Survey Methodology, December 1988
Vol. 14, No. 2, pp. 241–263
Statistics Canada

Total Error in the Dual System Estimator: The 1986 Census of Central Los Angeles County

MARY H. MULRY and BRUCE D. SPENCER[1]

ABSTRACT

The U.S. Bureau of the Census uses dual system estimates (DSEs) for measuring census coverage error. The dual system estimate uses data from the original enumeration and a Post Enumeration Survey. In measuring the accuracy of the DSE, it is important to know that the DSE is subject to several components of nonsampling error, as well as sampling error. This paper gives models of the total error and the components of error in the dual system estimates. The models relate observed indicators of data quality, such as a matching error rate, to the first two moments of the components of error. The propagation of error in the DSE is studied and its bias and variance are assessed. The methodology is applied to the 1986 Census of Central Los Angeles County in the Census Bureau's Test of Adjustment Related Operations. The methodology also will be useful to assess error in the DSE for the 1990 census as well as other applications.

KEY WORDS: Nonsampling error; Post enumeration survey; Coverage evaluation, Undercount; Capture-Recapture.

1. INTRODUCTION

The dual system estimator (DSE) is used in several contexts for estimating the size of a population. Its applications range from wildlife populations to human populations. DSEs of births are used at the U.S. Bureau of the Census in the formation of the demographic analysis estimates of the national population. Currently, the Census Bureau intends to use DSEs for measuring coverage error in the 1990 Decennial Census. This paper focuses on the application of the DSE in the census context where the two systems are the original enumeration and a Post Enumeration Survey (PES).

The obvious estimator based on the DSE of census undercoverage is $\widehat{UC}$, given by $\widehat{UC}$ = DSE - CEN, with CEN referring to the size of the original census enumeration. Since DSE = CEN + $\widehat{UC}$, the DSEs also provide alternative estimates of population. A more general class of alternative estimates based on the DSE (Spencer 1980; 1986) is $(1 - f) \times$ CEN + $f \times$ DSE, or equivalently

$$\text{CEN} + f \times \widehat{UC}$$

with $0 \leq f \leq 1$.

Estimates of total error of the DSE are essential for determining what value of f leads to the most accurate estimator of population size. Since the range of values for f include 0 and 1, the selection of either CEN or DSE is possible. The criteria for improvement of one set of population estimates over another may be based on measures of the quality of the distribution of the population (Hogan and Mulry 1987; Spencer 1986). Estimates of total error in the

[1] Mary Mulry, Undercount Research Staff, Statistical Research Division, U.S. Bureau of the Census, Washington, D.C. 20233. Bruce Spencer, Department of Statistics, Northwestern University, Evanston, IL 60201 and NORC. The views expressed are attributed to the authors and do not necessarily reflect those of the Census Bureau.

DSE are also important for statistical planning purposes, *e.g.*, how much money should be spent and how big a sample should be fielded in the PES.

DSEs are subject to several components of nonsampling error, in addition to sampling error. We present models of the total error and the components of error in the DSE. The models relate observed indicators of data quality to the first two moments of the components of the error. We then use techniques of propagation of error to estimate the bias and variance of the DSE. In doing so, we assess the total error, or the joint effect of the errors. Previous work on error models for the DSE includes Seltzer and Adlakha (1974).

The methodology is applied to the 1986 Census of Central Los Angeles County, also known as the 1986 Test of Adjustment Related Operations (TARO) conducted in Los Angeles (Diffendal 1988). The PES in TARO comprised about 6,000 housing units and over 19,000 people. A sensitivity analysis shows how the component errors interact, which ones cancel, and which ones compound each other. The methods described here to estimate the error in the TARO DSE can be extended to estimate the error in the 1990 DSEs.

We have tried to organize this paper to facilitate incomplete reading of the paper. Section 2 introduces and presents the rationale for the TARO DSE and its major components. Our strategy for assessing the component errors and combining them to estimate the total error in the DSE is described next (Section 3). A detailed description of the DSE, with notation, is necessary for precise description of the component errors (Section 4). Following that description is an assessment of the component errors (Section 5). A synthesis of the component errors leads to estimates of the total error of the DSE (Section 6). Our major conclusions are then presented (Section 7).

2. DUAL SYSTEM ESTIMATOR

The application of the dual system estimator requires assuming that there are two lists of the population. The first list is the original census enumeration, and the second is an implicit list of those covered by the sampling frame for the P sample of the PES, whom we will call the P-sample population. The sampling frame itself is not a list of people, but of census blocks.

The P sample is one of the two samples that comprise the PES. The PES is composed of the E sample, which is a sample of census enumerations, and the P sample, which is a sample of the population. The E sample is selected to estimate the number of enumerations that are erroneous. The P sample is selected to estimate, through dual system estimation, the number of people missed by the original enumeration.

Table 1

Probabilities of Inclusion in a Cell

	Original Enumeration		
	In	Out	Total
P sample In	p_{i11}	p_{i12}	p_{i1+}
Out	p_{i21}	p_{i22}	p_{i2+}
Total	p_{i+1}	p_{i+2}	p_{i++}

Table 2
True Population Size in Each Cell

	Original Enumeration		
	In	Out	Total
P sample In	N_{11}	N_{12}	N_{1+}
Out	N_{21}	(N_{22})	(N_{2+})
Total	N_{+1}	(N_{+2})	(N_{++})

The dual system estimator is based on a model that the probabilities that the i-th individual in the population of size N is in the census or not and in the P sample or not are as shown in Table 1 (Wolter 1986a); see Wolter (1986a) for discussion and references to earlier work. The true population size in each category is defined in Table 2.

In Table 2, $N_{++} = N$, the total population size. Even if we could observe the N_{ij}'s in the first row and first column, the N_{ij}'s in parentheses would not be observed directly, but would have to be estimated from the model. The DSE of N then would have the form $N_{1+}N_{+1}/N_{11}$, which we will refer to as the ideal DSE.

In estimating population size for measuring census coverage error, the N's are replaced by estimates from the original enumeration and two sample surveys, the P sample and the E sample. The survey data are weighted by the reciprocals of the selection probabilities. In the following definitions, the estimates with " ^" reflect the possible presence of nonsampling error:

N_p = the weighted number of P-sample selections

$\hat{N}_p$ = the estimate of the total population from the P sample.

CEN = the size of the original enumeration

II_1 = the number of persons imputed

II_2 = the weighted number of census enumerations with insufficient information for matching

EE = the weighted number of erroneous enumerations in the original enumeration, based on the E sample

$\widehat{EE}$ = the estimate of the number of erroneous enumerations in the original enumeration

C = CEN - II_1 - II_2 - EE = the weighted number of distinct people in the original enumeration from the E sample,

$\hat{C}$ = CEN - II_1 - II_2 - $\widehat{EE}$ = the estimate of the number of distinct people in the original enumeration from the E sample,

M = the weighted number of people in the census and the P sample

$\hat{M}$ = the estimate of the number of people in the census and the P sample.

With this notation, $\hat{N}_p$ estimates N_p, which unbiasedly estimates N_{1+}. The *ratio* $\hat{C}/\hat{M}$ is used to estimate the *ratio* N_{+1}/N_{11}. (By themselves, $\hat{C}$ and $\hat{M}$ are not good estimators of N_{+1} and N_{11}.) Thus, the estimator has the form $\hat{N}_{++} = \hat{N}_p \hat{C}/\hat{M}$. The ratio $\hat{C}/\hat{M}$ contains a correction for erroneous enumerations and for cases with insufficient information for matching, II_1 and II_2, so that cases with no chance of being included in the denominator are also excluded from the numerator.

The DSE is used to estimate the percent net undercount, or the *net undercount rate*, in the original enumeration,

$$\hat{U} = 100\,(\text{CEN} - \hat{N}_{++})/\hat{N}_{++}.$$

For the TARO site (*i.e.* Central Los Angeles County) as a whole, CEN = 355,352, $\hat{N}_p$ = 336,707, $\hat{C}$ = 343,567, $\hat{M}$ = 298,204, and $\hat{N}_{++}$ = 388,040. Using these numbers, the estimate of the net undercount rate is 8.42.

3. STRATEGY FOR ASSESSING TOTAL ERROR

The DSE is subject to various sources of error, including error due to incorrect addresses from the P sample, error due to missing data (unit and item nonresponse), response errors, interviewer errors, correlation bias, sampling error, *etc.* We wish to estimate the effects of these diverse sources of error on the DSE.

The first step in our strategy is to express the DSE as a function of components. We have constructed the components so that, for the most part, the different sources of error act either independently or perfectly dependently on different components. By isolating the effects of the various errors, we are better able to identify the major distinct sources of error.

Next, we estimate the first two moments of the component errors, one component at a time. In doing so we draw upon the results of various TARO evaluations and quality control programs. The way we constructed the components implies that correlation between component errors typically equals either 0 or 1.

To study the propagation of errors we have used computer simulation methods. A multivariate distribution of the error components, say F, was assumed. The specification of F was consistent with the first two moments as estimated in Section 5. Realizations of the component errors were simulated by pseudo-random draws from F and then the DSE was calculated; this procedure was repeated 10,000 times and the resulting empirical distribution of the DSE was used as an estimate of its actual distribution. The first two moments of the latter distribution provide numerical estimates of the total error of the DSE.

Sensitivity analysis was performed to discover the importance of using one distributional form for F rather than another. The results suggest that the exact distributional form (beyond the first two moments) is relatively unimportant (see Section 6).

We adopted a Bayesian approach in investigating of the error in the DSE. We estimated the first two moments of the distributions for the error components, then we derived the posterior distribution of the undercount rate conditional on the observed values of $\hat{C}$, $\hat{N}_p$, $\hat{M}$, *etc.*

4. COMPONENTS OF THE DSE

The DSE is subject to sampling errors and nonsampling errors, including failure of assumptions underlying the DSE model. The DSE does have a bias, but the bias in the census context is negligible (Wolter 1986a). Nonsampling errors may affect the accuracy of estimation of N_{+1}, N_{1+}, and N_{11}. Descriptions of the nonsampling error follow.

The error in the estimation of N_{+1} is defined by $\hat{C} - N_{+1} = (\hat{C} - C) + (C - N_{+1})$. The first term $(\hat{C} - C)$ is the net nonsampling error, which contributes to both bias and variance, and the second term $(C - N_{+1})$ is the sampling error, which contributes only to the variance. Define the net nonsampling error as $c = \hat{C} - C$.

The net error c arises during the processing of the E sample when respondents are misclassified as to whether they are correctly or erroneously enumerated in the original enumeration. Therefore, c has three components: c_e, which occurs during the data collection and processing; c_b, caused by a PES design that fails to balance estimates of the gross overcount and gross undercount; and c_i, caused by missing data, $c = c_e + c_b + c_i$. Sections 5.5, 5.6 and 5.7 cover c_e, c_b and c_i, respectively.

The error in the estimation of N_{1+} is defined by $\hat{N}_p - N_{1+} = (\hat{N}_p - N_p) + (N_p - N_{1+})$. The first term $(\hat{N}_p - N_p)$ is the nonsampling error, which contributes to both bias and variance and the second term $(N_p - N_{11})$ is the sampling error, which contributes only to the variance. The net nonsampling error is defined by $n_p = \hat{N}_p - N_p$.

The net error n_p arises during the interviewing for the P sample when the P-sample selections are not interviewed. This situation occurs when household members are fabricated or when there is missing data. Therefore, n_p has two components: n_{pf}, the error due to fabrication and n_{pi}, the error due to missing data, $n_p = n_{pf} + n_{pi}$. Section 5.3 discusses n_{pf}, and Section 5.7 covers n_{pi}.

The error in the estimation of N_{11} is defined by $\hat{M} - N_{11} = (\hat{M} - M) + (M - N_{11})$. The first term $(\hat{M} - M)$ is the net nonsampling error, which contributes to both bias and variance, and the second term $(M - N_{11})$ is the sampling error, which contributes only to the variance.

To facilitate the description of the nonsampling error in the estimation of N_{11}, consider the following tables of P-sample selections and respondents. Entries in Table 3 are the weighted number of P-sample selections in each category. Entries in Table 4 are the weighted number of P-sample responses in each category. Entries in Table 5 are estimates of the number of people in each category based on the P-sample interviewing, responses, and matching operation.

Table 3

P-sample Selections

P-sample Selections	Census Enumeration Status	
	Enumerated	Not Enumerated
Not reported	D_{11}	D_{12}
Reported		
Correct Census Day Address	D_{21}	D_{22}
Wrong Census Day Address	D_{31}	D_{32}

Table 4

Enumeration Status of P-sample Respondents

P-sample Status	Census Enumeration Status	
	Enumerated	Not Eumerated
Fabricated	A_{11}	A_{12}
Not Fabricated		
Correct Census Day Address	A_{21}	A_{22}
Wrong Census Day Address	A_{31}	A_{32}

Table 5

Match Status of P-sample Respondents

P-sample Status	Match Status	
	Matched	Not Matched
Fabricated	B_{11}	B_{12}
Not Fabricated		
Correct Census Day Address	B_{21}	B_{22}
Wrong Census Day Address	B_{31}	B_{32}

Since the P-sample selections who appear as reported in Table 3 are the respondents who are not fabricated in Table 4, $D_{21} = A_{21}$ and $D_{31} = A_{31}$. Also, $A_{11} = 0$ since a case fabricated during the PES cannot be enumerated in the census. Therefore,

$$M = D_{11} + D_{21} + D_{31} = D_{11} + A_{21} + A_{31}.$$

Since a case fabricated during the PES would not have a corresponding census enumeration, we assume $B_{11} = 0$. Therefore, $\hat{M} = B_{11} + B_{21} + B_{31} = B_{21} + B_{31}$.

Then the nonsampling error in the estimation of N_{11}, called m, may be defined as follows:

$$\begin{aligned} m &= \hat{M} - M \\ &= (B_{11} + B_{21} + B_{31}) - (D_{11} + D_{21} + D_{31}) \\ &= -D_{11} + (B_{21} - A_{21}) + (B_{31} - A_{31}). \end{aligned}$$

The error m has three components: $(B_{21} - A_{21})$, which is the error introduced in the matching operation (Section 5.2); $(B_{31} - A_{31})$, which is the error introduced by respondents giving the wrong Census Day address (Section 5.3); and $-D_{11}$. D_{11} has two components: missing match status m_i and fabrication m_f. Section 5.7 covers missing match status, and Section 5.4 covers fabrication.

The ideal DSE can be written as follows:

$$N_{1+}N_{+1}/N_{11} = (\hat{C} - c)(\hat{N}_p - n_p)/(\hat{M} - m).$$

5. COMPONENTS OF PES ERROR

Estimates of the first two moments of the posterior distribution of the undercount rate derive from estimates of the first two moments of the components of PES error. The components are correlation bias, matching error, accuracy of the reported Census Day address, fabrication in the P sample, measurement of erroneous enumerations, balancing the estimates of the gross overcount and the gross undercount, missing data, and sampling error. We next describe the source of each component of PES error and give models for each component. We model the component errors in terms of observable indicators of data quality. We estimate the first two moments of the distributions of the errors for use in the total error model in Section 6.

5.1 Correlation Bias

5.1.1 Source of Error

An important concern for dual system estimation is that the estimate of the proportion of the population enumerated in the census, based on the P sample, is accurate. The violation of one of the independence assumptions underlying dual system estimation may cause the estimate of the proportion of the population enumerated in the census, and thereby the estimate of the population, to be biased.

Three independence assumptions are made for dual system estimator:

Causality. The event of being included in the census is independent of the event of being included in the PES. That is, the cross-product ratio satisfies

$$\theta_i = p_{i11}\, p_{i22} / p_{i12}\, p_{i21} = 1, \text{ for } i = 1, \ \ldots, N.$$

Homogeneity. The capture probabilities satisfy $p_{i1+} = p_{1+}$ or $p_{i+1} = p_{+1}$ for $i = 1, \ \ldots, N$, within each of the post-strata.

Autonomy. The census and the PES are created as a result of N mutually independent trials.

The homogeneity assumption follows combination model M_{th} in Wolter (1986a). All the development for the Peterson model M_t in Wolter (1986a) also applies to model M_{th} when enough information is available to form post-strata where M_t holds.

To control heterogenity in the population the Census Bureau post-stratifies the data based on demographic and geographic variables, a technique originally recommended by Sekar and Deming (1949). An estimate of the population in each post-stratum is calculated and then all the estimates are summed to give an estimate of the total population. Unless the failure of the homogeneity assumption is severe, the estimate lies between the census and the truth.

Research by Wolter (1986b) and Cowan and Malec (1986) has demonstrated that the failure of the autonomy assumption has a negligible effect on the bias of the DSE but causes an increase in its variance. Wolter's formulation allows household members to act individually (autonomy) or together (failure of autonomy). Cowan and Malec present a model that permits clustering of the census misses (failure of autonomy). Next, we model the combined effect of the sources of correlation bias on the DSE.

5.1.2 Definition

For insight into the effect of correlation bias, assume all $\theta_i = \theta$ and write the true population size as

$$N = N_{11} + N_{12} + N_{21} + \theta\ (N_{12}N_{21}/N_{11}),$$

where θ_i is the cross-product ratio defined in Section 5.1.1.

The correlation bias affects only the last term because the other three may be estimated directly. The parameter θ represents the effect of the failure of the independence assumptions. When the independence assumptions hold, $\theta = 1$.

The correlation bias, arising when θ does not equal 1, is the only contributor to t, the error due to failure of the model. The population size can be written as follows:

$$\begin{aligned} N &= N_{1+}N_{+1}/N_{11} + t \\ &= N_{1+}N_{+1}/N_{11} + (\theta - 1)(N_{12}N_{12}/N_{11}). \end{aligned}$$

Therefore, the correlation bias, $t = (\theta - 1)(N_{12}N_{21}/N_{11})$.

5.1.3 Measurement

The parameter θ may be estimated at the national level for racial and ethnic subgroups using demographic analysis estimates of the population size. Note, however, that this technique presumes that the demographic analysis estimates are accurate. Even so, this formulation also permits varying θ to assess the sensitivity of the DSE to the estimate of the effect of the violation of the independence assumptions.

5.1.4 Estimation

Estimates for θ were not made for the 1986 TARO because an alternate source for population estimates did not exist, *e.g.*, no demographic analysis estimates were feasible. However, Ericksen and Kadane (1985) made three estimates of θ for blacks for the 1980 census: 2.1, 2.7, and 3.7. Since the population in the 1986 TARO was predominantly minority (73 percent Hispanic, 12 percent Asian, and 15 percent non-Asian and non-Hispanic), the Ericksen and Kadane estimates for 1980 will be used in this paper: $E(\theta) = 2.1, 2.7$, or 3.7, $Var(\theta) = 0$. We are treating θ as fixed, but unknown. A sensitivity analysis is conducted in Section 6 to demonstrate the effect of alternative values of θ.

These estimates of θ are consistent with the reports of the participant observers in the Los Angeles test site (Childers *et al.* 1987). Our professional judgment is that correlation bias is higher for urban areas than for the country as a whole. This implies that these estimates may be conservative for the Los Angeles test site because it was urban.

5.1.5 Summary

In the total error model the first two moments of the posterior distribution of the correction factor for correlation bias are assumed to be $E(\theta) = 2.1, 2.7$, or 3.7, and $Var(\theta) = 0$.

5.2 Matching Error

5.2.1 Source of Error

Matching error in this discussion refers to errors that occur in the operation where the P sample is matched to the original enumeration. Therefore, matching error does not encompass response errors that arise in the data collection. Although other types of errors may result in an inaccurate assignment of a P-sample respondent's census enumeration status, these sources are treated in other components of error.

After the P-sample interviewing is completed, a search of the census is conducted to determine if the respondents are enumerated. Then the P-sample respondents are designated as matching an enumeration in the census or as not enumerated in the census. Errors in assigning the enumeration status to P-sample persons which occur during the processing of the data are known as matching error. Errors may occur in either direction. People may be designated as matching a census enumeration although they are not in the census, called a "false match," or people may be designated as not enumerated although they are, called a "false nonmatch." Matching error will cause a bias in the estimate of the number of people in both the census and the P-sample population and thereby introduce a bias into the estimates of the number of people missed by the census.

5.2.2 Definition

The denominator N_{11} of the dual system estimator is estimated from sample survey data, the P sample. The following were introduced in Section 4:

A_{21} = the weighted number of people who were enumerated,

B_{21} = the estimate of the number of people who match.

Then the net error due to incorrect classification of enumeration statuses, m_m, may be defined as $m_m = B_{21} - A_{21}$. The conditional expected value and variance of m_m given observed value $\hat{M}$ are denoted by $\mathrm{E}(m_m)$ and $\mathrm{Var}(m_m)$.

5.2.3 Measurement

Measurement of m_m is possible by processing a sample of the cases a second time *i.e.*, by having highly trained personnel rematch them. The assumption underlying an independent rematch of a sample is that the personnel with more training make fewer mistakes in classifying enumeration statuses although they have the same materials and information available as the original workers. The original match codes and the evaluation match codes can be reconciled, and the discrepancies can be resolved.

Two evaluations of the clerical matching were conducted with the 1986 TARO data. One study evaluated the clerical matching for movers, and another evaluated the clerical matching for nonmovers.

In the evaluation of matching for nonmovers (Corby and Mulry 1988), a probability subsample of 35 blocks was chosen for a rematch by professionals from headquarters. The sample was stratified by match rate, and blocks with low match rates were sampled at a disproportionally high rates so that the quality control staff could learn as much as possible about matching errors. Adjacent blocks were not searched so the false nonmatches are possibly underestimated.

The second evaluation study considered matching error for movers (Childers *et al.* 1987). There were 90 movers who were not matched in TARO, and all of these movers were rematched. Eleven matches were found, two of which had been lost during the computer editing.

5.2.4 Estimation

We now use the results of the evaluation subsamples to estimate the moments of the distribution of m_m from the PES sample. Not conducting an extended search in the evaluation for the nonmovers probably reduced the number of false nonmatches found. Experience with extended searches implies that adding an additional 20 percent of the net error of 70 (Hogan and Wolter 1988) is a conservative way to compensate for the lack of one. The results from the two evaluations yield a net error of 95 in the PES sample. Therefore, the net error rate is $-.0055$. We apply the net error rate to only the P-sample cases with a resolved match status because the error in the imputation for the unresolved cases is covered in the Missing Data Section 5.7. The expected value of m_m becomes $\mathrm{E}(m_m) = -1831$, when the overall sampling weight of 17 is used.

An estimate of the variance of the estimate of net matching error for nonmovers has not been calculated. The sample variance of the number of errors for movers is zero because all the nonmatched movers were rematched. However we do not believe that the true variance is zero. One way to obtain a variance specificiation would be to assume that the errors occurred in the manner of a mixture of Poisson processes, *e.g.*, matching errors for movers followed one Poisson processs and matching errors for nonmovers independently followed another Poisson process. Treating the errors as arising from a simple Poisson process would then lead to a conservative estimate of variance; in this case the variance would be estimated by 17×107. However, the Poisson model may not be conservative if the errors occur in clusters. In an attempt to develop conservative estimates of variance, we have (somewhat arbitrarily) multiplied the variance estimate under the simple Poisson model by the overall sampling weight to obtain

$$\mathrm{Var}(m_m) = (17)^2 \times 107 = 30{,}923.$$

5.2.5 Summary

For the total error model, the first two moments of the posterior distribution of the net matching error for the PES sample are assumed to be $E(m_m) = -1831$ and $Var(m_m) = 30{,}923$.

5.3 Quality of the Reported Census Day Address

5.3.1 Source of Error

Some of the respondents in the P sample have moved between Census Day and their PES interview. The respondents may misreport whether they have moved during the time lapse. If they have moved, they may not report their previous address accurately, or their previous address may not be geocoded correctly by the staff. Any of these types of errors may cause the matching operation to search the census in an area other than where the respondent was enumerated. These errors may lead to assigning a nonmatch status to respondents who actually were enumerated because the matching operation is unable to locate their enumerations. Inappropriate assignment of the status of nonmatch will cause the estimate of the number of people missed by the census to be biased upward.

Circumstances under which inaccurate reporting of the Census Day address by a PES respondent will not cause a false nonmatch do exist. If the Census Day address is inside the search area for the reported address, and the reported address is geocoded correctly, then the matching operation will find the person.

5.3.2 Definition

The denominator N_{11} of the dual system estimator is estimated from sample survey data, the P sample. The following were introduced in Section 4:

A_{31} = the weighted number of people with an inaccurate Census Day address who are enumerated,

B_{31} = the estimate of the number of people with an inaccurate Census Day address who match at another address.

Then the net error due to inaccurate reporting of the Census Day address, m_a, may be defined as $m_a = B_{31} - A_{31}$. The conditional expected value and variance of m_a given the observed value $\hat{M}$ are denoted by $E(m_a)$ and $Var(m_a)$.

5.3.3 Measurement

Measurement of m_a is based on a follow-up of a sample of P-sample respondents whose enumeration status is "not enumerated". Data from the follow-up are used to estimate the error that arises when people who were enumerated misreport their Census Day address when they respond to the PES.

An evaluation of the quality of the reporting of the Census Day address was conducted after the 1986 TARO. A post-production follow-up which reinterviewed a sample of 903 of the nonmatches was aimed at determining the number of nonmatches caused by misreporting mover status. Another search to match respondents who reported they in fact had moved within the test site was made at the new address.

5.3.4 Estimation

The sample cases found to have errors in their reported Census Day address may be used to estimate

L_e = the weighted number of people who erroneously report their Census Day address in their P-sample interview.

A search of census enumerations at the newly reported addresses produces

r_{am} = the estimator of the percentage of people with errors in the location of their reported Census Day address who match census enumerations.

Then the expected value of the error m_a is estimated by

$$\mathrm{E}(m_a) = -\ r_{am}L_e.$$

The results of the post-production follow-up (Hogan and Wolter 1988) yielded a misreporting rate of at most 3.1 percent in the P sample. A match rate of 33 percent was estimated for those who misreported their Census Day address and moved within the test site. If we assume the match rate for those who reported a census day address outside the test site is also 33 percent, then the expected value $\mathrm{E}(m_a) = -3481$.

An estimate of the variance of the error due to misreporting has not been made. Our professional judgment is that a conservative estimate of the variance at the PES sample level is 900. Therefore, the variance at the TARO site level is

$$\mathrm{Var}(m_a) = (17)^2 \times 900 = 260{,}100.$$

5.3.5 Summary

For the total error model, the first two moments of the distribution of the error due to misreporting of Census Day address for the PES sample are assumed to be $\mathrm{E}(m_a) = -3481$ and $\mathrm{Var}(m_a) = 260{,}100$.

5.4 Fabrication in the P sample

5.4.1 Source of Error

Interviewers may fabricate people in P-sample housing units. Research has shown that interviewer fabrication during the PES may result in a substantial bias in the estimates of census coverage error based on the dual system estimator. Basically, the creation of fictitious individuals may decrease the PES match rate, causing the estimate of coverage error to be too large.

Experience at the Bureau of the Census has shown that fabrication of the members of a whole household is the problem for household surveys. Rarely is there a fabrication of the household member in a household where the other members are the real residents.

The quality control operation for the interviewing phase of the P sample is designed to check for fabricated interviews and to interview the real household members. Therefore, no statistical correction for fabrication in the P sample is made in the formation of the dual system estimates.

5.4.2 Definition

The N_{11} and N_{1+} in the dual system estimator are estimated from sample survey data, the P sample. The following were introduced in Section 4:

m_f = the weighted number of people who were replaced by fabricated P-sample interviews and who were enumerated,

n_{pf} = the error in N_{pf} due to households that were fabricated in the P sample.

The posterior expected values and variances of m_f and n_{pf} are denoted by $\mathrm{E}(m_f)$ and $\mathrm{E}(n_{pf})$ and $\mathrm{Var}(m_f)$ and $\mathrm{Var}(n_{pf})$.

5.4.3 Measurement

In the 1986 TARO, the estimate of the fabrication rate based on the quality control of the interviewing was approximately 0.6 percent. The estimate of the fabrication rate based on a post-production follow-up was approximately 1.2 percent (Hogan and Wolter 1988).

5.4.4 Estimation

We now estimate the moments of the posterior distributions of n_{pf} and m_f from the PES sample. We believe it is reasonable to assume n_{pf} is negligible in TARO. Therefore, the expected value and variance are given by $\mathrm{E}(n_{pf}) = 0$ and $\mathrm{Var}(n_{pf}) = 0$.

The quality control data may be used to estimate r_f = the rate at which P-sample interviews are fabricated.

The search of the census enumerations for people in the P sample who were found by the quality control operation to not have been properly interviewed produces r_{fm} = the match rate for people not interviewed because their household was fabricated in the P sample.

In TARO, records were not kept so that the people who were discovered by the quality control not to have been interviewed properly could be identified. Therefore, no search was made for matching enumerations. Since we have no data available for a direct estimate of r_{fm}, we conservatively assume that the people not interviewed properly are like the people who were. We set r_{fm} equal to the final overall P-sample match rate.

We use the conservative results from the post-production follow-up to yield a fabrication rate of 1.2 percent. The match rate for TARO is 88.6 percent (Diffendal 1988). Therefore, the expected value of the error m_f is given by $\mathrm{E}(m_f) = -2502$.

An estimate of the variance of the estimate of fabrication error has not been calculated. Our professional judgment is that a conservative estimate of the variance can be derived by the reasoning discussed in Section 5.4.2. Thus, we estimate that the variance for the TARO site is

$$\mathrm{Var}(m_f) = (17)^2 \times 206 = 59{,}534.$$

5.4.5 Summary

For the total error model, the first two moments of the distribution of the net error due to fabricated interviews are assumed to be $\mathrm{E}(m_f) = -2502$ and $\mathrm{Var}(m_f) = 59{,}534$. The net error due to fabricated interviews in is assumed to be negligible, and therefore, $\mathrm{E}(n_{pf}) = 0$ and $\mathrm{Var}(n_{pf}) = 0$.

5.5 Measurement of Erroneous Enumerations

5.5.1 Source of Error

Some enumerations may have been entered in the census as the result of mistakes. These enumerations are called erroneous enumerations. Since the dual system estimator requires estimating the number of distinct people captured in the census, a correction is made for erroneous enumerations in the estimate of total population. Subtracting the estimate of the number of enumerations that do not correspond to distinct people from the census count provides an improved estimate of the number of distinct people captured in the census. This estimated correction is obtained from the E sample in the PES.

The following types of enumerations are considered erroneous: (1) people who died before Census Day, (2) people who were born after Census Day, (3) enumerations that do not refer to real people, (4) people duplicated, (5) people enumerated outside the search area where the matching operation looks for their enumeration. The search area for a case includes the block for its address and the ring of adjacent blocks.

This component is caused by errors in measuring census error. An error in the estimation of the number of erroneous enumerations occurs either when an enumeration in the E sample is designated as erroneous although it is correct, or when an enumeration is designated as correct although it is really erroneous. Therefore, both positive and negative error can occur in the estimation of the number of erroneous enumerations.

The types of enumerations that are the most vulnerable to misclassifiction as to whether they are erroneous include the duplicated and fabricated enumerations. These errors are the only ones considered because the others are either inconsequential or are treated separately. Errors in identifying enumerations for people who died before Census Day and people who were born after Census Day have a trivial effect. Errors in classifying the eumeration status because a person was enumerated outside the search area is covered in Section 5.6 on balancing the estimates of the gross overcount and the gross undercount.

5.5.2 Definition

The bias in the DSE due to misclassification of enumeration status is caused by error in the estimation of N_{+1}. In the formation of the estimate of the number of distinct people in the original enumeration $\hat{C}$, a correction is made for the number of erroneous enumerations, $\widehat{EE}$. $\widehat{EE}$ and therefore $\hat{C}$ are estimated from sample survey data, the E sample. Errors in the estimate $\hat{C}$ occur through the misclassification of the enumeration status of E-sample cases. Let

c_e = the difference between the weighted number of erroneous enumerations misclassified as correct and the weighted number of correct enumerations misclassified as erroneous.

The expected value of c_e, conditional on the observed value $\hat{C}$, is denoted by $E(c_e)$. The variance of c_e, conditional on the observed value $\hat{C}$, is denoted by $\mathrm{Var}(c_e)$.

5.5.3 Measurement

Processing error may be measured directly using a rematch of a sample of cases. Errors from other sources, such as duplications due to violations of census residency rules, can be assessed by viewing the frequency distributions of the erroneous enumerations. This is preferable to direct measurement of these errors because of the difficulties in obtaining accurate data in additional follow-ups. When tests confirm that the gross errors from these sources are under control, the net error can be assumed to be negligible. For example, the distribution of the erroneous enumerations by age group is expected to have a large number of duplications in the highly-mobile groups of the population where there are more opportunities for the census residency rules not to be followed.

In the 1986 TARO, an evaluation of the E-sample processing was conducted in conjunction with the evaluation of the P-sample matching operation discussed in Section 5.2.3 (Corby and Mulry 1988). The data for the E sample from the same subsample of 35 blocks were reprocessed.

5.5.4 Estimation

We now estimate the moments of the distribution of c_e from the PES sample. The results of the reprocessing (Hogan and Wolter 1988) yield a net error rate of 0.0007 in the identification of correct enumerations. The expected value of c_e is $E(c_e) = -238$. This estimate is based on the E sample with a resolved enumeration status because the error in the imputation for the unresolved cases is covered in the Missing Data Section 5.7.

An estimate of the variance of net error has not been calculated. Our professional judgment is that a conservative estimate of the variance can be derived by the reasoning discussed in Section 5.2.2. Thus, we estimate that the variance for the TARO site is $\mathrm{Var}(c_e) = (17)^2 \times 14 = 4{,}046$.

5.5.5 Summary

For the total error model, the first two moments of the posterior distribution of the net error in identifying correct enumerations are assumed to be $E(c_e) = -238$ and $Var(c_e) = 4,046$.

5.6 Balancing the Estimates of the Gross Overcount and Undercount

5.6.1 Source of Error

Both the E sample and the P sample measure enumeration errors in the census. The E sample measures the gross overcount in the form of erroneous enumerations. The P sample measures the gross undercount in the form of those not enumerated. Ideally, the entire census would be searched before a P-sample person was declared to be not enumerated. Ideally, the entire country would be searched to determine if an E-sample enumeration is a duplicate or fictitious. Of course, such extensive searches are simply not feasible in the performance of the PES. These searches must be limited in the reasonable manner. The way chosen has to preserve the net error although the measured gross overcount and the measured gross undercount may increase due to limiting the search area. The gross overcount and the gross undercount have to balance to equal the net coverage error.

Failure to have procedures which balance the estimated gross overcount and the estimated gross undercount may cause an incorrect number of enumerations in the E sample to be designated as erroneous when they are correct. This error may cause either an upward or downward bias.

Balancing is not an issue for the design of the PES planned for 1990 and tested in the 1986 TARO, as it was in 1980. The design calls for overlapping the P sample and the E sample. The same blocks are included in the P sample as in the E sample. The P-sample search area is, by definition, the proper search area. The E-sample search area is chosen to be consistent with the P-sample search area.

5.6.2 Summary

Error due to geocoding error is believed to be negligible in the 1986 TARO and will not be included in the total error model. The appendix contains a model for balancing error.

5.7 Missing Data

5.7.1 Source of Error

Both the E sample and the P sample have missing data. The E sample has cases where the information required to determine whether the person is correctly or erroneously enumerated in the census is not available. The P sample has cases where the information needed to determine whether the person is enumerated in the census is not available. The probability of being enumerated is imputed statistically to compensate for the inablility to resolve the case.

An unresolved status may occur in more than one way. The interviewer may be unable to obtain an interview during the P-sample interviewing or during the PES follow-up. A P-sample or E-sample questionnaire may not have all the demographic and housing information required for the estimation. Even with all the information requested on the questionnaires, the circumstances may be so unclear that the enumeration status can not be resolved.

5.7.2 Measurement

We assess the error in the DSE caused by missing data instead of considering each component c_i, m_i and n_{pi} separately. Our approach is to perform a sensitivity analysis of reasonable alternative models for compensating for missing data. First a preferred method of imputation for unresolved P-sample and E-sample enumeration statuses is specified prior to the implementation of the PES. Reasonable alternative treatments of the missing data can be suggested by problems that arise during the collection and processing of the PES data. The DSE can be computed under these alternative models for compensating for missing data. The range of the alternative estimates indicates the sensitivity of the DSE to the method of imputation. For example, a narrow range implies that the estimates are robust, and the missing data cause little uncertainty in the estimates.

5.7.3 Estimation

The effect of missing data on the estimates from the 1986 TARO was assessed by examining the range of estimates obtained when methods of imputation based on reasonable alternative assumptions were used in place of the preferred method. These included alternative treatment of proxy responses, movers, and designation of ficticious enumerations (Schenker 1988). The alternative treatment of the proxy interviews for P-sample cases classified them as noninterviews and applied the weighting adjustment. This essentially assigned proxy cases the same match rate as nonproxy cases. The alternative treatment of the P-sample movers reclassified them all as unresolved and imputed a match probability, instead of imputing for only those who were not resolved. This essentially assigned movers the same match rate as nonmovers. The alternative treatment of fictitious cases resulted from a review of the unresolved E-sample cases by experienced matching personnel who converted some unresolved cases to fictitious. This raised both the observed and imputed rates of erroneous enumeration.

Models 000 and 111 shown in Table 4 of Schenker's paper give the upper and lower bounds of the estimates of undercount rates, respectively. Both models differ from TARO in that they have inmovers as substitutes for outmovers. P-sample inmovers are P-sample respondents who moved into their housing unit between Census Day and PES interviewing. In the 1986 TARO the P-sample inmovers from areas outside the test site were omitted from the PES estimation. The omission of the outmovers from estimation essentially assumes that they had the same capture rate in the original enumeration as the included cases. Movers are believed to have a lower capture rate than nonmovers. Model 000 has the TARO treatments while Model 111 has all the alternative treatments.

5.7.4 Summary

The effect of missing data on the distribution of the total error is assessed by computing the distribution of the undercount rate under several reasonable imputation methods. The alternative methods which yield the upper and lower bounds for the undercount are used in the total error analysis.

5.8 Sampling Error

5.8.1 Source of Error

The observed DSE is subject to sampling error because $\hat{N}_p$, $\hat{C}$, and $\hat{M}$ are estimated from samples. The sample size for the PES is determined by the amount of sampling error and budget allowable. Other things being equal, the larger the sample size the lower the amount of sampling error introduced in the estimates. The sampling errror is affected by the estimator and the sampling design. In the TARO PES design, both the P-sample and the E-sample observations are collected from the same sample of blocks. All the people residing in the housing units in the selected blocks are included in the P sample. All enumerations assigned by the census process to the sample block are included in the E sample. The estimation of the sampling error takes into account the tendency for census misses and erroneous enumerations to be correlated within blocks and within housing units. Experience has shown that many hard-to-enumerate areas have both a higher rate of omissions and a higher rate of erroneous enumerations.

5.8.2 Measurement

The standard randomization theory model for survey sampling is appropriate for estimating the variance of the DSE. The coefficient of variation which is the ratio of the square root of the variance of the observed DSE to the mean of the distribution of the DSE provides information on the amount of sampling error in the DSE.

The Taylor series estimator of variance for the observed dual system estimator (Moriarity 1987), $v(\hat{N}_{++})$, is given by

$$v(\hat{N}_{++}) = \hat{N}^2_{++}\,(v(\hat{N}_p)/\hat{N}^2_p + v(\hat{M})/\hat{M}^2 - 2c(\hat{N}_p,\hat{M})/\hat{N}_p\hat{M})$$
$$+ \hat{N}^2_p\, v(\hat{E})/\hat{M}^2 + 2\hat{N}_{++}(\hat{N}_p c(\hat{E},\hat{M})/\hat{M}^2 - c(\hat{E},\hat{N}_p)/\hat{M}),$$

where

$\hat{E} = II_2 + \widehat{EE}$,
$v(X)$ = the estimator of the variance of an estimator X,
$c(X,Y)$ = the estimator of the covariance between X and Y.

The categories II_2, insufficient information for matching, and $\widehat{EE}$, erroneous enumerations, are treated as one group in the variance estimation. The variance and covariance estimators reflect the cluster sampling of blocks and block clusters.

5.8.3 Estimation

The standard deviation of the dual system estimate of 388,040 for the TARO site is 3,100.37. The coefficient of variation is 0.008. This implies the standard deviation for the estimated net undercount rate is 0.7 percent.

5.8.4 Summary

The sampling error for the TARO DSE is 3,100.37, and the sampling error for the TARO net undercount rate estimate is 0.70 percent.

6. SYNTHESIS OF TOTAL ERROR

The combined effect of the component errors will be summarized by posterior distibutions for the net undercount rate. The bias in the estimate of net undercount rate, $B(U)$, is estimated by the difference between and the mean of the posterior distribution. To construct the posterior distribution, we used a simulation method with 10,000 repetitions, generating pseudo-random component errors and adding them to the TARO estimates. Using the formulas in Section 5.1.2, we obtain the following formula:

$$\begin{aligned} N &= (\hat{N}_p - n_p) + (\hat{C} + c - (\hat{M} - m)) \\ &\quad + \theta(\hat{C} - c - (\hat{M} - m))(\hat{N}_p - n_p - (\hat{M} - m))/(\hat{M} - m) \\ &= (\hat{C} - c)(\hat{N}_p - n_p)/(\hat{M} - m) \\ &\quad + (\theta - 1)(\hat{C} - c - (\hat{M} - m))(\hat{N}_p - n_p - (\hat{M} - m))/(\hat{M} - m). \end{aligned}$$

Several different distributions were used to reflect alternative estimates of imputation error, alternative estimates of correlation bias (parameterized by θ), and alternative marginal distributional forms for the components - normal, gamma, and uniform.

In this study, the estimate of percent net undercount for the TARO site is 8.42 with a sampling standard deviation of 0.7. This estimate was selected because estimates of nonsampling error components are available only for the site as a whole. When a DSE is constructed for each post-stratum and then the DSEs are summed to give an estimate for the site, the percent net undercount estimate is 9.02.

Table 6 displays the means and standards deviations of the error components for the PES sample. Recall that the DSE for the TARO site is 388,040, $\hat{M}$ = 298,204, $\hat{C}$ = 343,567, and $\hat{N}_p$ = 336,707. The overall sampling weight, 17, was used consistently throughout all the simulations so that comparisons of the effect of alternative assumptions such as correlation bias parameter values, error distributions, and imputation models are appropriate. The methodology generalizes to other applications where a different sampling weight is used in each stratum.

Table 7 displays the effects of the individual errors on the posterior distribution of the undercount when the TARO imputation is used. The net matching Census Day address, and fabrication errors are all errors in $\hat{M}$. Therefore, the presence of only one of them alone causes the bias in the estimate of percent net undercount to be positive. The net E-sample error is an error in $\hat{C}$. The presence of E-sample error alone causes the bias in the estimate of percent net undercount to be negative. The estimate for correlation bias, was chosen to be 2.7, the median of Ericksen and Kadane's estimates. The presence of only correlation bias causes the bias in the percent net undercount estimate to be negative.

Table 6

Assumed Distributions of Error Estimates

	Mean	Standard Deviation
Net Matching Error	-1831	176
Census Address Error	-3481	510
Fabrication Error	-2502	244
Net E sample Error	-238	64

Table 7

Individual Effects of Errors on Posterior Distribution of Percent Net Undercount and Bias in the Estimate of Undercount

	E(U)	Std. Dev.	B($\hat{U}$)
Net Matching	7.86	0.06	0.56
Census Address	7.35	0.16	1.07
Fabrication	7.34	0.08	1.08
Net E sample	8.49	0.02	-0.07
Correlation Bias (2.7)	10.61	0.00	-2.19

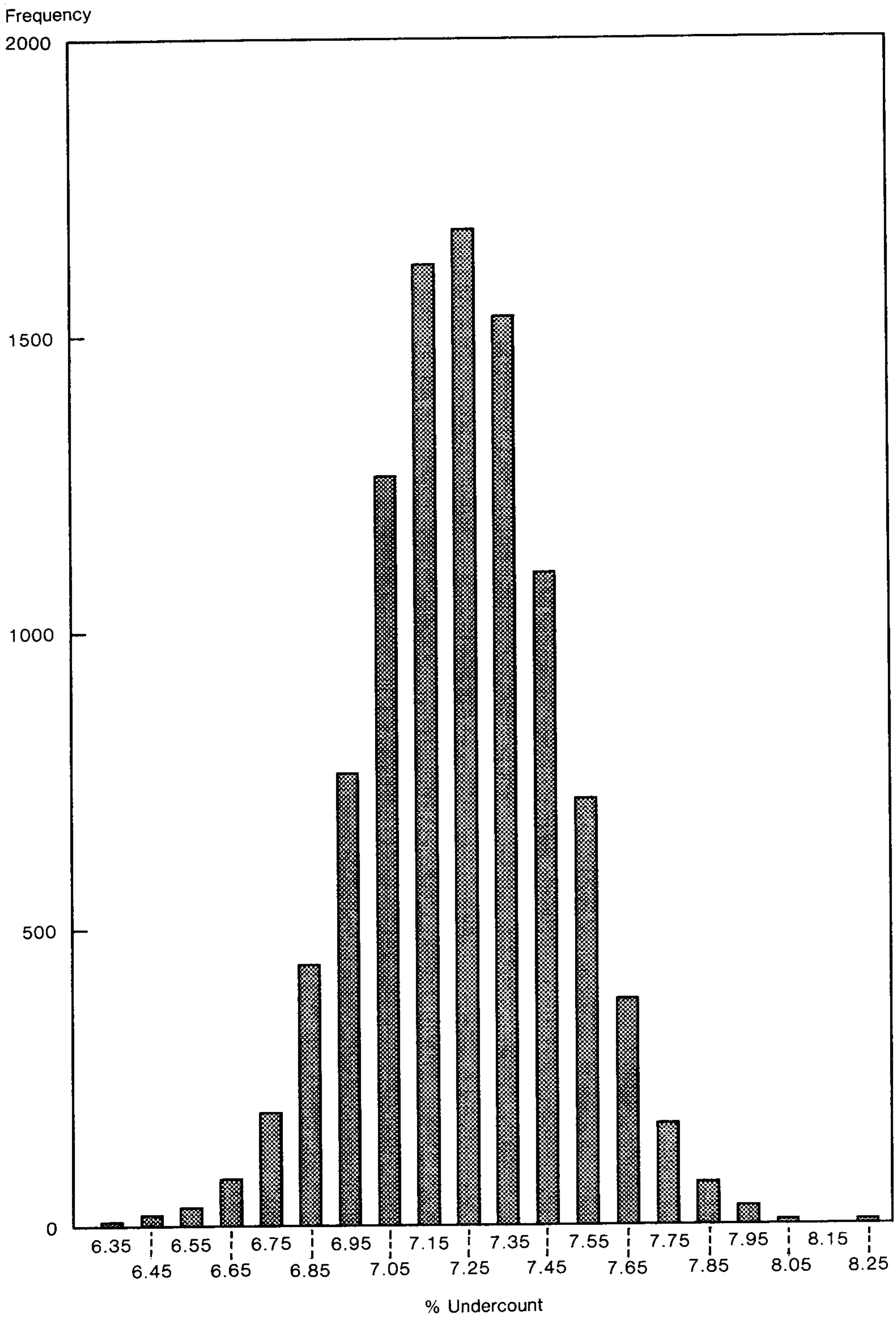

Figure 1. Percent Undercount when $\theta = 2.7$

Table 8

Percentiles of the Posterior Distribution of Percent Net Undercount for θ =2.7

	1	5	10	25	50	75	90	95	99
Normal	6.70	6.86	6.94	7.08	7.24	7.40	7.54	7.63	7.79
Uniform	6.75	6.86	6.93	7.07	7.24	7.42	7.55	7.62	7.73
Gamma	6.67	6.84	6.93	7.08	7.24	7.40	7.53	7.61	7.74

Table 9

Posterior Distribution of the Net Undercount Rate for Several Values of θ

θ	E(U)	St. Dev.	B($\hat{U}$)
1.0	5.75	0.18	2.67
2.1	6.72	0.22	1.70
2.7	7.24	0.23	1.18
3.7	8.09	0.27	0.33

Simulations were conducted where the first two moments for error n_p, c_e, m_m, m_f, m_a, and θ were held constant, but the distributions were varied. We assessed the total error when all the error distributions were normal, all were gamma, and all were uniform. Varying the distributions had minor effects on the distribution of the percent net undercount. In each case the distribution of the percent net undercount was very close to normal. Figure 1 shows the distribution of the undercount when $\theta = 2.7$, and it is illustrative of the results of the simulations.

Table 8 shows the percentiles of the distribution of the net undercount rate for different distributions for the component errors when θ is taken to be 2.7 and the TARO imputation is used. The standard deviation for the posterior distribution was 0.23. In all the cases, a normal distribution is an adequate approximation. The percentiles differed by at most 0.02 for the percentiles between 5 and 95. The 1 and 99 percentiles differed by at most 0.08.

Varying the value of the estimate of θ for the correlation bias did affect the moments of the posterior distribution of the undercount. The variation appears in the mean and in the standard deviation. Table 9 shows the results for the different values of θ, where the distribution for the errors are normal. The case where $\theta = 1$ portrays virtually no correlation bias, while for the other sources of error are present. In the cases where $\theta = 2.1, 2.7$, and 3.7, all the sources of error are taken into account. The distribution of the undercount shifts to the right as the estimate of θ for the correlation bias increases. The variance also increases as the estimate of θ increases. For all values of θ considered, the bias B($\hat{U}$) is positive although it decreases as θ increases.

The simulations were conducted with reasonable alternative models for the imputation for unresolved match status. Although there was some variation in the first two moments of the distribution of the net undercount rate, the estimate of net undercount rate in TARO appears robust to missing data. Table 10 illustrates the results of the simulations using models 000 and 111 described in Section 5.7.3. Models 000 and 111 yielded the upper and lower bounds of the undercount estimates under all the reasonable alternative imputation models. The bias in the estimate of the percent net undercount rate ranges from 0.93 to 2.79. In other words, the bias is between 11 percent and 33 percent of the net undercount rate estimate of 8.42. Varying the imputation model has almost no effect on the standard deviation.

Table 10

Posterior Distribution of the Percent Net Undercount Under Reasonable Alternative Imputation Models When $\theta = 2.7$

	$E(U)$	St. Dev.	$B(\hat{U})$
TARO	7.24	0.23	1.18
Model 000	7.49	0.23	0.93
Model 111	5.63	0.22	2.79

The total variance of the estimated net undercount rate may be estimated by the sum of the sampling variance and the nonsampling variance. For the case where $\theta = 2.7$, the standard deviation shown in Table 10 for both models 000 and 111 is 0.22 which translate to a non-sampling variance of 0.0005 when all errors are considered. The standard deviation of the estimate of net undercount rate is 0.70 which translates to a sampling variance of 0.49. Therefore, the total variance is 0.0054 and standard error is 0.73. The coefficient of variation of the net undercount rate is 0.083. The nonsampling variance contributes very little to the total variance relative to the contribution by the sampling variance.

7. CONCLUSIONS

When the post-stratification is used in the estimation, the undercount estimate for TARO is 9.02. The post-stratification increased the net undercount rate estimate by 0.6, which is less than one standard deviation of 0.73 from the estimate of 8.42. Although we expect the error in the post-stratified estimate is smaller, the result is consistent with the error analysis.

As we consider all the sources of error in the posterior distribution of the net undercount rate, we do not know the distribution of the correlation-bias parameter θ. Although we could assume a prior distribution for θ, others might disagree. If we were certain that θ is 2.7, then our 95 percent confidence interval for the net undercount rate would be

$$4.77 < U < 9.55.$$

We calculate this by taking the post-stratified estimate 9.02 and adjusting for the two bias estimates in Table 10, 2.79 and 0.93, and two standard deviations, 2×0.73. We feel this is a conservative estimate since we use two different bias estimates from imputation models 000 and 111. A very conservative 95 percent confidence interval for U for any value of θ between 2.1 and 3.7 is (4.43,10.32).

We believe the methodology described in this paper is applicable in the 1990 census with appropriate modifications. Areas for further research are nonsampling error estimates for post-strata, a distribution for the correlation-bias parameter, and models for address reporting error.

ACKNOWLEDGEMENTS

The authors wish to thank Aref Dajani for developing the computer software for the simulations and Chris Moriarity for the variance computation. We are also grateful to Kirk Wolter and Howard Hogan for helpful discussions and suggestions during the development of the models. Bruce Spencer thanks the Undercount Research Staff at the Census Bureau for supporting this research under a Joint Statistical Agreement. The authors thank the referees for their helpful comments and suggestions.

APPENDIX

Definition of Balancing Error

The non-linearity of the dual system estimator makes an additive model inadequate for viewing the technical implications of the balancing of the estimated gross overcount and the gross undercount. Therefore, a more appropriated multiplicative model is developed in this section.

Limiting the E-sample and the P-sample search areas affects two parts of the DSE. One effect is a bias in the estimate of the number of erroneous enumerations, $\widehat{EE}$. The other is a bias in the estimate of the number of people in both the census and the P-sample population, $\hat{M}$.

The following definitions are needed for examining the effects of limiting the E-sample and the P-sample search areas in the TARO design on the dual system estimate:

b = the proportion of the correct census enumerations that are in their P-sample search area.

g = the ratio of the number of correct census enumerations that are in their E-sample search area to the number that are in their P-sample search area.

The proportion g reflects error in the implementation of the survey committed when the E-sample search area is not equal to the P-sample search area. The way TARO was executed implies $g = 1$. To show what would happen if g does not equal 1, we will carry g through the discussion.

The limiting of the search area causes only a percentage b of the P-sample people who are in both the census and the P-sample population to be designated as matching a census enumeration. Under these circumstances, a systematic bias equal to $(1 - b)\ N_{11}$ is introduced into the esimation of the number of people in both the census and the P-sample population. Therefore, the observed really estimates bN_{11}.

Likewise, the limiting of the search area causes only a percentage b of the census enumerations to be available to be designated as correct. Then only a percentage g of those, the ones whose search areas are consistent with the proper E-sample search areas, will be designated as correct. Under these circumstances a systematic bias equal to $(1 - bg)N_{1+}$ is introduced into the estimation of the number of distinct people in the census. This bias occurs in the estimation of the number of erroneous enumerations, $\widehat{EE}$. With this formulation, the observed number of distinct people in the census really estimates bgN_{1+}.

If $g = 1$, no systematic bias is present in the estimation of the dual system estimate because $bgN_{+1}N_{1+}/(bN_{11}) = N_{1+}N_{+1}/N_{11}$.

The error in the estimation of N_{+1} due to the failure to balance may be defined by

c_b = the error in the number of erroneous enumerations due to the failure to define the E-sample search areas consistently with the P-sample search areas.

The error c_b would be nonzero if g does not equal 1. The ratio g may be greater than or less than 1. The error is given by $c_b = b(g - 1)N_{+1}$.

Measurement

In TARO, c_b was evaluated by testing to confirm that balancing was not an issue and that the design was under control. The percentage of matching enumerations found within the sample block was large, which implies that the design was under control. Since the design was under control, g is assumed to be approximately 1, and c_b is assumed to be negligible.

Estimation

The geocoding appeared to be very good in the TARO test site. However, no formal measurement of the effects of any misassignment on the estimation of $\widehat{EE}$ was conducted. Therefore, g is assumed to be 1, which implies $\mathrm{E}(c_b) = 0$ and $\mathrm{Var}(c_b) = 0$.

REFERENCES

CHILDERS, D., DIFFENDAL, G., HOGAN, H., SCHENKER, N., and WOLTER, K. (1987). The technical feasibility of correcting the 1990 Census. *Proceedings of the Social Statistics Section, American Statistical Association*, 36-45.

CORBY, C., and MULRY, M. (1988). Memorandum to K.M. Wolter, Subject: Matching Error Pilot Study. Statistical Research Division, U.S. Bureau of the Census, Washington, D.C.

COWAN, C.D., and MALEC, D.J. (1986). Capture-recapture models when both sources have clustered observations. *Journal of the American Statistical Association*, 81, 347-353.

DIFFENDAL, G., (1988). The 1986 test of adjustment related operations in Central Los Angeles County. *Survey Methodology,* 14, 71-86.

ERICKSEN, E.P., and KADANE, J.B. (1985). Estimating the population in a census year: 1980 and beyond. *Journal of the American Statistical Association*, 80, 98-108, 129-131.

HOGAN, H., and MULRY, M. (1987). Operational standards for determining the accuracy of census results. *Proceedings of the Social Statistics Section, American Statistical Association*, 46-55.

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

MORIARITY, C. (1987). STSD 1986 Test Census Memorandum II-12, Subject: Documentation of the Calculation of the Los Angeles Post-Enumeration Survey Block Weights and Dual System Estimate Variances. Statistical Support Division, U.S. Bureau of the Census, Washington, D.C.

SCHENKER, N. (1988). Handling missing data in the estimation of coverage error for the 1986 census of Central Los Angeles County. *Survey Methodology,* 14, 87-97.

SEKAR, C.C., and DEMING, W.E. (1949). On a method of estimating birth and death rates and the extent of registration. *Journal of the American Statistical Association*, 44, 101-115.

SELTZER, W., and ADLAKHA, A. (1974). On the Effect of Errors in the Application of the Chandrasekaran-Deming Techniques. Reprint 14, University of North Carolina, Laboratory for Population Statistics.

SPENCER, B.D. (1986). Conceptual issues in measuring improvements in population estimates. *Proceedings of the Second Annual Research Conference,* U.S. Bureau of the Census, Washington, D.C., 393-407.

SPENCER, B.D. (1980). Implications of equity and accuracy for undercount adjustment: a decision - theoretic approach. *Proceedings of the 1980 Conference on Census Undercount*, U.S. Bureau of the Census, Washington, D.C., 204-216.

WOLTER, K.M. (1986a). Some coverage error models for census data. *Journal of the American Statistical Association*, 81, 338-346.

WOLTER, K.M. (1986b). A Combined Coverage Error Model for Individuals and Housing Units. SRD Research Report Number Census/ SRD/RR-86/27, Statistical Research Division Report Series, U.S. Bureau of the Census, Washington, D.C.