ACCURACY OF 1990 CENSUS UNDERCOUNT ESTIMATES FOR THE POSTCENSAL ESTIMATES

Mary H. Mulry*, U.S. Bureau of the Census and Bruce D. Spencer, Northwestern University and NORC
Mary H. Mulry, U.S. Bureau of the Census, Washington, D.C. 20233

1. INTRODUCTION

The Bureau of the Census is currently considering including new estimates of undercount for the 1990 Census in the postcensal estimates program. These estimates, known as the Post Census Review (PCR) estimates, use the 1990 Post Enumeration Survey (PES) data, but differ from the 1991 estimates in that they have a new poststratification and revisions which correct some errors discovered after July 15, 1991. The revisions make a new assessment of the accuracy of the dual system estimate (DSE) and the census necessary for the decision.

The evaluation of the total error uses a strategy developed by Mulry and Spencer (1988, 1991 and 1992) to assess the overall accuracy of the PES and census estimates of population size, as well as the census undercount rate. The method uses data from the Bureau's comprehensive program in 1991 to evaluate the components of error and the total error in the PES (Bateman, et al, 1991) and from an additional study conducted in 1992.

2. BACKGROUND

We use the same notation as in Mulry and Spencer (1991) to describe the Census Bureau's empirical DSE that is based on observable quanitities. The Census Bureau modifies the census count $\tilde{N}_C$ to account for erroneous and imputed enumerations and persons with insufficient information to allow for a match to obtain, $\hat{N}_{CE}$, the estimated size of the population that could possibly be matched. The size of the P-sample population, N_{1+}, is estimated unbiasedly by the weighted number of P-sample selections, $\tilde{N}_P$. Next, the Census Bureau estimates the weighted number of matches between the P-sample and the census.

The Census Bureau estimates N by the *empirical DSE*, $\hat{N} = \hat{N}_P\hat{N}_{CE} / \hat{N}_{CP}$. The empirical DSE is used to estimate the percent net undercount, or the *net undercount rate*, in the original enumeration, $\hat{U} = 100(\hat{N} - \tilde{N}_C) / \hat{N}$. To actually perform the adjustment, the Census Bureau calculates an empirical DSE for each of 357 poststrata and uses the DSE to calculated an adjustment factor for the poststratum, $f = \hat{N} / \tilde{N}_c$.

The poststratified PES estimates for states and other areas are formed using the adjustment factors. If f_i, i = 1,...357, are the adjustment factors and $\tilde{N}_{Cij}$ are the census counts in the intersection of poststratum i and state j, j = 1,....50, the PES estimate of the population of state j, $\tilde{N}_j$, is estimated by

$$\tilde{N}_j = \sum_i f_i \tilde{N}_{Cij} \tag{2.1}$$

Other areas are treated analogously.

3. MEASUREMENT OF COMPONENTS OF PES ERROR

In this section, we identify the components of error and describe the sources of the evaluation data. Estimates of the bias and variance of the undercount rate derive from estimates of the first two moments of the components of PES error. The components are model bias (correlation bias), matching error, accuracy of the reported Census Day address and other P-sample data collection errors, fabrication in the P sample, E-sample operations error, E-sample data collection error, missing data, sampling error, ratio estimator bias and error due to excluded late census data. For operational reasons, balancing error is not treated separately, but is incorporated in other component errors (e.g., matching error, E-sample errors). Similarly, random nonsampling error estimates, except imputation error, are reflected in the sampling error estimates produced by jackknifing. The second-moment estimates reported below, except sampling error and imputation error, are estimates of the variance of the nonsampling bias estimates.

3.1 Measurement of Error

3.1.1 Model Error Based on Demographic Analysis

We are measuring model bias by comparing the PES estimates of population size with an independent estimate from demographic analysis as is done by Bell (1991). Demographic analysis uses vital records to provide alternative estimates of the population size in April, 1990 at the national level for sex, age, and race groups (Robinson, et al, 1991). Also, demographic analysis calculates alternative estimates of the sex ratios for age and race groups. Although demographic analysis may be subject to its own set of errors, using the sex ratios, as opposed to the estimates of population size themselves, is thought to minimize the effect of such errors.

Bell developed four estimators of model bias, also called correlation bias, using sex ratios from demographic analysis. Each method assumes no model bias for females . However, each method assumes a different parameter for males is constant across PES poststrata within an age-race group. We use the "theta estimator," which assumes that the odds ratio of the capture probabilities is constant over the PES poststrata within an age-sex group (Bell, 1991).

Data-based estimates of the variance of the estimates of were not available at the time the simulations were performed, so subjective estimates were derived. However, subsequent data-based estimates showed the subjective estimates were conservative.

We are examining alternative methods for incorporating model bias in the total error modeling. The present method adds 0.9 million males while Bell's method produces an estimate of model bias of 1.6 million males. Currently we are investigating the discrepancy in the total error results and the independent results from Bell's method.

3.1.2 Matching Error Study and Other Re-Match Studies

We base our estimates of P-sample matching error and E-sample office processing error on the results of re-match studies, the Matching Error Study (Davis, et al, 1991), the Selected Cluster Review, and the Hispanic Cluster Review. The focus for the E-sample was on the error in the assignment of the match codes for duplicated and fictitious enumerations. However, the error in the identification of those born after Census Day and those who died before Census Day also was examined.

In the Matching Error Study in 1991, the highest-skilled personnel conducted the re-match on a sub-sample of 919 block clusters selected for evaluation. The re-matchers also had more time in that they did not have the pressure of PES schedules. They reviewed all the cases in these block clusters and assigned new match codes. The re-match was "dependent" in that the re-matchers had access to PES match codes and the same information as the PES matchers. The re-match was "independent" in that the re-matchers were assigned to work in processing offices where they had not worked during the PES. Therefore, they had not previously seen the cases they worked on during the re-match and were not influenced by the PES operation in the processing office where they were re-matchers.

In the Selected Cluster Review, 104 PES block clusters with highest contribution to the estimated undercount, meaning a large difference in the weighted number of erroneous enumerations and weighted number of nonmatches, were processed again by the highest-skilled matching staff. Thirty-two of the 104 block clusters were in the evaluation sample for the Matching Error Study. Since the results of the re-processing have been added to the undercount estimates for the PCR, the matching error and E-Sample office processing error are assumed to be zero in these 32 block clusters.

In the Hispanic Cluster Review, 100 PES block clusters with high contribution to the Hispanic undercount and which had not been included in the Matching Error Study or the Selected Cluster Review, were re-matched by the highest-skilled matching staff. These blocks comprise a supplement to the evaluation sample for the Matching Error Study. The goal of the re-match of these block clusters was to provide more precise estimates of the matching error for Hispanics since the estimates with the 1991 data had high variances. Therefore, the estimates of matching error and E-sample office processing error are based on 1019 block clusters.

3.1.3 Evaluation Follow-up

The Evaluation Followup assessed the data collection error in the P Sample (West, et al, 1991) and the E Sample (West, 1991). Measurement of the error in the reported census day address and other P-sample errors is based on data collected in the P-sample portion of the Evaluation Follow-up. The sub-sample consisted of the whole household and partial household nonmatches in the 919 block clusters selected for evaluation. A sample of the matches, both whole household matches and partial household matches, were also included as a control group. The sub-sample also included both nonmovers and movers. The questionnaires were the PES Follow-up questionnaire for the matches and nonmatches who had not been to PES Follow-up, and a specially designed Revisit questionnaire for those cases that had been interviewed in the production PES follow-up. The Revisit questionnaire contained more probes concerning the respondent's Census Day address.

The E-sample cases in the 919 block clusters chosen for evaluation who had been in the PES Follow-up were selected for the Evaluation Follow-up. These cases were not matched to P-sample people during the first phase of the PES matching. Therefore, these cases were believed to be the most vulnerable to error. The most experienced and highly trained matching and interviewing personnel performed the Evaluation Follow-up. The matchers and interviewers were not allowed to work on cases that they had been assigned in PES.

3.1.4 Analysis of Reasonable Alternative Imputation Models

The noninterview rate for the P-sample interviews was 1.6 percent: however, 4.3 percent of the P-sample responses were proxy interviews. A weighting adjustment, as opposed to imputations, compensated for these noninterviews. Of the P-sample cases that were interviewed, 1.9 percent were unresolved and their enumeration status had to be imputed. In the E sample, 1.4 percent of the cases had to be imputed.

The source of data for the evaluation is the set of reasonable alternative imputation models developed by Mack (1991). When the preferred method of imputation, the one used in the PES estimation, is included, there are eight reasonable alternative models.

The estimation of the variance due to imputing probabilities used the reasonable alternatives and a Bernoulli-like estimator (Schafer, 1991a). The estimation of the variance due to model selection used the total error simulations with an equal number of replications for each of the eight imputation models.

Mack also has calculated bootstraps of the production PES imputation model. Three E-sample bootstraps and three P-sample bootstraps are available. There are 16 separate bootstrap DSEs when all possible combinations are made, including combinations with the PES model. The estimation of the variance due to parameter estimation in the imputation model uses the 16 bootstraps and an analysis of variance estimator (Schafer, 1991b).

Although the individual components of variance due to missing data were calculated for 1991 estimates, the reasonable alternatives and the bootstrap samples have not been calculated for the PCR estimates. The motivation for not performing the calculations is the limited resources, and the size of the error for the 1991 estimates is small. Therefore, for the PCR estimates, the variance due to missing data V_M is assumed to be six percent of the random error S^2 which is the average percentage observed in the 1991 estimates.

3.1.5 PES Sample

Stratified jackknife estimators of sampling variance and covariances for the DSE are estimated by the VPLX computer program (Fay, 1990). For the smoothed estimates, the covariance matrix is calculated using the results of the model fitting (Isaki, et al, 1991).

The ratio estimator bias is estimated using the entire PES sample and the stratified jackknife estimator from the computer program VPLX (Fay, 1984). The variance is estimated by assuming a coefficient of variation for the estimate of the ratio estimator bias of 0.10. The motivation for the choice of the coefficient of variation is that the bias estimate has about the same relative reliability as the PES variance estimate itself. Since the PES variance estimates for each evaluation poststrata are based on an average of approximately 400 algebraic degrees of freedom, the stability of the variance estimate is comparable approximately to a chi-square variate on 200 degrees of freedom, which has a coefficient of variation of 10 percent.

3.1.6 Evaluation of Excluded Late Census Data

Data entered into the census in November and December of 1990 is known as "late late census data (LLCD)". This data included both additions and deletions of enumerations. Most of the data was the result of census coverage improvement programs such as the local review program, the search/match process, and the parolees and probationers program. The data was too late for the routine PES processing. However, the data for block clusters with three or more changes were included in the PES by special processing. Data in block clusters with two or less changes were not included. In PES estimation, the excluded data were assumed to be in error at the same rate as the E sample as a whole.

An evaluation (Alberti, 1991) examined the effect of not processing the excluded late census cases on the DSE. The 190 cases in the evaluation sample of 919 block clusters underwent office processing but not field followup. As a result 67 percent were unresolved. Imputations for the unresolved cases were based on the imputations for the late census data that were processed during the PES.

The method for estimating the bias due to the excluded late census data assumes that the error in the DSE is the net error calculated at the national level in the evaluation, 170,000 people. The bias is then distributed synthetically throughout the poststrata according to the number of additions to the E Sample from the LLCD that was processed during the PES. The variance of the net error is based on estimates of the variance of the net error calculated in the evaluation.

3.3 Estimation of Components of Error

In some cases, the evaluation samples are not large enough to support reliable direct estimates of moments of component errors at the poststratum level. In those cases we use model-based estimates to smooth the direct estimates. The smoothing technique we consider is synthetic estimation. In other cases (e.g., model bias, sampling errors, and imputation errors for all poststrata and all nonsampling errors for evaluation poststrata), we are able to estimate the moments directly.

There are four methods using synthetic estimation. Two methods apply synthetic estimation to the net component errors, and the other two use the gross errors.

4. TOTAL ERROR ANALYSIS

The estimates of the net bias and a component of the covariance matrix of the DSEs are based on simulations with 1000 replications for the 357 PES poststrata. The covariance matrix reflects imputation error, V_M, the variance of the nonsampling bias estimates, V_{ns}, the variance of the estimate of the ratio estimator bias, V_R, and the variance of the estimate of the error due to excluded late census data, V_L. The sampling errors are included in S^2 because its estimation uses the jackknife. The difference between the observed DSE and the mean of the replicated values is used to estimate the net bias of the DSE and the net bias of the estimated undercount rate $\hat{B}(\hat{U})$. The variance of all the replicated values estimates the sum $V_M + V_{ns} + V_R + .V_L$.

We then form 95 percent confidence intervals for U as $(\hat{U} - \hat{B}(\hat{U}) - 2V^{1/2}, \hat{U} - \hat{B}(\hat{U}) + 2V^{1/2})$ with $V = S^2 + V_{ns} + V_M + V_R + V_L$.

To estimate the total error for the DSE for other units, such as states, we first use the method above to derive estimates of the vector of biases and the covariance matrix for the vector of adjustment factors for the 357 PES poststrata. Then the bias and variance of the DSE may be readily estimated from (2.1). To estimate the errors in census estimates, we simply use the DSE minus its estimated bias minus the census figure; the standard error of the estimated bias in the census is $V^{1/2}$.

We also have used the simulation methodology to examine the individual effect of the sampling and nonsampling errors on the undercount rate at the national level when all other errors are assumed to be zero. Using the root mean square error (RMSE) in Table 1 to rank the error sources, we see that the major sources of error in the national estimates are P-sample data collection error, correlation bias, and E-sample data collection error. All of the RMSEs are in the neighborhood of 0.3 to 0.4. P-sample matching error, E-sample operations error, and sampling error have medium RMSE. Ratio estimator bias, imputation error, excluded late census data and the P-sample fabrication error have the smallest RMSE.

Table 1 Individual Effect of Errors on Bias, Standard Deviation, and Root Mean Square Error of Undercount Rate for the U.S. When All Other Errors Are Assumed to be Zero $\hat{U}=1.61$

Errors	$\hat{B}(\hat{U})$	Std. Dev.	$(MSE)^{1/2}$
Matching	0.22	0.03	0.22
P-Sample Collection	0.32	0.06	0.33
P-Sample Fabrication	-0.01	0.02	0.02
E-Sample Collection	-0.28	0.04	0.28
E-Sample Operations	0.20	0.03	0.20
Model Bias	-0.41	0.15	0.43
Ratio Estimator Bias	0.07	0.00	0.07
Sampling	0.00	0.19	0.19
Imputation	0.00	0.08	0.08
Excluded Late Census	-0.07	0.01	0.07

Table 2 displays the 95 percent confidence intervals in addition to the estimates of the nonsampling bias, the standard deviation of the nonsampling bias, and the total standard deviation for the 10 evaluation poststrata and the U. S. The confidence intervals for evaluation poststrata 4, 6, and 8 do not cover zero. These are all evaluation poststrata for renters. The confidence interval at the national level also does not cover zero. In Table 2 the lowest coefficient of variation for $\hat{B}(\hat{U})$ at the evaluation poststratum level is 0.25. The coefficient of variation for $\hat{B}(\hat{U})$ at the national level is 0.45.

Table 2 Total Error of the Net Undercount Rate Assuming Synthetic Estimation of Net Component Errors

EPS	$\hat{U}$*	$\hat{B}(\hat{U})$	St. Dev. $\hat{B}(\hat{U})$	Total St. Dev.	95% Interval
1	-0.50	0.31	0.99	1.06	(-2.94, 1.32)
2	0.11	0.18	0.25	0.34	(-0.76, 0.62)
3	-0.22	0.81	0.88	1.00	(-3.03, 0.97)
4	2.33	-0.68	0.76	1.07	(0.87, 5.16)
5	2.92	1.54	0.84	1.14	(-0.90, 3.65)
6	5.30	-0.12	0.90	1.45	(2.52, 8.31)
7	1.33	0.80	0.45	0.68	(-0.83, 1.87)
8	7.13	-1.37	1.30	1.54	(5.42, 11.56)
9	2.07	2.23	0.95	1.28	(-2.71, 2.41)
10	6.44	3.55	1.05	1.70	(-0.50, 6.28)
U.S.	1.61	0.35	0.33	0.38	(0.50, 2.03)

*Based on post-stratified DSE.

The calculations are made assuming that the mean and variance of the model bias factor are both zero show the confidence intervals for the same evaluation poststrata covering zero. Although the estimated bias increases to 0.49 at the national level, the confidence interval still does not cover zero.

5. LOSS FUNCTIONS AS MEASURES OF ACCURACY

Examination of confidence intervals, is one method of evaluating the accuracy of the undercount estimates. Another method of evaluating the undercount estimates is using loss functions to examine the accuracy of the distribution of the population estimates from the PES and the census. Loss functions provide a conceptual framework for comparing estimators of population (Mulry and Hogan, 1986 and Spencer, 1986). The total error simulations provide estimates of a target population value for such an analysis comparing the census and adjusted estimates

Let X and θ denote vectors whose ith elements are X_i and θ_i, with X_i and θ_i denoting an estimate and its target value, respectively, for unit i, $1 \leq i \leq n$. For a summary measure of the error in X as an estimate of θ we will use a real-valued loss function $L(X, \theta)$. Using loss functions as analytical representations of the preference ordering, we will say that X is more accurate than Y if the expected value of $L(X, \theta)$ is less than the expected value of $L(Y, \theta)$. We will also use the expected value of the loss function, or its risk, as a quantitative measure of the inaccuracy of the statistic.

In keeping with governmental tradition (Office of Federal Statistical Policy and Standards 1978), the Census Bureau used loss functions of the form

(5.1) $L(X, \theta) = \Sigma_{i=1}^{n} |X_i - \theta_i|$,

(5.2) $L(X, \theta) = \Sigma_{i=1}^{n} w_i |X_i - \theta_i|$,

(5.3) $L(X, \theta) = \Sigma_{i=1}^{n} (X_i - \theta_i)^2$,

or

(5.4) $L(X, \theta) = \Sigma_{i=1}^{n} w_i (X_i - \theta_i)^2$,

taking a unit's estimate and target value to be the unit's population share. In (5.2), w_i is taken to be $1/X_i$. In (5.4), w_i equals $1/X_i$ for weighted squared error or $1/X_i^2$ for relative squared error.

Using the average of the 1000 total error simulations as the target when estimating the expected loss, or risk, produces a biased estimate. The bias occurs because we are using an estimated target and not the true value. We calculate a correction using a bootstrap. Fay (1992) has criticisms of the bootstrap bias corrections for the absolute error loss functions. They tend to overestimate the loss for both estimators, but particularly for the census.

For states, Table 3 shows the range of estimates of the difference in loss for squared error, relative squared error, weighted squared error, absolute error, and relative absolute error over all four models of bias. The range covers the four ways of modeling the component errors. We also calculated the loss functions when correlation bias was assumed to be zero and with our estimate of correlation bias.

Table 3 Differences in Expected Loss ($x10^6$)

Errors	No Model Bias	With Model Bias
Squared	1.5 - 2.5	2.7 - 3.6
Relative Squared	-236 - 880	604 - 1601
Weighted Squared	19 - 32	36 - 48
Absolute	2281 - 3112	3667 - 3887
Relative Absolute	19107 - 113658	82948 - 136926

A positive value says that the PES has less loss. A negative value says that the census has less loss. The loss was negative for the relative squared error for only one of the four ways when correlation bias was assumed zero.

The next question is: Are these numbers different from zero? We have computed their variances using a bootstrap. However, since we are performing an hypothesis test and estimation at the same time, we are unclear as to the appropriate significance level. Royce (1992) has developed an hypothesis test for the census adjustment decision in Canada which addresses this issue.

6. SUMMARY

We have applied a methodology for organizing and summarizing information about sources of sampling and nonsampling error in the 1990 census and PCR estimates of population from the PES. We have synthesized the sources of error into a description of their overall effect on the PES estimates of census undercount. The synthesis is in the form of a 95 percent confidence interval for net undercount rate. Examination of confidence intervals is one method of evaluating the accuracy of the undercount estimates. Another method of evaluating the undercount estimates is using loss functions to examine the accuracy of the distribution of the population estimates from the PCR and the census. The total error simulations have provided estimates of a standard for such an analysis comparing the census and adjusted estimates

REFERENCES

Alberti, N. (1991) Post Enumeration Survey Project P18: Evaluation of Late Late Census Data in the 1990 Post Enumeration Survey. Coverage Studies and Evaluation Memorandum Series #T-1, Statistical Support Division, U. S. Bureau of the Census, Washington, D.C.

Bateman, D., Clark, J., Mulry, M., and Thompson, J. (1991) 1990 Post-Enumeration Survey Evaluation Results, *Proceedings of the Social Statistics Section*, American Statistical Association, 21 - 30.

Bell, W. (1991) Post Enumeration Survey Evaluation Project P13: Use of Alternative Dual System Estimators to Measure Correlation Bias. Coverage Studies and Evaluation Memorandum Series #O-3, Statistical Support Division, U. S. Bureau of the Census, Washington, D.C.

Cochran, W.G. (1977) *Sampling Techniques*, third edition, New York, Wiley.

Davis, M., Biemer, P., Mulry, M., and Parmer, R. (1991) The Matching Error Study for the 1990 Post Enumeration Survey, *Proceedings of the Survey Research Section*, American Statistical Association, 248 - 253.

Diffendal, G. and Belin, T. (1991) Results of Procedures for Handling Noninterviews, Missing Characteristic Data, and Unresolved Enumeration Status in 1990 Census/Post-Enumeration Survey, STSD Decennial Census Memorandum Series #V-112, Statistical Support Division, U. S. Bureau of the Census, Washington, D. C.

Fay, R. E. (1992) "Absolute Values in the PES Loss Function Analysis." Memorandum for the Committee on Adjustment of Postcensal Estimates dated June 29, 1992, Washington, D.C., U. S. Bureau of the Census.

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

Fay, R.E. (1984) Properties of Estimates of Variance Based on Replication Methods, *Proceedings of the Section on Survey Research*, American Statistical Association.

Isaki, C.T., Huang,E.T., and Tsay, J.H. (1991) Smoothing Adjustment Factors from the 1990 PES, *Proceedings of the Social Statistics Section*, American Statistical Association, 338 - 343.

Mack, S. (1991) An Analysis of the Effect of Reasonable Imputation Alternatives on Estimates of Coverage Error in the 1990 Decennial Census, *Proceedings of the Survey Research Section*, American Statistical Association, 652 - 657.

Mulry, M. H. and Hogan, H. (1986) "Research Plan on Census Adjustment Standards". *Proceedings of the 1992 Annual Research Conference*, Washington, D.C., U. S. Bureau of the Census, 381-392.

Mulry, M. H. and Spencer, B.D. (1992a) Accuracy of Undercount Estimates for the 1990 Census, *Proceedings of the 1992 Annual Research Conference*, Bureau of the Census Washington, DC in press.

Mulry, M.H. and Spencer, B.D. (1991) Total Error in PES Estimates of Population, *Journal of the American Statistical Association*, 86, 839-855.

Mulry, M.H. and Spencer, B.D. (1988) Total Error in the Dual System Estimator: The 1986 Census of Central Los Angeles County, *Survey Methodology*, 14, 241- 263.

Robinson, G., Ahmed, B., Das Gupta, P., and Woodrow, K. (1991) Estimating Coverage of the 1990 United States Census: Demographic Analysis, *Proceedings of the Social Statistics Section*, American Statistical Association, 11 - 20.

Royce, D. (1992) To Adjust or Not to Adjust? A Comparison of Some Estimators of a Set of Population Totals, unpublished manuscript. Statistics Canada, Ottowa.

Schafer, J. (1991a) Memorandum to D. Bateman and M. Mulry, Subject: Calculation of Imputation Variance for PES Evaluation Project P16. Statistical Support Division, U.S. Bureau of the Census, Washington, DC.

Schafer, J. (1991b) Memorandum to D. Bateman and M. Mulry, Subject: Calculation of Bootstrap /Variance for PES Evaluation Project P16. Statistical Support Division, U.S. Bureau of the Census, Washington, DC.

Spencer, B. D. (1986) Conceptual Issues in Measuring Inprovements in Population Estimates, *Proceedings of the Second Annual Research Conference*. Washington, D.C., U. S. Bureau of the Census, 393-407.

West, K.K. (1991) Post Enumeration Survey Evaluation Project P9A: Accurate Measurement of Census Erroneous Enumerations -- Evaluation Followup. Coverage Studies and Evaluation Memorandum Series #K-2, Statistical Support Division, U.S. Bureau of the Census, Washington, DC.

West, K., Mulry, M., Parmer, R., and Petrik, J. (1991) Address Reporting Error in the 1990 Post-Enumeration Survey, *Proceedings of the Survey Research Section*, American Statistical Association, 336 - 241.

Acknowledgement

The authors wish to thank George Train for excellent computer software development, Mary Davis and Phil Gbur for important contributions during the data processing, and Lawrence Bates and Alfredo Navarro for expert computational work. Bruce Spencer thanks the Census Bureau and Northwestern University for supporting this work under a Joint Statistical Agreement.

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