NON-RESPONSE CONVERSION: THE EXPERIENCE OF THE 1990 POST-ENUMERATION SURVEY

Irwin Anolik and Howard Hogan, Bureau of the Census*

ABSTRACT

The field work for the 1990 Post-Enumeration survey was conducted during the summer of 1990. The goal was to interview an independent sample of households, which were to be matched to the 1990 Decennial Census in order to measure census coverage. After the initial interviewing was complete, survey designers reviewed the level and distribution of the whole household nonresponse cases. A special nonresponse conversion operation was then conducted for all such nonresponse cases in areas with low levels of response. This operation used better trained interviewers. This paper analyzes the results of this operation. It compares the distribution of characteristics from the converted cases with that of households initially interviewed.

1. INTRODUCTION

The 1990 PES was a coverage measurement survey conducted to evaluate the completeness of the 1990 Census of Population and Housing. The PES uses sample survey methods to measure components of census error for a sample of persons, thereby to estimate the net error of the census. Persons residing in PES sample households at the time of the PES interview (i.e., the P sample) were matched against census enumerations to estimate the proportion of the population omitted from the census. A separate overlapping sample, referred to as the E sample, was selected from the census enumerations in order to estimate erroneous inclusions in the census.

The basic design of the 1990 PES involved an attempt to match persons sampled in the PES (P sample) to their Census Day address where they should have been enumerated by the Census. In the PES, one of the main components of error is missing data in the P sample (see Fay et. al., 1988). A major category of this type of missing data involves nonresponse, both for individual persons and for whole households. In an attempt to control the error component associated with whole household nonresponse, a special nonresponse conversion operation (NRCO) was conducted in areas with relatively high levels of such nonresponse. In section 2, we outline some of the field procedures for this operation and the results obtained. Sections 3 and 4 compare demographic characteristics and results of matching associated with the households interviewed during NRCO to those households successfully interviewed during earlier PES field operations. Section 5 looks at the results of coverage estimation in the PES and shows how these estimates would change if NRCO was not implemented. Section 6 summarizes the major findings and presents some conclusions.

2. FIELD METHODOLOGY AND RESULTS

The 1990 PES sampled 167,046 housing units. Field interviewers determined that 144,412 of these units were occupied. Of these occupied housing units, 4,279 or 2.96% were classified as household noninterviews. These noninterviews included 1,269 "refused", 978 "not at home", 283 "other" and 1,749 "last resort" in which sketchy and incomplete information was obtained from a nonhousehold member. These last resort cases were treated as noninterviews. A low household noninterview rate is essential in order to measure small coverage errors accurately. The overall low household noninterview rate achieved by the PES is, in large measure, a result of the new PES design implemented in the test cycle for the 1990 census and related operations. (See Hogan and Wolter 1988). In short, this new design involves a larger interviewing period and a block or cluster sample which allows the interviewer to visit a housing unit several times without extreme travel cost.

Table 1

Results of Interviewing for the 1990 PES Nonresponse Conversion Operation

Outcome	Number	Percent
Converted to complete interview with a household member	1981	(53.1)
Converted to complete interview with a nonhousehold member	619	(16.6)
Remained a household noninterview	540	(14.4)
Household determined to be vacant housing unit	327	(8.8)
Household deleted from PES (i.e. listed in error)	29	(0.8)
Cases not attempted* (transmitted in error)	236	(6.3)
Total	3732	(100.0)

*These cases were previously either interviewed, found to be vacant or listed in error in the initial phase of PES interviewing. Even though these cases were transmitted to the field offices in error, they were not separated from the NRCO universe on the computer files used for analyzing the data for this paper. The computer flags in the data files identifying them as part of NRCO remained.

NRCO was conducted in order to further reduce the household noninterview rate. In the process, we hoped to gather valuable information on the nature of the nonresponse households.

NRCO used only current survey interviewers. These interviewers are highly trained and have experience working on other complex surveys. They attempted to convert as many household noninterviews as possible in District Offices (DOs) where the PES household noninterview rate was 2% or greater. This enabled PES clerical matching operations to continue on schedule by processing those DOs with a household noninterview rate of less than 2%.

As we see in Table 1, NRCO obtained complete interviews for almost 70% of the workload. About one third of these complete interviews were with non-household members. About 1 in 7 (14.4%) of the NRCO households noninterviews. A lesser number of cases were either found to be vacant, listed in error, or transmitted in error. Because some of these households remained inaccessible, the data presented are, to some extent, biased estimates for the nonresponse households. However, the results seem consistent with what we would expect of such households, and may understate the differences between the nonresponse and response households. In other words, the cases that remained as household noninterviews may differ from the response group even more than those NRCO households which were interviewed.

3. CHARACTERISTICS OF THE NRCO HOUSEHOLDS

The information obtained from NRCO interviewing provides some insight into the character of PES nonresponse households and the persons within them. One would expect that the chances of finding a respondent at home at any given time is related to the household size. Table 2 shows the household size distribution for both the response cases[1] and the NRCO cases. Not surprisingly, 37% of the NRCO flagged households were single person households as compared with 24.1% for the response cases. The average size of a household interviewed during NRCO was 2.26, as compared with 2.55 for the response cases.

Table 2

Size of Household for Response Cases and NRCO Cases in the 1990 PES

Size of household	Response Cases (%)	NRCO Cases (%)
All households	100.0	100.0
1 person	24.1	37.0
2 persons	31.4	30.7
3 persons	17.4	14.3
4 persons	15.2	9.6
5 persons	7.1	4.6
6 persons or more	4.8	3.7
Average number of persons per household	2.55	2.26

Table 3 compares the characteristics of persons included in NRCO households to persons in the response households. The table shows that persons in NRCO households were (1) distributed similarly to response households in terms of age and sex, (2) more likely to be minority, particularly black, and (3) less likely to own the home they lived in. These results are not unexpected, since the DOs with the highest rate of nonresponse were more likely to be in large urban areas, such as Los Angeles and New York, which more closely share the characteristics displayed by the NRCO cases.

Table 3

Comparison of Characteristics for Persons Included in NRCO Households to Persons in the Response Households

Person Characteristics	NRCO Households	Response Households
Age Group		
0-9	13.77	16.01
10-19	12.24	14.94
20-29	16.21	15.09
30-44	25.59	23.09
45-64	20.56	18.74
65+	11.63	12.13
Sex		
Male	49.53	48.20
Female	50.47	51.80
Race		
White, Nonhispanic. Other	55.13	66.74
Black	29.71	18.78
Hispanic	11.68	10.74
Asian	2.76	2.01
American Indian	0.72	1.73
Tenure		
Home Owned	56.78	66.12
Home Not Owned	43.22	33.88

In summary, PES households exhibited the same distribution for age and sex, regardless of whether they were in NRCO. Importantly, the results of NRCO suggest that certain households, because of their composition, are more likely to be nonresponse.

4. RESULTS OF MATCHING NRCO HOUSEHOLDS

The demographic characteristics discussed above - namely age, sex, race, and owner/renter status - are some of the variables used in matching PES sample persons to an overlapping sample of census enumerations. The results of this matching are used to estimate the proportion of the population omitted from the census. The weighted number of matches between the PES and census is used to calculate the dual system estimates and thereby the percent undercount for variously defined poststrata. Diffendal (1988) discusses in detail the issues of dual system estimation and postsratification in a PES.

Table 4 illustrates the general results of matching for (1) NRCO households, (2) response households in block clusters containing NRCO households, and (3) all response households. Each person in a household is classified as either matched (enumerated in the census), nonmatched (not enumerated in the census), or unresolved. Match rates for persons in NRCO households are substantially lower than those for persons in response households, regardless of whether or not we include cases with unresolved match status. The match rates for the response cases in NRCO block clusters are only slightly lower than the match rates for all response cases, but still significantly higher than the rates for NRCO cases. This suggests that nonrespondents in the PES are less likely to be enumerated in the census.

Note that a much higher percentage of persons in NRCO households are classified as unresolved. A match probability was imputed for each of these cases . See Schenker (1988) for a full discussion of the methods used to handle missing data in a PES. Most of the unresolved cases represent persons in "last resort"

Table 4

Results of Matching Persons
in Response Households and NRCO Households in the 1990 PES

Type of Household	All Cases			Resolved Cases	
	Match Rate (%)	Nonmatch Rate (%)	Unresolved Rate (%)	Match Rate (%)	Nonmatch Rate (%)
NRCO	69.4	16.9	13.7	80.4	19.6
Response (in NRCO clusters)	87.0	11.1	1.9	88.6	11.4
Response (all)	90.1	8.3	1.6	91.5	8.5
All Households	89.8	8.5	1.7	91.4	8.6

households (see Section 2), who, according to matching rules, could not be assigned a resolved match status. Anecdotal evidence suggests that some of the interview forms containing NRCO data were never processed during the PES matching operation. This likely explains the high percentage of persons in NRCO households with unresolved match status.

The noninterview weighting adjustment used in the PES basically assumes that the distribution of people, characteristics, and match statuses for households not interviewed within a block cluster are the same as for households interviewed (Schenker 1988). The results presented here indicate that match statuses for noninterview households within a block cluster are different than those for the interviewed households in that block cluster. Specifically, persons in NRCO households were less likely to be matched (enumerated in the census). This suggests that the method for handling missing data due to P-sample household noninterviews is a conservative one that leads to an overestimate of the number of P-sample persons who were matched and therefore enumerated in the census.

5. THE EFFECT OF NRCO ON ESTIMATION OF COVERAGE ERROR IN THE CENSUS

This section examines the effects of conducting NRCO on estimates of census coverage error for various subgroups of the population. Table 5 compares the estimates obtained (with NRCO) to what would have been obtained if NRCO flagged cases were allowed to remain as household noninterviews.

In order to estimate census coverage error, dual-system estimates (DSEs) were produced for a number of poststrata (see Hogan 1990). These poststrata are cross tabulated by sex and age.

The DSE can be written as

$$DSE = NP\ (CEN - SUB - EE)/M \qquad (1)$$

where:

N_P = weighted number of people in the P sample

CEN = census count

SUB = number of census whole person imputations

EE = weighted estimate of the number of erroneous enumerations and unmatchable persons in the census

M = weighted estimate of the number of matches between the PES and the census

The estimated percent net undercount can be written as

$$\text{Estimated net undercount (\%)} = 100\ (1 - CEN/DSE) \qquad (2)$$

Table 5

Estimated Undercount Rates (in %) for Various Population Subgroups

Population Subgroup	With NRCO	Without NRCO
Total	2.11	1.97
Race		
White Nonhispanic and Other	1.09	0.96
Black	5.56	5.52
Hispanic	5.66	5.36
Asian	3.33	2.93
American Indian	12.72	12.89
Tenure		
Owner	0.63	0.44
Renter	5.90	5.88
No Classification	1.88	1.73
Census Division		
New England	0.72	0.67
Middle Atlantic	1.26	1.01
South Atlantic	3.52	3.46
East South Central	1.26	1.26
West South Central	2.14	1.86
East North Central	0.94	1.07
West North Central	1.34	1.16
Mountain	2.56	2.42
Pacific	3.43	3.03

In order to calculate what the estimated net undercount would have been without implementing NRCO, a noninterview weighting adjustment was applied to each poststratum (instead of each block cluster) which assumes that match statuses for households not interviewed within a poststratum are the same as for households interviewed. This reduced the amount of computer processing involved and preserved the ability to present alternate estimates by population subgroup.

The "without NRCO" undercount estimates are generally lower for the population subgroups shown in table 5. Two exceptions are American Indians[2] and the East North Central census division. The overall estimated undercount rate is 0.14% lower without NRCO. This result is not surprising since, without NRCO, noninterviews would be weighted using the higher match rates found for the response cases. These higher match rates would, in turn, yield a lower estimate of census undercount.

6. CONCLUSION

This paper discusses the methodology and results of a special nonresponse conversion operation (NRCO) conducted in areas with relatively high levels of nonresponse in the 1990 PES. It compares the distribution of characteristics from the converted cases with that of households initially interviewed.

The results from NRCO highlight differences between the NRCO cases and the response cases. A higher percentage of single person households and, on average, smaller household size among the NRCO cases are in agreement with previous research on the nature of nonresponse households (see Palmer 1967).

Match rates for persons in NRCO households are substantially lower than those for response households. This suggests that nonrespondents in the PES are less likely to be enumerated in the census. It also suggests that the method for handling missing data due to P-sample household noninterviews is a conservative one, overestimating the number of P-sample persons who were matched (and therefore enumerated in the census) thus yielding a lower estimate of census undercount. The two sets of undercount estimates (with and without NRCO) presented herein illustrate this point.

Based on the results of this operation, future coverage estimation activities may want to consider including a nonresponse indicator in models for handling missing data.

REFERENCES

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

Fay, R.E., Passel, J.S., Robinson, G., and Cowan, C.D. (1988). The Coverage of Population in the 1980 Census. Evaluation and Research Re port PHC 80-E4. Bureau of the Census, Washington, D.C.

Hogan, H. (1990). The 1990 Post-Enumeration Survey: An Overview. Paper presented at the Population Association of America Annual Meetings, Toronto, Ontario.

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

Palmer, S. (1967). On the Character and Influence of Nonresponse in the Current Population Survey, *Proceedings of the Social Statistics Section*, American Statistical Association, 73-80.

Schenker, N. (1988). Handling Missing Data in Coverage Estimation, with Application to the 1986 Test of Adjustment Related Operations, *Survey Methodology*, 14, 87-97.

Footnotes

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

[1] In this paper, response cases refer to those households interviewed during initial PES field operations and hence not part of NRCO.

[2] The American Indian poststratum contained only 15 P-sample persons, all of whom were matched to the census.