverage rates.

When movers are included in the estimates, we subdivide each cell (which has been cross-classified by inclusion in census, PES, and administrative list sources) into non-movers, in-movers, and out-movers. Estimates of x_{000} involve adding the number of in-movers or out-movers (depending on whether PES-A or PES-B is the sample of interest) to the number of non-movers, in each of the eight cells of the 3-way table.

People who move out of PES blocks after census day but before the PES (out-movers) are not directly seen by the PES, but through interviewing of current residents and neighbors, information about these people is collected by the PES. Ethnographers also gather information about peoplw who move out of the block, so when PES-A is the sample of interest, the estimate $\hat{x}_{000}$ using the DSE:$(E\bigcup P)\times A$ estimator is:

$\hat{x}_{000}=\hat{x}_{000n}+\hat{x}_{000o}=(x_{001n}+x_{001o})\times(x_{110n}+x_{110o}+x_{100n}+x_{100o}+x_{010n}+x_{010o})/(x_{111n}+x_{111o}+x_{101n}+x_{101o}+x_{011n}+x_{011o})$.

With the DSE:$(E\Delta P)\times A$ estimator,

$\hat{x}_{000}=\hat{x}_{000n}+\hat{x}_{000o}=(x_{001n}+x_{001o})\times(x_{100n}+x_{100o}+x_{010n}+x_{010o})/(x_{101n}+x_{101o}+x_{011n}+x_{011o})$.

When PES-B is the sample of interest, the situation is somewhat more complicated. Not only do we have no direct information as to the number of people who are in none of the three sources (x_{000n}, x_{000i}, and x_{000o}) but we also do not know the number of in-movers who are in the census, but not in the PES or ethnographers' lists (x_{100i}). The latter cell can not be observed because the only information about the addresses for these people is their census day address, which is not in the PES sample block. Under the stated assumptions, it is possible to count the number of people in all the remaining cells. When administrative lists are used, the counts in several additional cells are not observable.

Using ethnographers lists, x_{100i} is the only cell which is not directly observable. One method of estimating this cell relies on two assumptions: (1) the number of people who move into the PES blocks is equal to the number of people who move out of the PES blocks in the period between census day and the PES, in each poststratum; and (2) census coverage of in-movers is equal to census coverage of out-movers. Both of these assumptions reflect a view that the size and characteristics of the poststratum are unchanging, i.e. that in-movers are numerically and qualitatively similar to out-movers. Under these assumptions, the number of in-movers in the census equals the number of out-movers in the census, so $x_{111o}+x_{101o}+x_{110o}+x_{100o}=x_{111i}+x_{101i}+x_{110i}+x_{100i}$. Then

$$\hat{x}_{100i} = (x_{111o}+x_{101o}+x_{110o}+x_{100o})$$

$$-(x_{111i}+x_{101i}+x_{110i}). \quad (1)$$

Another method of estimating this cell relies on the assumption that among people in the census, PES coverage for non-movers is the same as PES coverage for in-movers. Since in-movers cannot be on the administrative lists, the appropriate reference group for them is all non-movers regardless of whether or not they were on an administrative list. Under this assumption we have $(x_{111n}+x_{110n})/(x_{101n}+x_{100n}) = (x_{111i}+x_{110i})/(x_{101i}+x_{100i})$, so

$$\hat{x}_{100i} = ((x_{111i}+x_{110i}) \times (x_{101n}+x_{100n})/(x_{111n}+x_{110n})) - x_{101i}. \quad (2)$$

The DSE:$(E \bigcup P) \times A$ estimate for x_{000} using PES-B is:

$\hat{x}_{000} = \hat{x}_{000n} + \hat{x}_{000i} = (x_{001n}+x_{001i}) \times (x_{110n}+x_{110i}+x_{100n}+\hat{x}_{100i}+x_{010n}+x_{010i})/(x_{111n}+x_{111i}+x_{101n}+x_{101i}+x_{011n}+x_{011i})$.

The DSE:$(E \Delta P) \times A$ estimate for x_{000} using PES-B is:

$\hat{x}_{000} = \hat{x}_{000n} + \hat{x}_{000i} = (x_{001n}+x_{001i}) \times (x_{100n}+\hat{x}_{100i}+x_{010n}+x_{010i})/(x_{101n}+x_{101i}+x_{011n}+x_{011i})$.

3 Results from Two Ethnographers' Sites

There were 29 sites used in the 1990 ethnographic evaluation program, of which 28 were in the continental US. All sites were selected to be in areas which were difficult to enumerate, and which had a large concentration of minorities, including Blacks, Hispanics, Asians, and American Indians (de la Puente 1993, Martin and de la Puente 1993). Each site consisted of about 100 housing units, usually in 1 or 2 blocks. Four of these sites were put into the PES and PES data were collected, but the four sites were not actually used for PES evaluation. It should be noted that ethnographic sites were selected in areas where the ethnographers already had a relationship with people in the area. Since this was neither a random sample of the country nor a random sample of areas with high undercount, results from these sites may not be generalizeable to a population.

In one of these sites, the ethnographer did not proceed far enough with the coding to allow for quantitative estimates. In a second site, some missing coding has lead to uncertainties in the data, and we did not attempt to resolve these uncertainties. We have examined the data from the remaining two sites.

The first site is in rural North Carolina, and is part of the community of the Waccamaw Siouan Indian tribe (Lerch, 1992). Eighty-seven percent of the residents were Indian and the remaining 13% were white spouses and children. A household in this site does not consist of a stable set of people separate from people in other households. Rather, households form and then regroup as people move to other addresses (Lerch 1992). Adult children of residents in the site often live in mobile homes or newly-built houses close to the house of their parents. Mail is delivered to numbered mailboxes along the side of the road, and most of the mailboxes serve more than one household.

The second site is an urban site in downtown Fort Lauderdale, Florida. Haitians comprised about 70% of the residents in the sample area, Blacks (African Americans) about 25%, and 5% were of other races (Wingerd, 1992a). The site is one block away from a major drug dealing area, and drug transactions were common in the site. There were bullet holes and multiple deadbolts on doors, many drawn curtains, and people carried guns and knives (Wingerd, 1992b). The Haitians were Creole-speaking recent entrants, and often did not speak English. Although many of them were undocumented aliens, they were more approachable by the ethnographer than were the Blacks. There was a high rate of mobility among the Haitians, as some found better places to live and some returned temporarily to Haiti. The Black residents of the site were suspicious of anything relating to the government (Wingerd 1992a). Of the 4 census forms filled out from this site, none were from the Black community (Wingerd, 1992b).

The ethnographers consistently found people that the census and PES missed. A common theme in the ethnographers' reports was that the ethnographers were able to enumerate hard-to-count people because they had gained the trust of the residents (Hamid 1992, Lerch 1992, Wingerd 1992a).

3.1 Population Estimates

In the following discussion we have excluded people who were determined to be erroneously enumerated. In the first site, 275 non-movers, 3 in-movers, and 8 out-movers were found by the census, PES, and/or ethnographers. The DSE:$(E \bigcup P) \times A$ and DSE:$(E \Delta P) \times A$ estimators, under PES-A and PES-B, all gave an estimate of the number of people missed by all 3 sources as less than 1 person. Because 8 people moved out of the site, the PES-A population estimate was 283 people, with an estimated coverage rate of 80%. Under PES-B, $\hat{x}_{100i} = 5$ under (1), and $\hat{x}_{100i} = 0$ under (2). Under (1), the population estimates were similar to PES-A, whereas the population estimate under (2) was only 278. The estimated coverage rates in all PES-

B cases were similar to those under PES-A. Had movers been ignored in the triple system estimate, the estimated population size would have been 275, with an estimated coverage rate of 80%.

If the ethnographers had not collected data in these sites and if only non-movers were considered, under the assumption of independence of census and PES, the DSE would have predicted that 1.28 people were missed by both the census and PES in the first site. The population estimate would have been 268 people under the DSE, with an estimated coverage rate of 82.5%.

In the second site, 74 non-movers, 17 in-movers, and 1 out-mover were found by at least one source. As in the first site, less than 1 person was estimated to be missed by all 3 sources, under all estimators. The DSE:$(E \bigcup P) \times A$ and DSE:$(E \Delta P) X A$ estimators using PES-A gave a population estimate of 75 people, similar to that had movers been ignored. The estimated coverage rate was 80% (81% if movers were ignored). Because so many people moved into the site and so few moved out, when PES-B is used, (1) would estimate that $\hat{x}_{100i} = -5$, so we take $\hat{x}_{100i}$ to be 0. Under assumption (2), $\hat{x}_{100i} = 1$. These gave population estimates of between 91 and 93, with coverage rates of about 71%.

In contrast, a DSE in this site would have estimated that 3.00 people were missed by both census and PES. The population estimate would have been 69, with a coverage rate of 87%.

In both sites, the DSE underestimated the population and overstated the coverage rate, when compared with any of the triple system estimates. Including movers into these estimates was one reason for the larger population estimates and smaller coverage rates when compared to a DSE. A second reason was the greater number of people found by the ethnographers who were missed by both the census and PES, as compared to the DSE estimate of the number missed by census and PES.

In each site, the ethnographers found 8 non-movers missed by both census and PES. In the first site, the 8 people consisted of 6 households, of which 2 households (3 people) were missed completely by the census and PES, and 4 households (5 people) had other members who were found by the census and/or the PES. In the second site, the 8 people came from 6 households, of which 5 households (7 people) were missed completely by the census and PES, and 1 household (1 person) had other household members who were found by the census and/or PES. In addition, in this site the ethnographers found 5 in-movers who were missed by both census and PES.

3.2 Statistical Dependency of the Census and PES

The odds ratio in the fully observed table of non-movers who were on the ethnographers list, was 7.41 in the first site, and 2.81 in the second site. Under the assumption of independence of the census and PES, this number would be about 1. Ninety-five percent confidence intervals for the odds ratio were generated under 6 methods. Three methods modeled the individual as the unit, and did not use poststratification. In the first method (a parametric bootstrap), the counts in each of the 4 cells were modeled as being Poisson random variables with means equal to observed counts. One observation was drawn from each of the 4 cells, the odds ratio for that table was calculated, and this was repeated 100,000 times. The resulting odds ratios were sorted, and 2.5 percent of the observations from each tail were removed. The remaining range of the observations gave a 95 percent confidence interval.

In the second method (a non-parametric bootstrap), a sample was drawn with replacement from the observed counts in the fully observed subtable. The odds ratio was computed and a 95 percent confidence interval was created in the manner just described. The third method was a Fisher exact confidence interval, calculated using StatXact (Mehta and Nitin 1991).

In each site, these 3 intervals were very similar within each site. In the first site, none of the intervals included 1 (and in fact none included 2), whereas in the second site, the intervals did include 1.

There are several reasons why these first 3 methods might overestimate the statistical dependency of the census and PES. One reason is that the individuals within a site may come from a mixture of subpopulations, each with different probabilities of capture by the 3 sources. If this is the situation in these sites, calculating the odds ratios separately for the different subpopulations and then calculating a common odds ratio for the subpopulations should result in a more accurate estimate. We attempted to address this issue by poststratifying on first 1, then 2 variables.

The fourth method used poststratification on age, using year of birth (before 1960, and 1960-1990). In each site, this separated the individuals into 2 nearly equal groups (with 1 missing year of birth in the second site). An exact confidence interval for the common odds ratio was calculated. The fifth method used poststratification on age and race. The 95 percent confidence intervals for the common odds ratio using 1 or 2 poststratifications were very similar to each other and to the interval when poststratification was not used, although intervals using poststratification had a smaller lower endpoint than intervals without poststratification.

A second reason that the odds ratio may have been overestimated is that individuals within a household may not be independent. If one individual is found by all 3 sources for example, it is likely that other individuals in the household were also found by all 3 sources. To address this issue, the sixth method of generating confidence intervals for the odds ratio used the household as the unit of analysis. Households were sampled with replacement from the households within the site, and all individuals within each sampled household were included in the sample. The odds ratio for the individuals was calculated. This was repeated 10,000 times, then the odds ratios were sorted and the range of the middle 95% of the odds ratio gave the 95% confidence interval. As expected, the clustering of individuals within households gave a wider interval. The interval for the first site was (1.52, ∞), while for the second site it was (0.51, 24.96).

It should be noted that while the odds ratios in the subtable of people on the ethnographers lists may not be 1, the observed odds ratio in the other fully observed subtables in both sites are even further from 1. This indicates that the ethnographers' lists are not at all independent from the census or PES.

4 Discussion

Some aspects of the ethnographic study contributed to initial errors and uncertainties in the data. In particular, the coding required of the ethnographers was very meticulous, making room for errors, and errors were made. The error rate may be reduceable in future years if the coding were simplified. Also, the ethnographic data was linked to the census and PES data at a later time, so the ethnographers were unable to comment on some uncertain PES cases. In addition, the ethnographers did the initial 3-way matching, but they are not trained matchers and the results didn't necessarily agree with the results that trained matchers would have obtained. It should be noted that after the initial matching by ethnographers, some of the data was then clerically matched, and the entire data set was then reviewed and corrections were made.

In estimates based on PES-B, whether based on administrative or ethnographic lists, x_{100i} is unobservable and must be estimated. The assumption underlying estimator (2) is not likely to be accurate, as we would expect that the PES coverage rate for in-movers is smaller than that for non-movers. This has the effect that our estimates of x_{100i} and ultimately of x_{000} are too small. One of the assumptions underlying estimator (1), that the number of in-movers equals the number of outmovers between census day and the PES, may be inaccurate, especially when the number of movers in an area is small. This was the case for the second ethnographic site examined here. However, unless there are systematic population shifts between Census and PES time, (1) is likely to be less biased than (2) on the average.

Both PES-A and PES-B estimates using ethnographers' lists may lead to underestimation of x_{000} (Thurston and Zaslavsky 1994). It should be noted, however, that PES-A data were not considered critical in the 1990 ethnographic program. Consequently estimates based on PES-A may be less reliable than those based on PES-B.

Census experience shows that the non-match rate among movers is typically much greater than among non-movers (Schafer 1991). Although a large part of the reason for this high non-match rate is due to matching error, movers may be more prone to both over and undercounting (Citro and Cohen 1985, chapter 5). One way to improve the population estimates may be to consider movers separately from non-movers, drawing inferences about movers only from the mover population. In areas with a large number of movers, a separate triple system estimate for movers, combined with a triple system estimate for non-movers, may lead to more accurate estimates. However, when the number of movers is small, we would expect a large sampling variability from estimates based on movers alone. In this case, some way to pool estimates of movers across similar poststrata would be desirable.

5 Conclusions

Use of a triple system estimator is likely to lead to more accurate estimates of population and of census coverage rates than are possible using a DSE. In addition, an estimate of the statistical dependency of the census and PES is possible using triple system estimation. Proper consideration of movers may give more accurate estimates of population and coverage rates than when movers are dropped from the roster.

Estimates based on PES-A and estimates based on PES-B are both likely to give underestimates of x_{000}. If the coverage rate among people in the PES is of interest, estimates should be based on PES-B so that movers are included in this estimate.

Since estimates based on administrative lists tend to underestimate x_{000} to a greater extent than estimates based on ethnographers lists for both PES-A and PES-B (Thurston and Zaslavsky 1994), using ethnographers' lists is preferable to using administrative lists when possible. A larger sample size within each site would be desirable in order to obtain more precise estimates, and some form of random selection would be necessary if estimates are desired for a larger population. Due to the difficulties in collecting and processing ethnographic data however, it may be necessary to use administrative lists or some other source of names and addresses when large sample sizes are needed. Using an updated administrative list containing changes to the earlier version for a follow-up at a later time should help to improve the accuracy of estimates based on administrative lists.

Including movers did affect estimates of population and coverage rates in the two ethnographic sites considered here. Evidence from one of these sites strongly suggests that the census and PES are not independent. In both sites, using a DSE instead of a triple system estimator would have led to an underestimation of the population and overestimation of the census coverage rates.

6 References

Alvey, W. and Scheuren, F. (1982), "Background for an Administrative Record Census", *American Statistical Association Proceedings of the Social Statistics Section*, 137-146.

Brownrigg, L. and de la Puente, M. (1992) "Alternative Enumeration Methods and Results", *American Statistical Association Proceedings of the Section on Survey Research Methods*, 199-204.

Citro, C.F. and Cohen M.L., editors (1985), *The Bicentennial Census: New Directions for Methodology in 1990*, Washington, D.C.: National Academy Press

de la Puente, M. (1993), "A Multivariate Analysis of the Census Omission of Hispanics and Non-Hispanic Whites, Blacks, Asians and American Indians: Evidence from Small Area Ethnographic Studies", *American Statistical Association Proceedings of the Section on Survey Research Methods*, 641-646.

Diffendal, G. (1988), "The 1986 Test of Adjustment Related Operations in Central Los Angeles County", *Survey Methodology*, 14, 71-86.

Hamid, A. (1992), "Ethnographic Follow-up of a Predominantly African American Population in a Sample Area in Central Harlem, New York City: Behavioral Causes of the Undercount of the 1990 Census ", *Ethnographic Evaluation of the 1990 Decennial Census Report Series, Report 11*, Center for Survey Methods Research, Bureau of the Census, Washington D.C.

Hogan, H. and Wolter, K. (1988), "Measuring Accuracy in a Post-Enumeration Survey", *Survey Methodology*, 14, 99-116.

Lerch, P.B. (1992), "Coverage Differences in the Census of a Rural Minority Community in North Carolina: The Little Branch Area of the Waccamaw Sioux Tribe", *Ethnographic Evaluation of the 1990 Decennial Census Report Series, Report 20*, Center for Survey Methods Research, Bureau of the Census, Washington, D.C.

Marks, E.S., Seltzer, W. and Krotki, K.J. (1974), *Population Growth Estimation: A Handbook of Vital Statistics Measurement.* New York: The Population Council.

Martin, E. and de la Puente, M. (1993), "Research on Sources of Undercoverage Within Households", *American Statistical Association Proceedings of the Section on Survey Research Methods*, 1262-1267.

Mehta, C. and Nitin, P. (1991), *StatXact User Manual, Version 2*, Cambridge, MA: Cytel Software Corp.

Schafer, J.L. (1991), "A Comparison of the Missing-Data Treatments in the Post-Enumeration Program", *Journal of Official Statistics*, 7, 475-498.

Thurston S.W. and Zaslavsky A.M. (1994), "Using Ethnographic Data to Evaluate Dual System Estimates" unpublished report, Harvard University Department of Statistics, Cambridge, MA.

Vigil, J.D. (1988), "Counting the Hard-to-Enumerate Population", *Proceedings, Bureau of the Census Annual Research Conference*, 25-27.

Wingerd, J. (1992a), "Urban Haitians: Documented/Undocumented in a Mixed Neighborhood", *Ethnographic Evaluation of the 1990 Decennial Census, Report 7*, Center for Survey Methods Research, Bureau of the Census, Washington, D.C.

Wingerd, J. (1992b) "Triple Match Analysis: Mixed Urban Haitian", unpublished report, Center for Survey Methods Research, Bureau of the Census, Washington, D.C.

Zaslavsky, A.M. and Wolfgang, G.S. (1993), "Triple System Modeling of Census, Post-Enumeration Survey, and Administrative-List Data". *Journal of Business and Economic Statistics*, Vol. 11 No 3, 279-288.