AN INVESTIGATION INTO THE INDEPENDENCE BETWEEN THE CENSUS 2000 DRESS REHEARSAL AND THE INTEGRATED COVERAGE MEASUREMENT/POST ENUMERATION SURVEY[1]

Anne T. Kearney, Katie M. Bench, and Rita J. Petroni
U.S. Bureau of the Census, PRED, Mail Stop 9200, Washington, DC 20233-9200, anne.theresa.kearney@census.gov

1.0 BACKGROUND AND INTRODUCTION

1.1 Background

The U.S. Census Bureau conducted the Census 2000 Dress Rehearsal in 1998 in three sites: Sacramento, CA; Columbia, SC and 11 surrounding counties; and Menominee, WI. As in any census, not all housing units returned a census form. In Sacramento and Menominee, we followed up a sample of nonrespondents, while in South Carolina we did 100 percent nonresponse follow-up. To evaluate the effectiveness of the census to completely cover the population in those sites, the Census Bureau conducted a survey of a sample of the dress rehearsal block clusters. In Sacramento and Menominee we called this survey the Integrated Coverage Measurement (ICM) survey, and in South Carolina we called it the Post Enumeration Survey (PES). For the purposes of this analysis we will refer to them jointly as ICM/PES.

In order for the ICM/PES to provide data needed to measure coverage error in the census, three *a priori* assumptions about the survey must be maintained (Hogan, 2000; Hawala, 1999). The first assumption, the causality assumption (also known as the independence assumption), maintains that inclusion in the ICM/PES is not altered by inclusion in the census, and vice versa. If this assumption fails then there could be bias in the coverage estimates.

To ensure independence between the operations, the Census Bureau starts with an independent list of housing units and uses field staff who are specially trained on the importance of independence.

The second assumption is the homogeneity assumption. This assumption maintains that the marginal probabilities of the ICM/PES and census are uncorrelated. For example, we will maintain this assumption if the capture probabilities are homogeneous in at least one of the lists. (Wolter, 1986). The Census Bureau addresses this assumption by post-stratifying the ICM/PES respondents before calculating dual system estimates so that respondents in the same post-stratum cell have similar inclusion probabilities.

The third assumption is the autonomy assumption. This assumption maintains that inclusion in the ICM/PES or the census is a result of mutually independent trials (Wolter, 1986). It is believed that if this assumption fails and the inclusions are positively correlated, there will be an increase in the random error associated with the estimated undercount due to clustering in the data.

1.2 Introduction

As discussed in section 1.1 we assume that the Census 2000 Dress Rehearsal and the ICM/PES are independent, and, hence, that ICM/PES operations do not contaminate the census and vice versa. While both are important, this paper examines the impact of ICM/PES on census data. Hence, we analyze the difference between census data collected in ICM/PES block clusters and census data collected in non-ICM/PES block clusters during the Census 2000 Dress Rehearsal. If significant differences exist between the two sets of data, contamination of the census by the ICM/PES survey could be the cause.

A respondent or census interviewer becoming aware of those block clusters where the ICM/PES was being conducted could affect the overall coverage of the Census 2000 Dress Rehearsal. For example, becoming aware of the ICM/PES could remind a respondent about the Census 2000 Dress Rehearsal leading him/her to find and return a Be Counted Form. (Be Counted Forms are census forms that the Census Bureau made available in hard to enumerate areas.) Or if a census interviewer knew that the ICM/PES was being conducted in his/her assignment area, he/she may do a better job enumerating people in those areas. Such situations could have caused contamination and lead to differences in the census data

[1] This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a more limited review by the Census Bureau than its official publications. This report is released to inform interested parties and to encourage discussion.

from ICM/PES block clusters and census data from non-ICM/PES block clusters.

To investigate the potential contamination effects of the ICM/PES on the census, Hawala (1999) paired block clusters that were in ICM/PES with block clusters that were not in ICM/PES. He tested the difference in census estimates from these pairs using the Bonferroni multiple comparison procedure.

This paper expands on Hawala's research. Here we use census data in a whole-group analysis. That is, we compare national level estimates for all ICM/PES block clusters to national level estimates for all non-ICM/PES block clusters. The methodology is discussed in Section 2. To test for significant differences we use a multiple comparison procedure called the False Discovery Rate (FDR) offered by Benjamini and Hochberg (1995). We also used the Bonferroni multiple comparison method. Results are presented in Section 3 and conclusions are discussed in Section 4.

2.0 METHODOLOGY

2.1 Paired Block Design

Hawala's (1999) approach to the test for independence was based on an approach used in the 1990 Census (Davis, 1991), except that Hawala paired block clusters that were in ICM/PES with up to four block clusters that were not in the ICM/PES. He matched the block clusters that were in the same site, county, type of enumeration area (TEA) (which indicates the method used to collect the data: mail-out/mail-back or update/leave), and sampling stratum. He matched the block clusters within the above cells on the number of housing units in the block cluster based on a pre-census count. He aggregated the data to the block cluster level, calculated averages and proportions within the block clusters, and then calculated averages across all ICM/PES block clusters and all paired non-ICM/PES block clusters and tested differences in averages.

Hawala chose variables that were related to population coverage, respondent reaction, and housing unit status to detect contamination.

2.1.1 Bonferroni Procedure

Hawala used a Bonferroni procedure to adjust the overall α value to maintain an overall ten percent level of significance in his t-tests. He used $\alpha_R = 1 - (1 - \alpha_N)^{1/C}$, where α_R is the significance level used to maintain an overall significance level of $\alpha_N = 0.1$. C is the number of comparisons. α_R ranged from 0.0009 to 0.002 in the paired block analysis.

2.2 Whole Group Analysis

For whole group analysis, we group all census block clusters selected for the ICM/PES and group all census block clusters not selected for ICM/PES. National level estimates (either proportions or averages) for the two groups are calculated and compared using t-tests. Variances used in forming t-tests are calculated using a stratified jackknife procedure in VPLX. VPLX, developed by Robert E. Fay of the Census Bureau, is a variance estimation software package able to handle complex sampling designs (Fay, 1989).

2.2.1 Multiple Comparison Procedures

We used the FDR multiple comparison procedure. This procedure looks at the proportion of errors committed by falsely rejecting the null hypothesis. The FDR has some advantages for our study over other procedures such as the familywise error rate and Bonferroni procedures. For instance, when more of the hypotheses are not true, the potential for increase in power is larger for the FDR procedure, and the power of the FDR procedure is uniformly larger than that of the other methods (Benjamini and Hochberg, 1995).

The procedure below controls the FDR at $q^* = 0.1$ for independent test statistics and for any configurations of false null hypotheses. Consider testing $H_1, H_2, \ldots, H_m$ based on the corresponding p-values $P_1, P_2, \ldots, P_m$. Let $P_{(1)} \le P_{(2)} \le \ldots \le P_{(m)}$ be the ordered p-values, and $H_{(i)}$ be the null hypothesis corresponding to $P_{(i)}$. Benjamini and Hochberg define the following Bonferroni-type multiple-testing procedure: let k be the largest i for which

$P_{(i)} \le \frac{i}{m} \times q^*$, then reject all $H_{(i)}$ i=1,2, ..., k.

We also used a Bonferroni procedure like the one described in Section 2.1.1.

2.2.2 Variable Selection and Grouping

Demographic characteristics related to coverage, average number of persons per occupied housing unit, and missing and edited data rates are examples of the variables we estimated. We chose these variables because we thought they could be affected by contamination. For example, other relatives include the following relations to the reference person: brother/sister (and in-laws), mother/father (and in-laws), grandchild, grandparent, nephew/niece, uncle/aunt, or cousin. These groups may have a more tangential relationship to the household and may more often be left off household rosters and so may be more subject to contamination.

For South Carolina we calculated estimates for the entire site and also estimates for each TEA. The computations for Sacramento and Menominee were not broken out by TEA since the sites only consisted of one TEA. Sacramento is entirely mail-out/mail-back, and Menominee is entirely update/leave.

The variables described above are split into four groups for determination of significance. The person-level variables dealing with demographics and form types are one group, and the variables which deal with missing data are the second group. The third group is a housing-unit level variable: average number of persons per occupied housing unit. The fourth group is the distribution of number of housing units reported at a basic street address. (In future research we may do a chi-square test on these distributions.) An FDR and Bonferroni procedure are done for each grouping.

3.0 RESULTS

3.1 Paired Block Design

Hawala found few significant differences compared to the number of variables he tested. He concluded that there was not evidence of contamination between the two operations.

3.2 Whole-Group Analysis

Note that the Menominee site is small and for many of the comparisons there were fewer than 100 unweighted persons in the numerator of the proportions.

3.2.1 Form-Type and Demographic Variables

Table 1 shows the differences in proportions between non-ICM/PES block clusters and ICM/PES block clusters for group one, the variables relating to form type and demographic data. As the entries in the Difference column demonstrate, each site has some negative and some positive differences meaning there is not a general trend in the differences.

Most p-values in Table 1 are highly insignificant. There are no significant differences using either the FDR or Bonferroni procedure. Notice that the proportion Male is negative in every table except Menominee. This trend would lead one to suspect that the ICM/PES may have improved coverage of this group. Based on this table, we do not see any strong signs of contamination.

3.2.2 Variables Relating to Nonresponse

Table 2 shows the differences between non-ICM/PES block clusters and ICM/PES block clusters for the variables relating to nonresponse. Sacramento (which is entirely a mail-out/mail-back TEA) and the South Carolina mail-out/mail-back TEAs tend to have negative differences, and the South Carolina update/leave TEA tends to have positive differences. Menominee is entirely update/leave, but the same trend is not present. Note that a positive difference in the average number of data defined persons means non-ICM/PES block clusters have on average more data defined persons, where as a positive difference in the other variables means that non-ICM/PES block clusters have more missing data.

There are three significant differences in Table 2. In Sacramento, the difference between non-ICM and ICM on the proportion of times tenure was edited or imputed is significant by both the FDR method and the Bonferroni method. We do not suspect contamination in Sacramento, but there is weak evidence in the South Carolina update/leave TEA. With the two significant differences out of seven and with the difference in the amount of time relationship is edited or imputed is barely insignificant, and the signs of these three differences are positive, it appears that the ICM may have influenced respondents to supply more complete data. It is possible that early interaction with the ICM/PES listers who were required to record number of units at the basic street address and in rural areas were required to record householder name may have created a sense of awareness in the respondents that resulted in more complete response to the census. One theory on differences in the estimates due to the additional contact with Census Bureau staff (even though it be for the ICM/PES operation) is that respondents would be less cooperative; they could have felt that they helped enough already. But the reverse is found in the data. Another possibility is that the census interviewers became aware of the ICM/PES in their clusters and did a better job collecting data in these clusters.

3.2.3 Average Number of Persons per Occupied Housing Unit

Table 3 shows no significant differences in average number of persons per occupies housing unit in any site. Neither do we see a trend in the differences. The differences in Sacramento and the mailout/mailback TEA in South Carolina have different signs.

3.2.4 Proportion of Housing Units at a Basic Street Address

Table 4 shows the differences in proportion of units at the basic street addresses between non-ICM/PES block clusters and ICM/PES block clusters. Multi-unit structures are uncommon in update/leave (more rural) areas, so the last category was dropped from the South Carolina update/leave table, and Menominee is not presented because three of the six cells (number of units by non-ICM/PES and ICM/PES) had less than 10 observations.

In the South Carolina update/leave TEA, the FDR method detected two significant differences. These differences are not statistically significant by the Bonferroni method. It would appear that multi-unit structures (with 3-10 dwellings) are more often reported in the non-ICM/PES block clusters. Since this significant difference is positive and the single unit at basic street address is negative, it is possible that multi-unit structures in the ICM/PES block clusters are reported as single unit dwellings in the update/leave areas. On the other hand, in rural areas (update/leave areas tend to be rural), this variable is calculated based on the number of units on the file with the same mapspot. (In mailout/mailback areas, it is the number of units on the file with the same house number and street name.) It is not a respondent reported variable. Possibly only one unit at the address returned a form in ICM/PES block clusters so that the address appeared to be a single unit.

Due to time constraints we have not yet conducted a chi-square test on these distributions.

4.0 CONCLUSIONS

Hawala concluded there was no contamination of the census by the ICM/PES. Our analyses show that a weak area of concern for contamination is update/leave TEAs. The difference in the results between Hawala and this paper could be due to the different methodologies. Hawala included a large number of variables in his study, and found only a handful of significant differences. Here we included a smaller number of variables in the investigation, and used a different methodology to detect significant differences.

Most significant differences found were in update/leave TEAs. A more appropriate test on Table 4 would be a chi-square. It is unknown if this test would turn up significant differences. Based on this, we are not convinced that contamination of the census by the ICM/PES is present. Our analyses reveal a small number of differences.

5.0 FUTURE RESEARCH

The Census Bureau is planning another study on contamination for the Census 2000. The survey that evaluates the census in 2000 is called the Accuracy and Coverage Evaluation (A.C.E.). Here we will use Census 2000 data in non-A.C.E. block clusters and A.C.E. block clusters. Analysts will compare differences in similar proportions and averages to detect contamination.

6.0 REFERENCES

Benjamini, Y. and Hochberg, Y. (1995) Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. *J. R. Statist. Soc B*, 57, No. 1, 289-300.

Davis, M.C. (1991) Final Report for Post Enumeration Survey Evaluation Project P14, Part 1: Independence of the Census and the P-Sample: Comparison of Blocks, 1990 Coverage Studies and Evaluation Memorandum Series #P-4.

Fay, R.E. (1989) Theory and Application of Replicate Weighting for Variance Calculations. Proceedings of the Section on Survey Research Methods of the American Statistical Association, pp212-217.

Hawala, S. (1999) Validation of Causality for the Dual System Estimation Procedure in the Census 2000 Dress Rehearsal, Statistical Policy Working Paper 30, Proceedings of the 1999 Federal Committee on Statistical Methodology Research Conference

Hogan, H. (2000) Accuracy and Coverage Evaluation: Theory and Application. Report prepared for the February 2-3, 2000 Dual System Estimation Workshop of the National Academy of Science panel to review the 2000 Census.

Wolter, K. M. (1986) Some Coverage Error Models for Census Data. *Journal of the American Statistical Association*, June 1986, Vol. 81, No. 394, Survey Research Methods

ACKNOWLEDGMENTS

The authors would like to thank Roxanne Feldpausch and Aref Dajani of the Census Bureau for their insightful comments.

Table 1 Differences in Proportions Between Non-ICM and ICM Block Clusters on Person-Level Variables for all Sites

	Difference	*p*-Value
Sacramento		
Persons on BCFEs[1]	0.0000	0.9840
Other Relatives[2]	0.0018	0.6030
Nonrelatives	0.0004	0.9044
Renters	0.0015	0.9522
Males	-0.0057	0.0588
Hispanics	-0.0138	0.2984
Blacks	0.0025	0.7642
Asians	0.0049	0.6456
SC: Total		
Persons on BCFEs[1]*	0.0003	0.5754
Other Relatives[2]	0.0006	0.8808
Nonrelatives	0.0012	0.7338
Renters	-0.0259	0.4966
Males	-0.0009	0.8180
Hispanics	-0.0009	0.7794
Blacks	0.0057	0.8104
SC: Mail-out/Mail-back TEA		
Persons on BCFEs[1]*	0.0004	0.4412
Other Relatives[2]	-0.0003	0.9445
Nonrelatives	0.0013	0.7718
Renters	-0.0360	0.4354
Males	-0.0002	0.9680
Hispanics	-0.0015	0.6966
Blacks	-0.0128	0.6384
SC: Update/Leave TEA		
Persons on BCFEs[1]*	0.0001	0.9680
Other Relatives[2]	0.0030	0.6456
Nonrelatives	0.0014	0.6456
Renters	0.0109	0.5418
Males	-0.0034	0.6242
Hispanics*	0.0014	0.5754
Blacks	0.0656	0.1010
Menominee		
Other Relatives[2]	0.0277	0.0802
Nonrelatives*	-0.0031	0.8494
Renters*	0.0388	0.7114
Males	0.0105	0.3472
Hispanics**	0.0142	0.0750

[1] BCFEs are Be Counted Forms and other form types made available in hard to enumerate areas

[2] Other relatives include the following relations to the reference person: brother/sister (and in-laws), mother/father (and in-laws), grandchild, grandparent, nephew/niece, uncle/aunt, or cousin.

* Unweighted numerator less than 100 in ICM/PES block clusters.

** Unweighted numerator less than 100 in non-ICM/PES and ICM/PES block clusters.

Table 2 Differences in Averages and Proportions Between Non-ICM and ICM Block Clusters on Variables Pertaining to Nonresponse for all Sites

	Difference	*p*-Value
Sacramento		
Avg. Data Defined Persons[1]	-0.0524	0.4066
Housing Units in NRFU[2]	0.0143	0.1802
Edited or Imputed		
Tenure	-0.0085[FB]	0.0006
Sex	-0.0020	0.4778
Hispanic	0.0028	0.4532
Race	-0.0013	0.7490
Relationship	-0.0067	0.0466
SC: Total		
Avg. Data Defined Persons[1]	0.0171	0.7948
Housing Units in NRFU[2]	-0.0082	0.4716
Edited or Imputed		
Tenure	0.0037	0.2040
Sex	0.0017	0.5892
Hispanic	0.0020	0.6384
Race	0.0015	0.6100
Relationship	0.0021	0.4902
SC: Mail-out/Mail-back TEA		
Avg. Data Defined Persons[1]	0.0149	0.8572
Housing Units in NRFU[2]	-0.0067	0.6312
Edited or Imputed		
Tenure	0.0017	0.5824
Sex	-0.0023	0.5156
Hispanic	-0.0013	0.7872
Race	-0.0030	0.3788
Relationship	-0.0017	0.6312
SC: Update/Leave TEA		
Avg. Data Defined Persons[1]	0.0225	0.6170
Housing Units in NRFU[2]	-0.0141	0.4716
Edited or Imputed		
Tenure	0.0102	0.1260
Sex	0.0144[FB]	0.0118
Hispanic	0.0127	0.1336
Race	0.0157[FB]	0.0006
Relationship	0.0140	0.0150
Menominee		
Avg. Data Defined Persons[1]	0.3812	0.0734
Housing Units in NRFU[2]	-0.0385	0.3954
Edited or Imputed		
Tenure**	-0.0065	0.6170
Sex*	0.0103	0.5686
Hispanic	0.0176	0.3844
Race*	0.0062	0.6818
Relationship*	0.0191	0.2150

[1] This is a difference of averages; the remaining are differences in proportions. Two characteristics are required to be data defined, where name counts as a characteristic. Name must have at least three characters in the first and last name together. The characteristics

that are included in the counting are relationship, sex, race, Hispanic origin, and either age or year of birth.
[2] NRFU is short for Nonresponse Follow-up.
[F] Significant based on False Discovery Rate Method
[B] Significant based on Bonferroni Method
* Unweighted numerator less than 100 in ICM/PES block clusters.
** Unweighted numerator less than 100 in non-ICM/PES and ICM/PES block clusters.

Table 3 Differences in Average Number of Persons in Occupied Housing Units Between Non-ICM and ICM Block Clusters for all Sites

Site	Difference	*p*-Value
Sacramento	-0.0560	0.4122
SC: Total	0.0164	0.7794
SC: Mail-out/Mail-back TEA	0.0048	0.9522
SC: Update/Leave TEA	0.0530	0.2302
Menominee	0.3983	0.1118

Table 4 Differences in the Proportion of Housing Units at a Basic Street Address Between Non-ICM and ICM Block Clusters for Sacramento and South Carolina

Units at BSA[1]	Difference	*p*-Value
Sacramento		
1	-0.0267	0.4296
2	-0.0019	0.4592
3-10	0.0117	0.6242
11+	0.0170	0.5620
SC: Total		
1	0.0182	0.6966
2	-0.0129	0.1836
3-10	-0.0100	0.5418
11+	0.0047	09204
SC: Mailout/Mailback TEA		
1	0.0301	0.6030
2	-0.0175	0.1616
3-10	-0.0158	0.4592
11+	0.0032	0.5486
SC: Update/Leave TEA		
1	-0.0117[F]	0.0478
2*	0.0024	0.4902
3-10*	0.0093[F]	0.0602

[1] BSA stands for Basic Street Address
[F] Significant based on False Discovery Rate Method
* Unweighted numerator less than 100 in ICM/PES block clusters.