ESTIMATION OF PES FABRICATIONS FROM QUALITY CONTROL DATA

Antoinette Tremblay, U.S. Bureau of the Census
S. Lynne Stokes and Betsy S. Greenberg, The University of Texas at Austin
S. Lynne Stokes, MSIS Dept., CBA 5.202, The University of Texas, Austin, TX 78712

Key Words: 1990 Census, acceptance sampling, rectification

1. INTRODUCTION

The 1990 Post Enumeration Survey (PES) served as the main instrument in the measurement of coverage differentials by area in the 1990 Decennial Census. As such, it would form the basis of the Census adjustment decision, if the accuracy of the PES was deemed sufficient. This project focus is on P-Sample interviewer fabrication, which is a potential source of error in the estimates of Census undercount from the PES. The Quality Control (QC) operation of the PES interviewing phase was designed to detect fabricated data and correct it. This paper describes the use of QC records to produce estimates of the number of fabricated persons which remain after the QC operation was concluded.

This project was initially designed to evaluate household fabrications. However, from the quality control data, it was not always clear whether or not an interviewer had actually fabricated in cases where erroneously included or deleted household members were detected. For example, there were situations where one or two household members were correctly listed on the roster and the remaining members were either wrong or missing. We speculate that the inerviewer may have found some names on a mailbox, but we can't be sure. Similarly, for households incorrectly listed as vacant, the interviewer may have deliberately skipped the interview, or may have been unable to contact anyone in the household. Regardless of the intention of the interviewer, the effect on the roster was the same. Therefore, the project evolved into more than just an evaluation of whole-household fabrication; it has become an evaluation of all erroneous inclusions and omissions of persons from the PES. Thus, we use the term "collection error" rather than fabrication.

The creation of fictitious individuals and omission of 'real' persons in the PES impacts the estimate of coverage error; it can introduce a bias and increase the variance of the undercount estimate. For example, a fabricated person in the PES would not match to the census and would therefore falsely inflate the number of nonmatches and in turn, the undercount estimate. Fictitious persons included and real persons omitted can bias the P-Sample estimate of the population upward or downward. Also, this bias can differ across various subpopulations.

To see why this is true, the form of the dual system estimator (DSE) is reviewed. The data which would be obtained from the Census and PES if there were no collection errors in the PES is frequently displayed as follows:

		PES: In	PES: Out	Total
Census	In	N_1	N_2	N_c
	Out	N_3	N_4	
	Total	N_p		N

$N_c = N_1 + N_2$ is the population count observed from the Census; $N_p = N_1 + N_3$ is the P-Sample population count made from the P-Sample of the PES, and N_1 is the number of matches (i.e., people counted in the Census) in the P-Sample population. N_4 and N cannot be observed, but the other cells can. (Actually, all the other cells are estimates based on the sample design of the PES, but this will be ignored for purposes of simplifying the discussion.) The DSE is an estimate of the true population count and is calculated as

$$\hat{N} = \frac{N_p N_c}{N_1}$$

This estimator is approximately unbiased if there are no errors due to erroneously included or omitted persons in the PES (and if there are no other types of nonsampling errors). Collection errors in the PES can introduce errors into two of the three components of $\hat{N}$; only N_c cannot be affected by them. In order to see how this happens, M and I are defined to be the number of erroneously omitted and included persons, respectively, in the P-Sample. Further, it is assumed that a proportion μ of the erroneously omitted persons were counted in the Census, so they would have matched had they been counted in the PES. If there were no other types of non-sampling errors except collection errors, the table above would be altered by these errors to look like the one below:

		PES: In	PES: Out	Total
Census	In	$N_1-\mu M$	$N_2+\mu M$	N_c
	Out	$N_3+I-(1-\mu)M$	N_4	
	Total	$N_p+(I-M)$		N

Thus, the DSE would be calculated as:

$$\hat{N}_b = \frac{[N_p+(I-M)]N_c}{N_1-\mu M}$$

The DSE can be biased by the presence of both erroneously omitted and included persons. The numerator can be biased either upwards or downwards. The denominator can only be biased downward. The only instance in which no bias will be introduced into $\hat{N}$ is if $\mu = \mu_p(1- I/M)$, where μ_p is the match rate for all persons counted in the P-Sample (i.e., $\mu_p = N_1/N_p$). Otherwise, $\hat{N}$ will be biased upward if $\mu > \mu_p(1-I/M)$ and downward if $\mu < \mu_p(1-I/M)$. The relative size of the bias would be (in the absence of other nonsampling errors):

$$\text{Relbias}(\hat{N}_b) = \frac{E(\hat{N}_b) - E(\hat{N})}{E(\hat{N})}$$

$$= \frac{I - [(1-(\mu/\mu_p))M]}{N_p(\mu/\mu_p)M} \qquad (1)$$

A further consequence of the presence of collection errors is an increase in the variance of the DSE. Erroneously omitted persons decrease the sample size, and thus the precision of $\hat{N}$. This is complicated by the fact that this decrease in precision may not be apparent from the sample. For example, if I=M, the apparent size of N remains as planned, while the real size is smaller, so that the estimated sampling variance of $\hat{N}$ will be biased downward. In order to evaluate the effect of collection errors, information about the magnitude of I, M, and μ must be estimated.

2 THE QUALITY CONTROL OPERATION

The data for this study came from the Quality Control (QC) operation of the PES interviewing phase. The PES interviewing phase of households and group quarters units in special places involved approximately 158,000 households. This interviewing phase also included a QC reinterview operation. The operation was designed to detect inaccurate PES data and provide the means to correct that data before the results were processed. Overall, approximately 35 percent of the PES sample (i.e., 56,000) was reinterviewed in the QC operation through telephone calls and personal visits. The purpose of the verification check was to confirm that the PES interviewer visited the correct housing unit and conducted the interview according to the survey procedures.

A sample of questionnaires to undergo verification was selected from each work unit. A work unit is a group of completed questionnaires within a cluster for a single interviewer. These groups of questionnaires, by PES cluster, were for one or more days of work, depending on how often the interviewer turned work in. The number of questionnaires in the initial sample for a given work unit depended upon the size of that work unit. The overall sampling rate was 1 in 6, but the rate increased for smaller work units. In addition to sampled cases, questionnaires failing the office edit were included. The verification check involved contacting a household and asking some of the same questions that the PES interviewer asked during the original interview. If any questionnaires failed the verification, the remainder of that work unit was to have been verified and corrected; this was termed rectification.

A P-Sample questionnaire failed the QC check whenever the household roster was incorrect. This could happen if the pre-QC interviewer missed or listed additional people in a household where other members were correct, listed the household as vacant when the QC check showed that the household was not vacant, or when the pre-QC interviewer listed a household that was replaced during QC by either a vacant household or an entirely different household.

Throughout the PES QC process, the Interviewing Automated System (IAS) played a role. Data available from the IAS included indicators if the questionnaire was in the initial QC sample, rectified, or a replacement for a noninterviewed sample case. The IAS information included whether a case was subject to an office edit, whether the household roster was determined during QC to be in error (i.e., at least one person was erroneous), and whether the household was replaced by a completely different household during QC. The IAS also provided the number of persons erroneously omitted from the household detected during QC, and the number of persons erroneously included in the household detected during QC.

3. ANALYSIS

In this section, a procedure is described for estimating the number of undetected errors (such as whole-household fabrications or number of erroneously included persons) which remain in the PES data after QC. In addition, a method for weighting these estimates up to the Evaluation Poststrata and National levels is provided. Then, after accounting for those erroneous persons detected during the PES Follow-up operation, the impact on the DSE of these collection errors is provided. Also, demographic comparisons are made between the fabricated households/persons of whole-household fabrications detected during QC and their replacements.

3.1 Residual Errors

Suppose that the goal is to estimate the number of undetected errors (such as whole-household fabrications) which remain in the PES data after QC. Define T to be the number of work units in the PES, the ith of which is of size n_i. Let Y_i denote the number of errors in the ith work unit and mi the size of the initial random sample which is selected for reinterview from that work unit. Y_i will be observed

if the ith work unit is rectified, which will happen only if a roster error is found among the initial reinterview sample of m_i households. In that case, those errors will be removed from the PES data; designate the sum of all such errors Y_{obs}. Therefore, the only errors which remain undetected are those in work units which were not rectified. Therefore, the quantity of interest is:

$$Y - Y_{obs} = \sum_{i=1}^{T} Y_i - Y_{obs}$$

Since Y_{obs} is known, the only concern is the estimation of Y, the total of errors in all work units. For this parameter, the following estimator is suggested:

$$\hat{Y} = \sum_{i=1}^{T} \chi_i \frac{Y_i}{P_i},$$

where

$$P_i = 1 - \binom{n_i - E_i}{m_i} \Big/ \binom{n_i}{m_i}$$

if $n_i - m_i \leq E_i$, and equal to 1 otherwise. E_i is the number of housing units having collection errors in the ith work unit, and χ_i is an indicator of whether (χ_i=1) or not (χ_i=0) the ith work unit was rectified; that is, χ_i is an indicator of whether or not at least one collection error was detected. P_i is the probability that at least one housing unit having a roster error will be chosen from work unit i in the initial simple random sample from the work unit. (Note that $\hat{Y}$ requires that Y_i and E_i be known only for the work units that are rectified.) If the work units are considered to be the sampling units, is much like a Horvitz-Thompson estimator. The difference between and a true Horvitz-Thompson estimator is that in this case the 'sample size', that is, the number of work units rectified, is random rather than fixed. Then,

$$\hat{Y} - Y_{obs} = \sum_{i=1}^{T} \frac{\chi_i Y_i (1-P_i)}{P_i} \qquad (2)$$

with variance estimate

$$\hat{V} = \sum_{i=1}^{T} \frac{\chi_i Y_i^2 (1-P_i)}{P_i^2}.$$

$\hat{Y}$ and $\hat{V}$ are evaluated in Greenberg and Stokes (1992) and shown to work well.

The following is suggested as an estimator of the number of undetected errors remaining if the PES and QC operation were conducted on the entire P-Sample populations:

$$\hat{Y}^* - \hat{Y}^*_{obs} = \sum_{h=1}^{13} W_h (\hat{Y}_h - Y_{h,obs})$$

where W_h is the weight associated with the hth evaluation sampling group, and where and $Y_{h,obs}$ are as defined above, but for the hth sampling group. (The sampling groups arise from collapsing the PES sampling strata into 13 groups, with these groups approximating as closely as possible the PES Evaluation Poststrata.)

There are numerous instances (from the IAS data) in which the QC procedures appears not to have been followed precisely or consistently. For example, in some cases a roster error was found among the initial sample, but none, or only some, of the remaining households in the work unit were rectified. There was also some nonresponse among households chosen for reinterview. Both situations can be thought of as ones in which Y_i cannot be observed accurately, even when 'sampled'; i.e., when a roster error was found in the initial sample. In these cases, a type of imputation procedure was used. Y_i is replaced in estimator (2) by

$$\hat{Y}_i = \frac{n_i}{m_i + r_i} \sum_{j=1}^{m_i+r_i} Y_{ij}$$

where m_i and r_i are the number of households in the ith work unit which were actually initially sampled and rectified, respectively, and Y_{ij} is the number of errors in the jth household of the ith work unit. The effect of this procedure is to impute the number of errors in the unobserved portion of the work unit to be the same as that in the observed portion. The actual Y_i's, and not the imputed ones, are used in calculating Y_{obs}, however, since the unobserved errors were not removed from the PES data.

A further procedural complication to the theory of estimation is the office edit procedure. If the questionnaire failed the office edit, the questionnaire was reinterviewed. For estimation purposes, these housing units are removed from the sample. The rationale is that any household whose questionnaire is subjected to an office edit is not at risk of having an undetected roster error, since it is reinterviewed (theoretically) with probability one, and is therefore not a part of the population of interest.

Table 1 summarizes our results. The estimates are for total household errors at the PES level and weighted to the national level.

Table 1:
Undetected Household Errors after PES QC

	Omitted Persons	Included Persons	Household Fabrication	Nonvacant Households
PES Level	852 ± 49	336 ± 30	322 ± 29	139 ± 17
National Level	477,636	183,160	179,283	62,608

Table 2 summarizes the results for the 13 evaluations postrata weighted to the national level. The table gives both the number of erroneous persons and the results as a percentage of the P-Sample population.

Table 2:
Erroneous Persons by Poststratum

PES EVALUATION POSTSTRATA	Omitted (M)	Included (I)	Change (M-I)
01: NE, Central, Minority	34,473 (.65%)	11,857 (.22%)	22,616 (.43%)
02: NE, Central, Nonminority	48,533 (.51%)	344,444 (.36%)	14,089 (.15%)
03: US, Noncentral, Minority	101,477 (.42%)	18,534 (.08%)	82,913 (.34%)
04: NE, Noncentral, Nonminority	33,973 (.11%)	10,987 (.03%)	22,986 (.08%)
05: S, Central, Nonminority	61,090 (.65%)	37,080 (.39%)	24,010 (.26%)
06: S, Central Minority	30,273 (.21%)	24,721 (.17%)	5,552 (.04%)
07: S, Noncentral, Nonminority	137,830 (.30%)	71,539 (.15%)	66,291 (.15%)
08: MW, Central, Minority	37,374 (.91%)	23,944 (.58%)	13,430 (.33%)
09: MW, Central, Minority	25,818 (.25%)	1,001 (.01%)	24,817 (.24%)
10: MW, Noncentral, Nonminority	61.404 (.15%)	33,178 (.08%)	28,226 (.07%)
11: W, Central, Minority	42,483 (.69%)	18,704 (.30%)	23,779 (%.39)
12: W, Central, Nonminority	26,238 (.24%)	10,055 (.09%)	16,183 (%.15)
13: W, Noncentral, Nonminor.+Ind	123,974 (.44%)	14,006 (.05%)	109,968 (%.39)
Total	764,910 (.32%)	310,050 (.13%)	454,860 (%.19)

Results show that the number of erroneously omitted persons remaining after PES QC is about 0.32% of the P-Sample population, while the number of fictitious individuals is about 0.13% of the population. The Evaluation Poststrata most affected by these errors are central city and minority ones. The problem of undetected collection error is minimal in most nonminority Poststrata.

Equation (1) shows the relative bias of the DSE due to collection error. The relative bias is seen to be a function of both erroneously omitted and included persons, M and I. However, the estimates of the numbers of undetected erroneously omitted and included persons shown in Table 2 include those from collection errors that were later detected and corrected during PES Follow-up. This reduction in M and I should be accounted for before calculating relative bias in the DSE. PES Follow-up included only nonvacant total household nonmatches. One hundred thirty-eight fabrications were found among this group of households and were corrected during Follow-up.. It is assumed that these 138 households are among those this project classifies as undetected whole-household non-vacant fabrications. From Table 1, the estimate of the number of such households in the PES sample undetected after PES QC was 139! Therefore, it appears that the PES Follow-up was remarkably successful in eliminating this type of collection error from the PES data. However, a minority of the whole household fabrications (139/322 = 43% from Table 1), and an even smaller fraction of all households with erroneously omitted persons (139/852 = 16%) are of this type. It is assumed in this project that all of the whole-household nonvacant fabrications were successfully removed during PES Follow-up, but that none of the remaining collection errors were removed.

Equation (1) shows that the relative bias of the DSE is also a function of the ratio of the match rates of persons erroneously omitted from the PES and of all persons in the PES (μ/μ_p). The IAS system did not provide adequate information to enable the tracing of the match status of the detected erroneously omitted persons in nonfabricated households. However, the match status of such persons in both types of fabricated households was able to be determined. Estimates of these match rates, as well as the match rates for all PES persons, are shown, by Poststratum, in Table 3*. The PES match rate is higher, and usually substantially so, than that of the erroneously omitted persons in fabricated households. This is expected, since it is believed that fabrications most often occur in hard-to-enumerate areas,where match rates are usually low.

Since data about the match rate of erroneously omitted persons from nonfabricated households was not available, it is not known who they resemble more closely: the persons missed in fabricated households or the average person in the PES. It seems reasonable to suppose that the match rate lies somewhere between these two extremes.

In the evaluation of relative bias in the DSE, the two extreme values for the ratio (μ/μ_p) are used. For

* The data from several Postrata were pooled: Poststratum 2 is a combination of poststrata 2 and 6; 10 is a combination of 10 and 4; 11 is a combination of 11 and 8; 9 and 12 are combined.

one estimate of relative bias (method 1), it is assumed that the erroneously omitted persons that remain in the PES data after Follow-up are like those from fabricated households; thus, the ratio of the match rates reported in Table 4 are used. For a second estimate of relative bias (method 2), it is assumed that the erroneously omitted persons that remain in the PES data after Follow-up are like those in other PES households; thus, $(\mu/\mu_p)=1$. It is believed that the true value of relative bias lies somewhere between the two values give in Table 4.

Table 3: Match Rates

PES EVALUATION POSTSTRATA	PES MATCH RATE	MATCH RATE 0F REPLACE-MENTS
01: NE, Central, Minority	78.5%	77.5%
02: NE, Central, Nonminority	90.1%	70.9%
03: US, Noncentral, Minority	84.4%	62.3%
04: NE, Noncentral, Nonminority	92.7%	95.5%
05: S, Central, Nonminority	86.6%	73.5%
06: S, Central Minority	92.8%	71.0%
07: S, Noncentral, Nonminority	92.0%	69.0%
08: MW, Central, Minority	83.7%	72.0%
09: MW, Central, Minority	93.0%	91.0%
10: MW, Noncentral, Nonminority	95.1%	96.0%
11: W, Central, Minority	84.8%	74.0%
12: W, Central, Nonminority	91.9%	91.0%
13: W, Noncentral, Nonminor.+Ind	89.6%	71.1%
Total	89.7%	75.1%

Table 4 uses the data of Tables 2 and 3 to produce two estimates of relative bias in the DSE, by Poststratum and nationally. The two National-level estimates are 0.02% and 0.05%. Most Poststrata are minimally affected. Three minority, central city Poststrata (5, 8, and 11) have estimates of relative bias which are 0.10% or larger, using method 2.

Table 4: Relative Bias in the DSE

PES EVALUATION POSTSTRATA	METHOD 1	METHOD 2
01: NE, Central, Minority	.01%	.02%
02: NE, Central, Nonminority	.02%	.03%
03: US, Noncentral, Minority	-.06%	.02%
04: NE, Noncentral, Nonminority	.006%	.005%
05: S, Central, Nonminority	.016%	.21%
06: S, Central Minority	.03%	.06%
07: S, Noncentral, Nonminority	.002%	.04%
08: MW, Central, Minority	.06%	.12%
09: MW, Central, Minority	-.003%	0%
10: MW, Noncentral, Nonminority	.08%	.08%
11: W, Central, Minority	.06%	.10%
12: W, Central, Nonminority	.07%	.07%
13: W, Noncentral, Nonminor.+Ind	-.04%	.008%
Total	.02%	.05%

3.2 Whole-Household Fabrication Detected During Quality Control

Comparisons can be made between the fabricated and replacement persons/households of whole-household fabrications detected during QC. Since, these data are from only those fabrications detected during QC, generalizations beyond them should be made cautiously. Only those persons that are excluded from the PES dual system estimation (not imputed) are included in the tabulations. That is, those persons born since census day, coded as outmovers, and coded as out-of-scope during matching (for the replacements) are included, as are those persons from households that are coded as last resort, whole-household duplicates, and whole-household fictitious.

Six hundred-eleven persons, nationally, were detected as fabrications during the QC operation. Every possible status is represented (nonmover, inmover/mover address, inmover/alternate address, born since census day, no status, and outmover), with the nonmovers claiming over 93%. Every age group is represented, with the two groups of 30-44 and 45-64 each having 24% of the total. When looking at only

those persons with a coded sex, 53% of the fabrications are female. All race/origin categories are represented, with 69% of the fabrications having a coded race/origin being coded 'Black'. Looking at all nonmissing categories of sex, age, race/origin and status, the largest cell frequency is 'nonmover, female, age 45-64, Black' with 9%.

When comparing the characteristics of the replacement persons and the fabricated persons, 1,007 persons replaced the 611 fabricated persons. Every possible status is represented (nonmover, inmover/mover address, inmover/alternate address, born since census day, no status, and outmover), with the nonmovers claiming over 85%. Every age group is represented, with the group of 30-44 accounting for 26% of all persons having a coded age. Fifty-three percent of all replacement persons with a coded sex are female. All race/origin categories are represented, with Black accounting for 49% of all persons having a coded race/origin. Overall, the match rate of the replacement persons is 71.8%. (However, when only those persons that go through imputation are included, the match rate is 75.1%.) Aside from the fact that this match rate is based on only 890 persons, this rate is probably biased downward since by the nature of the QC operation, nonmatched replacement persons are more likely to be detected than are the matched replacement persons.

The average household size of all fabricated households, excluding only outmovers, is 1.76 persons. When only those persons which feed into estimation are included, the average household size is 2.13 persons. The analogous average sizes for the replacement households are 2.44 and 2.67, respectively.

Four hundred-twenty households, nationally, were detected as whole-household fabrications during the QC operation. Fifty-nine percent of the fabricated households are coded as complete interview with a household member. Thirty-two percent are coded as vacant. Only 4% of the households received the outcome code of possible curbstone when put through the PES process. This code signals the possibility of a fictitious household. The criteria needed to receive this code is if the questionnaire is coded complete with household member, no phone number is listed for the household, and no year of birth is coded for any of the household members. Of the combined complete interview with a household member and complete interview with a proxy outcome categories, 58% of the homes which received a tenure code are not owned. In terms of the replacement households, there are only 417 to replace the 420 fabricated households. After investigation, it was discovered that there were no such map spot numbers that identified the other three households. That is, they were never to have been included in the PES sample. Of the 417, 64% are coded as complete interview with a household member. Ten percent are coded as vacant. Less than 1% of the households received the outcome code of possible curbstone when put through the PES process. Of the combined complete interview with a household member and complete interview with a proxy outcome categories, 52% of the homes which received a tenure code are not owned.

4. CONCLUSION

The estimates show that there are more undetected households that have erroneously included persons, when other household members are properly accounted for, than households that are complete fabrications. Furthermore, it is estimated that about 60% of the undetected whole-household fabrications are classified as vacant in the PES. Since only nonvacant households were included in PES Follow-up, less than half of the residual fabricated households have a chance of being caught.

At the PES (National) level, it is estimated that there are 852 (477,636) undetected households with at least one erroneously omitted person. Likewise, there are 336 (183,160) undetected households with at least one erroneously included person. (These figures include the 139 (62,608) whole-household nonvacant fabrications detected during the PES Follow-up operation.)

In every Poststratum the estimated net effect of collection errors is to make the P-Sample population estimate smaller. That is, the number of erroneously included persons is smaller than the number of erroneously omitted persons. This effect is largest in Poststratum 1 (Northeast, Central city Minority) and smallest in Poststrata 4, 6, and 10 (NE and MW Noncentral Nonminority and South, Central, Nonminority). The estimated number of erroneously included persons is largest in Poststratum 8 (Midwest, Central city, Minority).

When all aspects are considered, the effect on the DSE of collection errors is minimal, except possibly in selected Poststrata. When the match rate for all erroneously omitted persons in nonfabricated households is assumed to be the same as that for erroneously omitted persons in whole-household fabrications, the relative bias at the National level is 0.02%. When the match rate is assumed to be the same as that for persons in the PES, the relative bias at the National level is 0.05%.

REFERENCE:

Greenberg, B.S. and Stokes, S.L. 1992.. "Estimating Nonconformance Rates after Zero-Defect Sampling with Rectification," to appear in *Technometrics*.