Multiplicity Estimation for Service Based Enumeration in Census 2000

By

Felipe Kohn
Randal ZuWallack
Richard Griffin
U.S. Bureau of the Census
Washington, D.C. 20280

Multiplicity Estimation for Service Based Enumeration in Census 2000

Felipe E Kohn,   Randal Zuwallack and Richard Griffin
U.S. Census Bureau

This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress.

Introduction

The Service Based Enumeration (SBE) is designed to provide people with no usual residence an opportunity to be enumerated. The U.S. Census Bureau planned special procedures to enumerate such persons at shelters, soup kitchens, mobile food vans, and certain outdoor locations with no apparent means of shelter, referred to as TNSOLs (Targeted Non-sheltered Outdoor Locations). Because the SBE can only account for people at these facilities on the day of enumeration, the Bureau planned to apply multiplicity estimation to account for people who use these facilities but did not use them on the day of the SBE. The uncorrected census count for this population includes only those people actually enumerated (after unduplication) at these service locations on this one day. The corrected count includes the enumerated persons plus the estimated number of people who use these services but not on the day of the enumeration.

Although the SBE operations were completed for Census 2000, the Census Bureau decided NOT to use Multiplicity Estimation to estimate the number of people who use services but not on the day of enumeration. Instead only the enumerated persons are included in the total U.S. population count.

The paper discusses the enumeration procedures, the multiplicity estimator, the results of the enumeration and estimation as well as the decision not to use multiplicity estimation.

Background

In 1990, the Census Bureau conducted a special operation, called S-Night, which was designed to locate and enumerate persons without usual residence including those who use shelters, soup kitchens and other targeted service locations. However, the data collection procedures were entirely different from those for the Census 2000 SBE and no estimation techniques were used. The Census Bureau felt it could improve on the 1990 methodology, thus the Census 2000 SBE and multiplicity estimation procedures were devised. Because it is a new system, we do not have any estimation results from 1990 to compare to Census 2000 SBE multiplicity estimates.

As part of Census 2000, Census Bureau enumerators visited shelters to collect information on the 27th of March and the following day went to collect information at soup kitchens and mobile food vans. Approximately 280,000 persons were enumerated in over 14,000 locations (51% shelters, 15% soup kitchens and mobile food vans and 34% TNSOLS) during the SBE. The enumeration used two modes of data collection, self-administered questionnaires at shelters and enumerator-administered at soup kitchens, mobile food vans and TNSOLS. The results from the locations are then compared and unduplicated since persons can be enumerated in multiple places.

During the SBE data collection operation, respondents were asked the number of times in a week that they use that type of facility:

- In shelters, respondents were asked "Including tonight, how many nights during the past seven nights did you stay in a shelter?"

- In soup kitchens and mobile food vans, respondents were asked "Including today how many days during the past seven did you receive a meal from a soup kitchen or mobile food van?"

The responses to these questions range from 1 to 7 or a nonresponse. The multiplicity estimator is based on the responses to these usage questions. Additionally, persons enumerated in soup kitchens and mobile food vans were asked the question "Including last night, during the past seven nights did you stay in a shelter?" Theoretically, persons who answered "yes" to this question were already accounted for in a shelter, thus steps are taken to ensure these persons are not double weighted.

Multiplicity Estimator

Multiplicity estimation for SBE uses service usage responses to adjust enumeration counts for persons who use services but not on the day of enumeration. A person who only uses a service one night out of the seven nights in a week is given a weight of 7 since they only had one chance in 7 to be enumerated on the day selected for enumeration. Conversely, a person who uses a service all seven nights in a week is given a weight of 1 since they would be enumerated no matter what day was selected for enumeration. Although March 27, 2000, a Monday, was not randomly selected we are implicitly treating it as representative. For nonrespondents to the usage questions, the responses required for multiplicity estimation are imputed. The multiplicity estimator of persons without usual residence that use services has the formula shown below:

$$\hat{X} = \sum_{j=1}^{n} \frac{7}{A_j} + \sum_{j=1}^{m} \frac{7}{B_j} + X_{TNSOL} + X_{BCF}$$

where   n = the number of persons enumerated at shelters
m = the number of persons enumerated at soup kitchens or mobile food vans who said they did not use shelters in the last 7 nights
A_j = number of nights in the last 7 nights person j used a shelter
B_j = the number of days in the last 7 days person j received a meal from a soup kitchen or mobile food van
X_{TNSOL} = the number of persons enumerated at Targeted Non-Shelter Outdoor Locations (TNSOL)
X_{BCF} = the number of persons without a usual residence enumerated on Be Counted Forms (BCF)

Unduplication of all SBE persons including BCF persons without a usual residence was completed prior to estimation.

Decision not to use Multiplicity Estimation

We have evidence, cited below, that the level of response bias to these usage questions is unacceptably high. In addition the nonresponse rates to the usage questions are higher than we normally accept for most surveys. Thus, we decided not to correct the count of persons actually enumerated in SBE using the multiplicity estimate which relies on the responses to these usage questions. Due to the highly biased responses that we received in the shelters, we decided that the multiplicity estimator should not be applied to this population; thus, only those persons actually enumerated were included in the population count.

Note: this is not limited to n + m but in addition includes persons enumerated at soup kitchens or mobile food vans who said the DID use shelters in the last 7 nights.

Usage Patterns in SBE Facilities

A review of existing literature on the usage of services targeted to people without conventional housing shows the average length of stay at facilities is more than just a one or two night stay. These findings were confirmed in two studies at the national level, and various studies at the state and local level. Results from the National Survey of Homeless Assistance Providers and Clients (NSHAPC) show, "Examining the living situations of homeless clients during the eight-day period including the day of the interview and the preceding seven days reveals the extreme transiency of many homeless people. During this time period many clients slept in a number of different places, which could include places not meant for human habitation; emergency shelters or transitional housing; or living arrangements such as a house, apartment, or room in which someone is allowed to stay on a temporary basis".

The NSHAPC results show that during the eight day period in question, 73% of the homeless clients slept in one or more varieties of shelters, 54% slept in temporary housing, and 32% slept on the street. Only 34 % of people slept only in shelters, 28 % slept in both shelters and temporary housing, 5% slept in both shelters and on the street, and 6% slept in three places (shelters, temporary housing, and on the streets) during the eight day period. In addition, 7% slept only on the street, 14% slept on the street and in temporary housing, and 6% slept only in temporary housing.

The other national findings described in the study, "Emergency Shelter and Services: Opening a Front Door to the Continuum of Care", shows that across all shelters, the mean length of stay was 71 days and the median was 30 days. However, this varied greatly, from a few shelters with mean stays of less than 5 days to a handful reporting average stays over a year. The 24-hour shelters had the longest median stays, at 30 days; the small number of day-only shelters showed great variability in length of stay. However for 95 percent of these agencies, the average duration of residence per client was 9 months or less."

Results

This section presents results of the estimation at the national level. Table 1 presents the number of persons enumerated during SBE, the national level multiplicity estimate, and the ratio of the multiplicity estimate to the number of persons enumerated by type of facility.

The largest number of people were enumerated in shelters

and the largest ratio is also in that type of facility.

Tables 2 and 3 show the nonresponse percentages and the distribution of responses by facility type.

Table 2 explains in part, the size of the ratios for shelters. The multiplicity estimate is inversely proportional to the usage question response. In every region of the country at least sixty-five percent of the persons enumerated in shelters responded that they stayed just one day in the shelter in the past seven days. We know from the surveys discussed above that most persons use shelters more than 1 or 2 days a week. Thus, the data collected has a severe case of <u>response bias.</u>

A second factor affecting the reliability of the data is the high levels of non-response (to the service usage questions) in both types of service facilities. Over twenty percent of the people enumerated in soup kitchens and shelters did not respond to the usage question(s); the nonresponse rates in Tables 2 and 3 range from 18% to 28%.. Those records are imputed based on the respondents to the usage question. Thus, since the responses were heavily tilted to the single night so are the imputed ones and this fact adversely effects the validity of the imputation process.

Analysis

The main problem for the multiplicity estimation is the responses to the usage question in the shelters. For multiplicity estimation, the usage question responses are used for shelter respondents and soup kitchen respondents who did not use shelters in the last seven nights. Nationally about 71 percent of the persons enumerated in shelters and soup kitchens (after unduplication) were enumerated in shelters and further, persons enumerated in soup kitchens that respond yes to the shelter usage question are not used in the multiplicity estimation process. Thus usage responses in shelters are critical to the multiplicity estimate.

The multiplicity estimate is inversely proportional to the usage question response. Persons responding " 1" get an effective weight of 7, while persons responding "7" get an effective weight of 1. Since nationally 77 percent of the shelter persons responses to the usage question were '1', most of the persons enumerated in shelters had a effective weight of 7. We know from the surveys discussed above that most persons use shelters more than 1 or 2 days a week (no information on frequency of use of soup kitchens was available). In fact, the data suggest that the most frequent response ought to be 7, not 1. Using a preponderance of weights equal to 7 results in multiplicity estimates that are too high. Although we do not have independent data on soup kitchens for comparison, the

responses to the soup kitchen usage questions appear more reasonable.

Rationale for Decision not to use Multiplicity Estimation

The ratio of the multiplicity estimate to the number of persons actually enumerated in shelters (4.25 nationally) is probably too high due to the high percentage of persons responding "1" to the shelter usage question. We feel this percentage is too high based on results from NSHAPC and other national findings discussed above.

Although the total national level multiplicity estimate of nearly 1 million persons is reasonably close to what was expected (based on the surveys discussed above), using the multiplicity estimation results to distribute these persons to local areas and service facilities is not statistically defensible due to response bias to the usage questions, particularly in shelters.

Recommendations

In this paper we have exposed some of the problems that the SBE enumeration had during Census 2000. In its search for a perfect Census, the Bureau will have to do some things different to account and/or estimate this segment of the population. In this section, we present some ideas that can be the beginning of the restructure of the SBE for the 2010 Census and beyond.

- In the 2000 Census Dress Rehearsal the enumeration process worked better in part because of the limited number places where it was held and the control that was established to ensure that all field operations were followed as planned. In the 2000 Census with over five hundred Local Census Offices to manage, the control in the SBE operations was much weaker and as result some operations were not executed as planned. We feel that in future censuses a strict control in field operations is required for the enumeration of the SBE population.

- The response bias in the soup kitchens is likely less than in the shelters (tables 2 and 3). The soup kitchens questionnaires were filled by enumerators while the questionnaires in shelters were self administered; we recommend that the questionnaires for the entire SBE population be filled by enumerators.

- During the 2000 Dress Rehearsal there were some problems with the shelters and soup kitchen

questionnaires and the questionnaires were re-designed addressing those problems; however, due to the results in the Census 2000, more studies should be conducted about the design of the questionnaires and especially how the usage questions are phrased.

Table 1: Ratio of Multiplicity Estimate to Persons Enumerated at the national level

Region	Service Facility (Includes BCFs)	Persons Enumerated (1)	Mult. Estimate (2)	Ratio (2)/(1)
United States	Shelters	183,414	780,369	4.25
	Soup Kitchens	74,033	165,365	2.23
	TNSOLs	23,080	23,080	1.00
	Total	280,527	968,814	3.45

Table 2: Distribution of Responses for Shelters at the Regional Level

Region	Shelter Usage Question (nights in past 7)							Non-Response Rate
	1	2	3	4	5	6	7	
Region	Shelter Usage Question (nights in past 7)							Non-Response Rate
Northeast	64.97%	1.06%	1.98%	1.06%	1.04%	1.04%	28.85%	18%
Midwest	83.24%	2.95%	2.91%	2.95%	2.91%	2.05%	3.00%	20%
South	81.86%	3.00%	3.90%	3.10%	2.10%	2.10%	3.93%	19%
West	81.21%	2.97%	2.94%	2.97%	3.00%	2.03%	4.88%	28%

Table 3: Distribution of Responses for Soup Kitchens at the Regional Level

Region	Soup Kitchen and Mobile Food Van Usage Question (nights in past 7)							Non-Response Rate
	1	2	3	4	5	6	7	
Northeast	29.83%	3.99%	14.94%	6.99%	13.10%	13.95%	17.20%	21%
Midwest	36.03%	3.99%	10.83%	6.89%	11.88%	12.87%	17.50%	27%
South	31.70%	4.94%	12.93%	6.98%	11.06%	13.07%	19.31%	26%
West	43.90%	3.03%	9.01%	5.99%	10.02%	11.17%	16.88%	26%