ESTIMATION IN THE 1995 CENSUS TEST SERVICE-BASED ENUMERATION

David L. Ferraro, U.S. Census Bureau
Room 3784-3, DSSD, U.S. Census Bureau, Washington, DC 20233

Key Words: No Usual Residence, Coverage Improvement, Sampling, Estimation

I. INTRODUCTION

The 1995 Census Test was the first attempt at a fundamentally different approach to counting persons without a usual home than was used in the 1990 Census. The new methodology enumerates people at facilities where they receive services. The 1995 service-based enumeration (SBE) counted people at shelters and soup kitchens. The goal of the SBE project is to test operational methods and estimation methodologies to include in the census persons who use services and may be missed in the standard enumeration of households and other group quarters. The methodology was not designed to provide a count of the homeless population or service users.

We conducted a complete enumeration of shelters on the evening of March 6, 1995. We then conducted a complete enumeration of soup kitchens that provided services on March 7, 1995. This was done at all three of the Census test sites (Oakland, CA., Paterson, NJ. and six parishes in northwest Louisiana). McNally (1996) provides an evaluation of the methods used in this enumeration. For the purposes of this paper, I will refer to this enumeration as the Initial enumeration.

For the research on estimation, a secondary enumeration was conducted. We made a follow-up visit a week later at a sample of shelters and soup kitchens in Oakland only. We enumerated a sample of clients who used services within the sample of sites. I will refer to this as the Follow-up enumeration. Wellens and Gerber (1995) provide an evaluation of operational procedures used in the Initial and in the Follow-up enumerations.

The objective of the Follow-up enumeration is to test different statistical methodologies for estimating the number of people without a usual residence who used services during the enumeration period. There are three goals:

1. Determine the operational feasibility of the follow-up enumeration.
2. Determine the coverage improvement of the follow-up enumeration.
3. Test various estimation methodologies.

II. PROCEDURAL METHODOLOGY

The Initial enumeration was designed to include in the census people who used services and were missed in the enumeration of housing units and group quarters. It collected basic demographic and service-usage information. The Initial enumerations were unduplicated and the results used in both the 1995 Census Test results and in the estimation.

For the Follow-up enumeration, Thompson (1995) selected a stratified systematic sample of 20 service provider sites from the complete list of sites. Within these 20 sites, we selected a systematic sample of persons. Based on previous information from service providers in Oakland, about 800 persons were expected to stay in shelters at the time of the Census Test, and about 500 (with some overlap) were expected to use soup kitchens. It was estimated that a 20% follow-up sample would be sufficient to assure a CV less than 20%. Sampling rates varied from 20% to 100% depending on the size and type of provider. The overall expected sampling rate was about 22%.

Both enumerations had questionnaires that asked name, date of birth, Spanish/Hispanic origin and race. In addition, they both asked the service-usage questions (questions not shown either gave instructions or asked for addresses):

6. Which of the following best describes the place where you stay overnight most of the time?
8. Which of the following best describes the place where you stayed last night?

The Follow-up asked the additional service-usage questions:

11. The Census Bureau has been counting people in a census in this area. Did you complete a census questionnaire in a shelter, soup kitchen, or other service location on Monday, March 6, or Tuesday, March 7, 1995?
12. How many days during the past week, including today, have you visited shelters, soup kitchens, or other service locations in this area?
13. Which of the following best describes where you stayed Monday night, March 6, 1995?
15. Did you receive any free prepared meals in a soup kitchen, or from a mobile food van on Tuesday, March 7, 1995?

The same procedures were employed for the Follow-up as

the Initial enumeration with the exception of taking a sample of people. Any addresses given in response to question #6 were matched to the standard enumeration files. Matches were dropped from the SBE universe.

The original plan was to conduct two follow-up visits. The first follow-up would have been another complete enumeration in all three areas. The second one being a sample of service sites and a sample of persons in all three areas. In the end, only the second follow-up was done and only in Oakland. Since no other urban or rural areas were included, any conclusions drawn from these data will apply only to Oakland and cannot be generalized to other areas.

The original list of service providers in Oakland was not verified just prior to Census Day as was planned. This meant that only 17 of the original 32 sites were found and inscope on the Initial enumeration day. The original list of 32 sites was used to select the sample of service providers for the Follow-up visit. Because of this, out of the original 20 providers selected for the Follow-up, only seven sites were visited and had completed questionnaires.

III. UNDUPLICATION

Two unduplication procedures are required. The first unduplication is within the Follow-up cases. Since people could be enumerated at multiple sites, we clerically matched cases within the Follow-up sample by name and date of birth. Each questionnaire was classified into either having sufficient information to match or insufficient information to match. The first group was used to unduplicate, but the second group can still be used for estimation purposes.

We could not match any questionnaires within the Follow-up sample by using names and dates of birth as match keys. Table 1 summarizes the results of the SBE Initial Enumeration (all sites) and Follow-up unduplication. We included the initial counts for comparison purposes.

Table 1

Questionnaire Status	Initial	Follow-up
After census match	956	129
Sufficient data	467	88
Duplicate people	19	0
After unduplication	448	88
Insufficient data	489	41
Total	937	129

The second match was the Follow-up sample to the Initial Enumeration to get a count of people enumerated both times. The same procedures were used as before. Table 2 summarizes the Initial Enumeration to Follow-up Match using Questionnaires with sufficient data. The heading of "Insufficient" refers to questionnaires with no name and date of birth to match.

Table 2 Comparison of Follow-up to Initial

	Initial Enumeration			
Follow-up	In	Out	Insuffi-cient	Total
In	17	71	41	129
Out	431			
Insufficient	489			
Total	937			

We did the matching with all the sites in the Initial enumeration, i.e., we matched the 88 cases to the 448 cases.

IV. NONRESPONSE AND DATA EDITS

A critical aspect of all the estimators is that complete and matchable data are collected for each person captured. Nonresponse in this case has significant impact for two reasons. Including cases with insufficient data in both systems leaves the possibility that individuals aren't counted as matches between enumerations. Secondly, it affects the unduplication within each enumeration by not being able to identify duplicates and double counting. Both of these factors lead to potentially large biases of the population estimates.

As the above tables indicate, there was a high nonresponse rate of complete names in the Initial (52.2%) and in the Follow-up (31.8%). However, it was not uniform across all sites. For example, all 41 cases in the Follow-up that did not have complete names came from two shelter sites. Although not as prevalent, a similar trend exists in the Initial Enumeration. Out of the 17 sites, eight sites had all the cases with incomplete names. One other site had one incomplete name out of 33 cases. This means from the other seven sites, 488 (65%) cases had incomplete names and 230 cases had complete names. Even ignoring the one soup kitchen that had many more clients than expected (418 incomplete cases), there was a 31% nonresponse rate of complete names in the six sites.

A related issue that is critical to some of the estimators is item nonresponse in the Follow-up to the service-usage

questions. Certain estimators are dependent on these questions rather than on matching. Forty-nine percent of the questionnaires had nonresponse on all the service usage questions. It was also obvious from microdata review that some questionnaires were not filled out properly and some editing was required. Some boxes were left blank even though the space for a write-in was filled. This was consistent with Wellens (1995) who said skip patterns were not properly followed.

Some of the service-usage nonresponse could be alleviated by editing the missed skip patterns. For Questions #6, #8 and #13, incorrect responses were marked based on skip patterns and write-in answers.

After editing the file for these missed skip patterns, the nonresponse rates for the service-usage questions drops to 42% (75 complete service-usage responses). Again, this nonresponse rate was not uniform across sites. The two shelters that had all the incomplete names had 83% of the service-usage nonresponses while having 57% of the total number of questionnaires.

In addition, there was nonresponse in the demographic characteristics. There were eight missing values for Race. Six of these were item nonresponses which listed Mexican, Mexican-Am., Chicano as Spanish/Hispanic origin and Race blank. The other two listed "MEXICAN" and "MIXED", respectively. For this analysis, we classified all eight cases as Nonblack.

There were seven missing values for sex. All seven cases had names listed. We imputed sex for all seven based on first names. At this point, there is no model for predicting sex or race for the imputed cases.

Hogan and Cowan (1980) suggest that the simplest method to handle nonresponse is to subtract the questionnaires with insufficient data to match from both systems. Although this estimate does not use all reported data (it is data that is questionable) it avoids the possible increased variance and larger bias by making additional errors in the adjustments to the insufficient data.

Another approach is to use all unduplicated questionnaires from the Initial (ignoring the unduplication bias) and not using the questionnaires with insufficient data in the Follow-up. This way all first captures are used in the estimation.

For cases with insufficient data a nonresponse adjustment can be applied to the estimators. This too has problems. There is an implicit assumption that those with insufficient data have the same match rates and correct enumeration rates as those with sufficient data. This may be a faulty assumption with this population. It is further complicated by matching the two data sets. A nonresponse adjustment is beyond the scope of this paper and is not covered.

V. DATA CONSISTENCY

The questionnaire included service-usage questions that we can use to check the consistency of the usage answers each person gave. They also give a check of our matching results of who we should have matched based on their responses versus who we actually matched.

There are four service-usage questions that relate to each other as to whether the person should have been enumerated in the Initial. Comparison of the responses to these questions, show how many people we should have enumerated in the Initial and matched to the Follow-up:

- 17 matched by name
- 29 answered "Yes" to #11
- 42 answered "Every day" to #12
- 60 answered "Shelter for homeless persons" to #13 or "Yes" to #15

This suggests that our match rate was much less than the service-usage responses and the respondent's recall of being enumerated was less than of using services. There are many possible reasons for this. We did not collect sufficient information to match from the Initial. There could be response errors to the service-usage questions. Also since we only visited soup kitchens at their largest meal of the day, people could have used their services at a time when enumerators weren't present.

VI. STATISTICAL METHODOLOGY AND ESTIMATORS

Three classes of estimators are considered in this paper. The first class, based on capture-recapture methods, matches results from samples for two time periods to produce dual system estimates. The second type of estimator is a multiplicity estimator which relies on respondents' answers to "service-usage history" questions. A third estimator weights the data according to the case's first enumeration.

Service providers in Oakland provided estimates of the number of people who use their services. About 800 persons were expected to stay in shelters at the time of the Census Test and about 500 (with some overlap) were expected to use soup kitchens. This gives an estimate of the total population in the range from 800 (100% overlap) to 1300 (no overlap). It is unclear how accurate this estimate is since the original list of provider sites was not reliable and the estimates varied from our enumeration counts.

To make it clearer as to which subpopulation is being referred to for each estimator, define the following:

n_{11} = all Initial cases (937)
n_{12} = Initial cases with sufficient data to match (448)
n_{21} = all Follow-up cases (129)
n_{22} = Follow-up cases with sufficient data to match (88)
n_{23} = Follow-up cases with sufficient data to match and complete service-usage (75)

A. Dual System Estimator (DSE)

Schindler, Griffin and Navarro (1993) consider a DSE using a capture-recapture model. Information about persons in either the first or second capture, in both captures, or missed in both captures is used to produce an unbiased estimate of this population group.

For the series of two samples with sizes n_{1i} and n_{2j}, i=1,2 and j=1,2,3 from a total population of size N, let m_2 be the number of persons in the second enumeration who have been enumerated in the first enumeration.

Then the "Peterson estimate" equivalent to the DSE is

$$\hat{N} = \frac{n_{1i} \cdot n_{2j}}{m_2}$$

There are four assumptions (Wolters, 1986) that are made for the DSE.

1. The population is closed. That is there is no change in the total population during the reference period.
2. There is independence between the events of being enumerated in the Initial enumeration and the Follow-up enumeration.
3. Persons counted more than once on any given day can be identified and unduplicated.
4. Neither enumeration list contains duplicates for population members.

The third and fourth assumptions may not be valid since there is such a high percent of questionnaires with insufficient data to unduplicate in the Initial enumeration alone. Given this, the estimator is biased and will be an overestimate of the true population.

Using populations n_{12} and n_{22} given earlier in Table 2, the estimate is:

$$n_{12}=448 \qquad n_{22}=88 \qquad m_2=17$$
$$\hat{N} = 448 \cdot 88/17 = 2319$$

A second DSE uses an alternative capture-recapture method that uses the response to question #11 of whether they were counted in the Initial enumeration. Matching is not now necessary. Now, let m_2 be the number of persons who answer "yes" to this question. With this DSE, all the data from the Initial enumeration (n_{11}) can be used since we aren't matching and the service-usage question doesn't depend on answers from the Initial. Only the cases with a complete answer to question #11 (n_{23}) was used from the Follow-up. The second assumption is now that there is no response error.

Then the estimate is:

$$n_{11}=937 \qquad n_{23}=75 \qquad m_2=29$$
$$\hat{N} = 937 \cdot 75/29 = 2423$$

A third DSE uses responses to service-usage questions. This one assumes that if the persons said they used services (shelters or soup kitchens in questions #11 and #13) on the days of the Initial enumeration, then they should have been counted and matched. Populations n_{11} and n_{23} are used.

Then the estimate is:

$$n_{11}=937 \qquad n_{23}=75 \qquad m_2=60$$
$$\hat{N} = 937 \cdot 75/60 = 1171$$

To directly compare the estimators, the same subpopulations for each estimate is required. See Table 3 for a comparison of all the estimates. Estimates using different subpopulations are given for comparison purposes only. They artificially inflate or deflate the estimate.

B. Multiplicity Estimator

The second estimator is a multiplicity estimator that Schindler, Griffin and Navarro (1993) also consider that does not rely on matching, but rather on respondents' answers to a service-usage question about the number of times a client uses services. By determining how often a client uses services over a given time period and assuming that the service is utilized on a routine basis, an unbiased estimate of this population can be developed. The frequency of usage is the response to question #12.

In this case, let H be the number of days in a week and T be the number of sites in the area. Let M_{ij} be the number of persons enumerated at site j on day i. Let G_{ij} be the probability of selection of persons enumerated at site j on day i. Let A_{ijk} be the number of days person k enumerated at site j on day i reports using sites in the area. Generally, a two-stage cluster sample of h days and, for each day, a sample of n sites are selected. Due to budget constraints in this case, we only selected a sample of n sites for one (h=1) day.

Then the estimator is

$$\hat{N}=\frac{H\cdot T}{h\cdot t}\sum_{i=1}^{h}\sum_{j=1}^{t}\frac{1}{G_{ij}}\sum_{k=1}^{M_{ij}}\frac{1}{A_{ijk}}$$

Correct responses are critical for unbiased estimates. Even small response errors can lead to large biases. An alternative is to ask their usage over the past month. However, this may lead to even larger response errors due to the longer reference period.

Assumptions one and four made in the previous section apply here also.

In this case, let H=7 be the number of days in a week and T=17 be the number of sites in the area. Let h=1 since there was only one day of Follow-up sample. G_{ij} was either 0.2 or 0.333. This estimator only can use population n_{23} from the Follow-up. In this case, n_{23}=74 since there was one response of "don't know".

Using n_{23}, the estimate is:

$$\hat{N}=\frac{7\cdot 17}{1\cdot 7}\cdot(7.8+1.0+6.3+15.5+24.6+15.7+7.8)=1336$$

The multiplicity estimator is generally designed to have a sample of sites visited on multiple days. These data do not have that as only one day of sample interviews were collected at one sample of sites.

There are some inherent biases with this method when using a reference period of a week depending on the person's frequency of use. For example, if the persons visits a site every other day, this question will report four of every seven days which is an over count. In actuality it is one of every two days.

C. Reweighting Estimator

A third estimator is the reweighting estimation model. This measures additional coverage by assigning a specified weight associated with the first enumeration of a particular client and disregards any subsequent enumeration of that same client. A summation of the weights associated with the captured individuals will result in an estimate of this population group.

In this case, let T and t be the number of sites in the Initial and Follow-up, respectively. Let n_{1i} be the number of persons enumerated at site i in the Initial and let n_{2i} be the number of first enumerations of persons at site i in the Follow-up. Let W_1 be the weight of each person in the Initial and W_{2i} be the weight of each person at site i in the Follow-up. Then a simple estimator that ignores the data collection problems is

$$\hat{N}=\sum_{i=1}^{T}(n_{1i}\cdot W_1)+\sum_{i=1}^{t}(n_{2i}\cdot W_{2i})$$

All four assumptions made for the DSE apply to the reweighting estimator also.

Each person in the Initial enumeration has a weight of W_1=1.0 in the simple case since each was a first enumeration assuming we did a complete enumeration. In the Follow-up there was a two-stage sample design. Assuming a systematic sample of sites and a systematic sample of clients the weights (W_{2i}) for each site are derived from the probabilities of selections. The probability of selection for each site is what was planned for the Initial and Follow-up. It varied from 0.625 to 1.0. The sites that were

Table 3 Comparison of Estimates

Initial	Follow-up	DSE #1 (matching)	DSE #2 (recall)	DSE #3 (site-usage)	MULTIPLICITY	REWEIGHTING
n_{11}	n_{21}					1461
n_{11}	n_{22}					1271
n_{11}	n_{23}	4133	2423	1171		1220
n_{12}	n_{21}					
n_{12}	n_{22}	2319				782
n_{12}	n_{23}	2240	1158	560		731
	n_{23}				1336	

actually visited both times differed significantly. This will likely introduce a large bias into the estimate.

Ignoring nonresponse in the initial enumeration and using n_{11}, since they are all first enumerations, and using only the first enumerations from n_{22} of the Follow-up, since that is used for matching, is the most logical estimate. This estimate is biased but it's not clear which way from the true population due to the two different biases. It is an overestimate because people are possibly counted in their subsequent enumeration and an underestimate because only those with sufficient information are counted from n_{22}.

The estimate is:

$$\hat{N}=937+[63+36+76.8+65+62.4+31.5]=1271$$

See Table 3 for a comparison of all the estimates. Again, estimates using different subpopulations are given for comparison purposes only.

VII. CONCLUSION

In order for the SBE approach to give an accurate count of the population, better data has to be collected. Procedures need to be followed closer so that complete data are collected from each person at all the sites. The estimates from the 1995 Census test give a wide range of values due to the data problems that were encountered which makes an evaluation of the estimators difficult. It is also not clear what to compare the estimates to since the provider estimate was not reliable. Any conclusions drawn from these test data should be about the validity of the estimators and not the estimates. The estimators can be evaluated based on consistency between the estimates.

All of the estimators have assumptions that are violated due to the data collection problems. All the estimates are biased. The systematic sample of provider sites in the Follow-up in actuality was not a representative sample. It is also likely that both lists contain duplicates.

Based on comparison to the service-usage data, matching by name and date of birth gave very low match rates. This is also evident by the DSE estimate being twice as large as the closest one. More matchable data needs to be collected for the DSE to give a reliable estimate. By the same comparisons, using respondent recall (Question #11) and the DSE also gives large estimates and standard deviations. Using service-usage responses (#12 and #14) gives estimates that are in line with other estimators, but this assumes that if a person said they used services a week prior that we were able to capture them at that time. It says nothing of whether this is true or not.

The DSE (based on service-usage), multiplicity and reweighting all gave estimates in the range of 1150-1350 using n_{11} and n_{23} as a base. This was the only consistent result and also was in line with the provider estimate.

Research continues on improving and simplifying the questionnaire and procedures in order to collect better data. We are also researching using the multiplicity estimator with just a single enumeration.

III. REFERENCES

Hogan, H., and Cowan, C.D., (1980), "Imputations, Response Errors, and Matching in Dual System Estimation", Proceedings of the Section of Survey Research Methods, American Statistical Association, pp 263-268.

McNally, T., "Service-Based Enumeration Coverage Yield Results Evaluation", DMD 1995 Census Test Memorandum Series No.27, February 20, 1996.

Schindler, E., Griffin, R., and Navarro, A. (1993), "Sampling and Estimation for the Homeless Population", Proceedings of the Section of Survey Research Methods, American Statistical Association, pp 468-473.

Thompson, John H., "Sample for Follow-up for Persons Without a Usual Residence for Oakland", DSSD 1995 Census Test Memorandum Series K-3, April 19, 1995.

Wellens, Tracy R., and Gerber, E.R., "Evaluation of the Operational Procedures for the Service-Based Enumeration of the 1995 Census Test", DMD 1995 Census Test Memorandum Series No. 13, November 2, 1995.

Wolter, K.M., (1986), "Some Coverage Error Models for the Census Data", Journal of the American Statistical Association, Volume 81, Number 394, pp 338-346.