COMPARISON OF CENSUSPLUS AND DUAL SYSTEM ESTIMATION IN THE 1995 CENSUS TEST

Mary H. Mulry and Richard Griffiths*
Mary H. Mulry, Bureau of the Census, Washington, DC 20233

KEY WORDS: coverage measurement, census undercount, Census 2000

1. Introduction

The Census Bureau tested integrated coverage measurement (ICM) in the 1995 Census Test because the methodology has been expected to reduce the differential coverage error observed in previous censuses. ICM also is expected to reduce overall coverage error. One goal of the 1995 Census Test was to test two methodologies for integrated coverage measurement. The primary issue has been whether a new methodology known as CensusPlus, which uses ratio estimation, is effective. Another goal was to test dual system estimation (DSE) which was used for the 1990 Post Enumeration Survey (PES) as an alternative to CensusPlus with ratio estimation. This paper evaluates the effectiveness of the two methodologies by examining whether they add persons in the traditionally undercounted groups to the census numbers. The evaluations use data collected in the Oakland, CA and Paterson, NJ test sites. The rural test site in northwest Louisiana is not included in the analysis because data were collected for CensusPlus but not for DSE due to budget constraints.

Each estimation methodology has appeared to have its advantages and disadvantages. The CensusPlus method appears as though it can be completed more easily by the December 31 deadline for delivery of the census numbers to the President. However, CensusPlus requires assuming that the reinterview is "truth." While the dual system methodology may take longer to complete the data collection, it assumes that the reinterview is another list of the population, independent from the census list but not necessarily the truth. The data for CensusPlus is not independent from the census roster because the interviewer is allowed to see the census roster for the address while the data for DSE is considered independent because it is collected prior to accessing the census roster. The Census Bureau has experience with the dual system methodology and found that the dual system estimate generally agreed with demographic analysis estimates in 1990. However, the 1995 implementation of the dual system methodology was different in some important ways from the 1990 methodology which could have caused different results (Mulry and Singh, 1994). On the other hand, since the Census Bureau did not have any direct experience with CensusPlus, it was unable to predict the quality of the population numbers which would be produced by the new methodology.

The analysis examines the characteristics of the persons added by the ICM estimation methodologies after nonresponse followup in two ways. One is at the overall site level for Oakland and Paterson, and the other is by block clusters within site. Examining the census numbers at the site level indicates what types of persons are being added on the whole. Investigating whether the block clusters with additions are the ones that are expected to be hard to enumerate provides some insight into the validity of the methodologies.

2. Methodology

2.1 Site Level Analysis

The site level analysis uses the estimates for each poststratum in Oakland and Paterson and aggregates of the poststrata. The poststrata are defined by 7 age categories, 4 race/Hispanic ethnicity categories, and tenure. Estimates for all 56 poststrata are made for Oakland. There are not enough Asian and Pacific Islanders in Paterson to support separate poststrata so they are included in the Non-Hispanic White and Other poststrata. In Paterson, there are 42 poststrata.

To calculate the persons added by ICM, we first have to calculate the estimate of the population after nonresponse followup. After nonresponse followup (NRFU), an estimate of the population size may be made for a site,

$\hat{C}_j$ = post NRFU estimate for poststratum j .

After the ICM, estimates of the population size may be made for a site,

$\hat{C}_j^*$ = CensusPlus estimate for poststratum j

$\hat{D}_j$ = dual system estimate for poststratum j.

Adjustment factors for CensusPlus and for DSE are defined by

$\hat{C}_j^*/\hat{C}_j$ = CensusPlus adjustment factor in poststratum j

$\hat{D}_j/\hat{C}_j$ = DSE adjustment factor in poststratum j.

The adjustment factors are used in forming the estimates for areas within the poststrata by multiplying the adjustment factor by the area's post nonresponse followup estimate. The differences in the poststratum totals for the estimators $\hat{C}_j^*-\hat{C}_j$ and $\hat{D}_j-\hat{C}_j$ indicate whether the

ICM is adding people in the traditionally undercounted groups. Note that the "additions" could also be subtractions. Another comparison is based on the ratio of the differences

$$(\hat{C}_j^+ - \hat{C}_j)/(\hat{D}_j - \hat{C}_j) \quad .$$

This ratio indicates the success of CensusPlus in adding people relative to the success of DSE.

The second aspect of the analysis is a "face validity" check which compares the estimates from CensusPlus and DSE with the alternative estimates of the population size for Paterson and Oakland. Often the assumption is that the larger the estimates of population size, the better. The comparison with estimates based on other methods and data may add validity to this assumption.

2.2 Block Cluster Level Analysis

The block cluster analysis checks whether ICM adds persons in the block clusters considered hard to enumerate. Adding persons in areas considered hard to enumerate reinforces the validity of the methodology. The analysis considers only the ICM block clusters and unweighted data. Block level estimates from the weighted data are not available for this study.

The unweighted data for CensusPlus are assessed by the difference between the resolved rosters, RE_b and the number from the initial enumeration and nonresponse followup, C_b in a block cluster defined by $RE_b - C_b$. The unweighted data for DSE are assessed by the subtracting the self-responses C_b from the sum S_b of the matches, the other correct enumerations, and the nonmatches. This difference reduces to the difference between the nonmatches NM_b and the sum of the erroneous enumerations EE_b and the cases with insufficient information for matching II_b ,

$$S_b - C_b = NM_b - EE_b - II_b.$$

For the block cluster level analysis, the block clusters are assigned two ranks. One ranking is according to $(RE_b - C_b)$. The other ranking is based on $(S_b - C_b)$. The ranks are assigned in decreasing order of the values. The rankings based on $(RE_b - C_b)$ and $(S_b - C_b)$ are compared with two other rankings designed to indicate how hard a block cluster is to enumerate. The comparison of the sets of rankings use the Spearman rank correlation and Kendall's tau. One set of "hard to enumerate" rankings are those assigned by the planning data base using 1990 census data housing and socioeconomic data from the long form (Robinson and Kobilarcik, 1995).

The other set of rankings are calculated by using the estimated probabilities of being enumerated in the census based on conditional logistic regression models developed as part of the evaluation of the 1990 Post Enumeration Survey (Alho, Mulry, Wurdeman, and Kim, 1993). The covariates for four separate models for minorities and nonminorities in central cities in the Northeast and the West come from short form data. With the models appropriate for a site, the probability of enumeration can be estimated for every person in a block cluster. The two indices of difficulty to enumerate used in the analysis are based on the average probability of being enumerated in a block, p_i, and on the number of 1990 enumerations in the block, N_i:

> The average probability of not being enumerated.
> $(1.0 - p_i)$
>
> The expected number of people missed.
> $(N_i (1.0 - p_i))$

Indices defined by the percentage of people who have an enumeration probability lower than a threshold or the number of people who have an enumeration probability lower than a threshold were considered but did not perform as well.

3. Results

3.1 Site Level Results

3.1.1 Estimates and Variances

The estimates of the adjustment factors and their variances from VPLX (Fay, 1990) for the test sites Oakland, CA and Paterson, NJ are shown in Table 1. Because of space limitation, only shown are estimates for the sites, the racial /ethnic subgroups, and owners and renters. Estimates for all poststrata are given in Mulry and Griffiths (1996). These estimates do not include the segment of the population in groups quarters and institutions because ICM estimates are not made for these groups in the 1995 Census Test. Variance estimates for the difference in the adjustment factors for CensusPlus and DSE have not been made although the correlation between them should be high since they use some of the same data. However, the standard errors of the adjustment factors themselves can provide a conservative estimate of the variance of the difference. Even without the variances of the differences, some observations can be made.

Table 1. Adjustment Factors and Their Standard Errors

	Oakland		Paterson	
	CensusPlus	DSE	CensusPlus	DSE
Total	1.005	1.087	1.034	1.137
	(0.017)	(0.020)	(0.012)	(0.019)
Blacks	0.950	1.105	0.952	1.205
	(0.023)	(0.024)	(0.027)	(0.043)
Hispanics	1.132	1.203	1.128	1.138
	(0.038)	(0.044)	(0.018)	(0.016)
Asians & Pacific Is.	0.976	1.075	N/A	N/A
	(0.047)	(0.046)		
Others	1.039	1.007	0.998	1.026
	(0.032)	(0.025)	(0.021)	(0.014)
Owners	1.026	1.060	1.039	1.044
	(0.014)	(0.015)	(0.028)	(0.010)
Renters	0.989	1.107	1.032	1.185
	(0.028)	(0.031)	(0.022)	(0.027)

Unexpectedly, CensusPlus adjustment factors found that the post nonresponse followup estimates were too large in subgroups that have been traditionally undercounted and where DSE found the post nonresponse estimates to be too small. For example, the CensusPlus adjustment factor is 0.989 for renters in Oakland while the DSE adjustment factor is 1.107.

For Blacks, the CensusPlus adjustment factor in Oakland is 0.950 while the DSE adjustment factor is 1.105. For the poststrata considered most difficult, Black male renters, 18-29 years of age, the adjustment factor for CensusPlus in Oakland is 0.863 while the adjustment factor for DSE is 1.104 with standard errors of 0.022 and 0.024, respectively. In Paterson, the CensusPlus adjustment factor for this group is 0.952 while the DSE adjustment factor is 1.205 with standard errors of 0.027 and 0.043, respectively. The CensusPlus adjustment factor is less than 1.0 for all but three of the 14 poststrata for Blacks in Oakland and for all but one of the Black poststrata in Paterson while the DSE adjustment factor for all these poststrata is greater than 1.0. When taking sampling error into account, some of the factors can not be considered significantly different from 1.0, but there definitely is a trend of being less than 1.0

More in line with expectations, both CensusPlus and DSE found that the post nonresponse followup estimates were too small for Hispanics in both Oakland and Paterson. All the CensusPlus adjustment factors for Hispanic poststrata are greater than 1.0 in Oakland and Paterson with the exception of two poststrata in Oakland. The DSE adjustment factors are greater than 1.0 for all the Hispanic poststrata in Oakland and Paterson.

Estimates for separate poststrata for Asians and Pacific Islanders are made only in Oakland. CensusPlus shows post nonresponse followup estimates for Asians and Pacific Islanders, both owners and renters, are too large while DSE indicates that they are too small. CensusPlus and DSE consistently indicate the estimates for some poststrata are too large or too small. They also show opposite directions for other poststrata.

For Others, the adjustment factors for CensusPlus and DSE show mixed results. For some poststrata, they appear to agree well while they show dramatically different results in other poststrata. The primary example is the Oakland poststratum for males and females 0-17 years-of-age who are renters where the CensusPlus adjustment factor is 1.612 with a large standard error of 0.261 and the DSE factor is 0.960 with a much smaller standard error of 0.067. The other example is poststratum for male renters age 50 and over where the adjustment factor is 1.010 for CensusPlus and 1.191 for DSE in Oakland, and 0.753 and 1.019, respectively, in Paterson. The standard errors for both estimates in Oakland are 0.092 while in Paterson the standard errors for the adjustment factors are 0.112 for CensusPlus and 0.041 for DSE.

3.1.2 Differences Between Estimates

The values of the ratio of the difference between the CensusPlus and post nonresponse followup estimates to the difference between the DSE and post nonresponse followup estimates

$$(\hat{C}_j^* - \hat{C}_j)/(\hat{D}_j - \hat{C}_j)$$

is 0.061 for Oakland and 0.250 for Paterson. This ratio indicates that CensusPlus was unsuccessful in adding people in Oakland relative to the success of DSE. In Paterson, CensusPlus had some success in adding people but nowhere near as many as DSE. CensusPlus does not even appear comparable to DSE.

Looking at values of the ratio of the difference for subgroups provides some insight as to where CensusPlus performed so badly. The values of the differences and the ratio of the differences between CensusPlus and DSE are shown by race/ethnic group and by tenure in Table 2 for Oakland and Paterson, respectively. In these tables the ratios of the differences are negative when CensusPlus found people should be subtracted from the post nonresponse followup estimate, but DSE found some should be added. CensusPlus appears to have some success relative to DSE for Hispanics and for owners with ratios 0.650 and 0.430, respectively, in Oakland and ratios 0.925 and 0.895, respectively in Paterson. For Blacks, CensusPlus found large subtractions relative to the

additions by DSE with a ratio of -0.477 in Oakland and -0.232 in Paterson. For Others in Oakland, CensusPlus found over five times the additions that DSE found. When looking at even smaller subgroups, the ratios are even more varied.

Table 2. CensusPlus Adds Relative to DSE Adds (%)

	Oakland	Paterson
Total	6.1 (17.2)	25.0
Black	-47.7 (28.2)	-23.1
Hispanic	65.0 (20.0)	92.5
Asians & Pacific Is.	-32.4 (75.1)	N/A
Owners	43.0 (20.4)	89.5
Renters	-10.1 (25.7)	17.1

3.1.3 Comparison With Alternative Estimates

One could say maybe CensusPlus is telling the true story, implying that DSE is too big. However, CensusPlus estimates in Oakland and Paterson do not appear reasonable when compared with the independent population estimates. DSE estimates definitely are more reasonable (Robinson, 1996). For Paterson, the DSE is 148,394, and the postcensal estimate of 143,663 is within sampling error. The DSE for Oakland at 368,232 is lower than the postcensal estimate of 385,415 and the California state agency's estimate of 406,673. Neither of the independent estimates is within sampling error of the DSE for Oakland.

Table 3. Alternative Estimates and Their Standard Errors

	Paterson	Oakland	NW LA
CensusPlus	135,227 (2,154)	341,176 (5,860)	121,054 (1,695)
DSE	148,394 (2,746)	368,232 (6,938)	N/A
Postcensal	143,663	385,415	123,348
State	N/A	406,673	122,740

3.2 Block Level Results

3.2.1 Comparison With Indices

The comparison of the persons added with the planning database and several indices based on the estimated 1990 inclusion probabilities using Spearman rank correlation and Kendall's tau (Snedecor and Cochran, 1980) for the Oakland site are contained in Table 4. The same statistics for the Paterson data are shown in Table 5. In these tests, the null hypothesis is that the two sets of rankings are uncorrelated or negatively correlated (discordant). The alternative hypothesis is that the two rankings are positively correlated (concordant). No correction for multiple comparisons is included in the calculations of the p-values, but the results are so extreme such correction would not change the inferences.

Table 4. Kendall's tau and Spearman's rank correlation, Oakland, n=160

	tau	p-value	rho	p-value
C^+ vs PDB	0.00	0.46	0.01	0.45
DSE vs PDB	0.12	0.01	0.19	0.01
C^+ vs E(missed)	-0.04	0.55	-0.06	0.78
DSE vs E(missed)	0.18	0.00	0.27	0.00

In Oakland, the statistics for the comparison of the ranks from the DSE with the planning database and two of the indices based on the 1990 inclusion probabilities with the Spearman rank correlation range from 0.19 to 0.27, and the p-values range from 0.01 to 0.00. For Kendall's tau, the statistics for DSE versus the planning database and the 1990 indices range from 0.12 to 0.18, and the p-values range from 0.01 to 0.00. When doing the same comparison with CensusPlus instead of DSE, the values of Spearman's rank correlation statistic range from -0.061 to 0.001, and the p-values range from 0.452 to 0.778. Likewise, the values of Kendall's tau statistic range from -0.04 to 0.01, and the p-values range from 0.47 to 0.55.

The values of the statistics for Paterson are larger than those observed in Oakland and are all positive. The relative values of the statistics follow a pattern in Paterson similar to the one observed in Oakland. The p-values for all the statistics range from 0.04 to 0.00.

Many of the statistics are too small to be important although the amount of data used permits detecting that there is some concordance between the two rankings. The values of the statistics for the comparison of DSE with the planning database and the indices based on the inclusion probabilities in Paterson are large enough to be considered

as indicating concordance. For Spearman's rank correlation, the values range from 0.41 to 0.51 while the values of the Kendall's tau statistic range from 0.31 to 0.35. There also appears to be concordance between the rankings based on CensusPlus and DSE since the range of Spearman's rank correlation and Kendall's tau is from 0.44 to 0.61 when both sites are considered. The planning database and the indices from the inclusion probabilities also appear to be in concordance because the range of the two statistics across both sites is 0.45 to 0.82.

Table 5. Kendall's tau and Spearman's rank correlation, Paterson, n=114

	tau	p-value	rho	p-value
C^+ vs PDB	0.14	0.01	0.21	0.01
DSE vs PDB	0.32	0.00	0.47	0.00
C^+ vs E(missed)	0.11	0.05	0.16	0.04
DSE vs E(missed)	0.35	0.00	0.50	0.00

3.2.2 Mail Nonresponse and Indices

For an indication of the validity of the planning database and the indices based on the estimated inclusion probabilities as indicators of the hard-to-count areas, comparisons with the mail nonresponse data were made. Table 6 shows the ordinary correlation coefficients at the block group level and the block level in Oakland and Paterson.

The comparison of the indices of difficulty to enumerate and the mail nonresponse uses the ordinary correlation coefficient as opposed to its nonparametric counterparts, Spearman's rank correlation and Kendall's tau. For both block clusters and block groups, the mail nonresponse rates tend to be more positively correlated with the planning database and indices based on the inclusion probabilities which are rates. Similarly, the number of mail nonrespondents tends to be more correlated with the indices based on numbers of persons missed or numbers with a low enumeration probabilities.

Table 6. Correlation Between Mail Nonresponse and Indices of Difficulty to Enumerate

	Block Groups		Blocks	
	Oakland	Paterson	Oakland	Paterson
Rate & PDB	0.67	0.64	0.31	0.55
Rate & Ave(1-p_i)	0.64	0.81	0.67	0.65
No. & E(missed)	0.76	0.92	0.79	0.73

For both block clusters and block groups, the highest observed correlation is between the number of mail nonrespondents and the expected number of persons missed with values ranging from 0.73 to 0.92 across both sites. For all the indices based on the inclusion probabilities, the mean probability of missing a person has the highest correlation with the nonresponse rate at the block group level and the block level. The data for Oakland indicate that the planning database and the mean probability of missing a person in blocks and block groups are comparably correlated with the mail nonresponse rate with correlation coefficients of 0.67 and 0.64, respectively. In Paterson, the same correlation coefficients are 0.65 and 0.81, respectively. At the block level, correlation between the mail nonresponse rate and the mean probability of missing a person is 0.67 in Oakland and 0.65 in Paterson, while the correlation with the planning database is 0.31 in Oakland and 0.55 in Paterson. The mail nonresponse rate appears more highly correlated with the mean probability of missing a person than with the planning database. The correlations with the planning database appear higher for block group level than at the block level probably because the planning database is constructed with data measured at the block group and tract level.

4. Summary

As designed and implemented in the 1995 Census Test, DSE increases the estimate after nonresponse followup in the traditionally undercounted groups, primarily Blacks and renters, while CensusPlus does not. However, both CensusPlus and DSE increase the post nonresponse followup estimate for Hispanics. Only DSE increases the estimate for Asians and Pacific Islanders.

The conclusion is supported by the following aspects:

The CensusPlus adjustment factor for the post nonresponse followup estimate is less than 1.0 for all but three of the 14 poststrata for Blacks in Oakland and for all but one of the Black poststrata in Paterson while the DSE adjustment factors for all these poststrata are greater than 1.0.

CensusPlus added only 6.1 percent of the number of people added by DSE after nonresponse followup in Oakland. CensusPlus was better in Paterson but still added only 25.0 percent of the amount DSE added.

The relative pattern of additions of persons to block clusters for the 1995 DSE is concordant with the indices of difficulty to enumerate, namely the planning database and the indices based on the estimated probabilities of being enumerated from 1990 PES data.

As a side benefit, the expected number of persons missed based on the 1990 inclusion probabilities is an excellent predictor of the number of mail nonrespondents in block groups and block clusters. The planning database has some prediction ability for block groups, but not for areas as small as block clusters.

The results shown are dependent on the selection of the test sites and the design and implementation of the CensusPlus and DSE methodologies in the 1995 Census Test. These results may or may not apply in general.

The 1995 design of CensusPlus is considered unacceptable for ICM. There were 13 evaluations of the 1995 ICM (Vacca, Mulry, and Killion, 1996). The major component of error for CensusPlus could be called rostering error. The evaluations showed that not even all the correct census enumerations were included in the resolved rosters for CensusPlus. CensusPlus is being redesigned and tested again in the 1996 Community Census in 7 tracts in Chicago and 3 Indian reservations.

For DSE, the major component of error appeared to be a high nonresponse rate. An evaluation showed that the noninterviewed households were smaller and had a lower match rate than the average of their blocks. The 1995 design of DSE is being refined for the 1996 Community Census.

If CensusPlus looks promising after the results of the 1996 test are analyzed, both CensusPlus and DSE will be tested again in the Dress Rehearsal in 1998. The design of 1996 ICM attempts to avoid the major sources of error observed in the 1995 test. We should have those results by the fall of 1997.

* This article reports results of research undertaken by staff members of the Census Bureau. The views expressed are attributable to the authors and do not necessarily reflect those of the Census Bureau.

References

Alho, J., Mulry, M. H., Wurdeman, K. and Kim, J. (1993) "Estimating Heterogeneity in the Probabilities of Enumeration for Dual System Estimation." Journal of the American Statistical Association, 88, 1130-1136.

Fay, Robert E. (1990) "VPLX: Variance Estimates for Complex Samples, " Proceedings of the Section on Survey Research Methods, American Statistical Association, 266-290.

Mulry, M. H. and Griffiths, R. (1996) "Comparison of CensusPlus and Dual System Estimates." 1995 Census Test Evaluation Memorandum No. 42, Decennial Methods Division, Bureau of the Census.

Mulry, M. H. and Singh, R. P. (1994) "New Applications of Sampling and Estimation in the 1995 Census Test," Proceedings of the Survey Research Method Section, American Statistical Association, 742-747.

Robinson, J. Gregory (1996) "Evaluation of CensusPlus and Dual System Estimates with Independent Benchmarks." 1995 Census Test Evaluation Memorandum No. 43, Decennial Methods Division, Bureau of the Census.

Robinson, J. Gregory and Kobilarcik, Edward (1995) "Identifying Differential Undercounts at Local Geographic Levels: A Targeting Database Approach," paper presented at the Annual Meetings of the Population Association of America, San Francisco.

Snedecor, George W. and Cochran, William G. (1980) Statistical Methods, The Iowa State University Press, Ames, Iowa.

Vacca, E. A., Mulry, M. H., and Killion, R. A. (1996) "The 1995 Census Test: A Compilation of Results and Decisions." 1995 Census Test Evaluation Memorandum No. 46, Decennial Methods Division, Bureau of the Census.