

GAO

July 1987

DECENNIAL CENSUS

Coverage Evaluation and Adjustment Activities



RESTRICTED---Not to be released outside the General Accounting Office except on the basis of specific approval by the Office of Congressional Relations.

RELEASED



United States
General Accounting Office
Washington, D.C. 20548

General Government Division

B-222824

July 9, 1987

The Honorable Mervyn M. Dymally
Chairman, Subcommittee on Census
and Population
Committee on Post Office and Civil Service

The Honorable Constance A. Morella
Ranking Minority Member
Subcommittee on Census and Population
Committee on Post Office and Civil Service
House of Representatives

In your letter dated March 25, 1987, you requested that we assist the subcommittee in its evaluation of the Census Bureau's (Bureau) decision relating to coverage evaluation and adjustment of the 1990 census figures. In a subsequent meeting with subcommittee representatives, we were asked to report on the Bureau's 1980 experience and the 1990 activities to date, including timing and costs. This fact sheet responds to your request. Appendix I contains information on the Bureau's coverage evaluation activities for the 1980 census, and in appendix II we discuss the Bureau's coverage measurement and adjustment activities for the 1990 census.

This fact sheet was prepared on the basis of interviews with officials of the Bureau; our review of Bureau documents including evaluation and adjustment plans, pretests' results and evaluations, and Bureau papers presented to outside advisory groups; and our review of prior census hearings. In addition, we attended meetings of the Census Advisory Committees and the National Academy of Sciences Panel on Decennial Census Methodology. We also used information gathered during our previous work on the 1980 census. However, we did not evaluate or verify the information.

Coverage evaluation measures errors in the census counts, such as missed persons and double counting of others. Because census data is used to apportion seats in the House of Representatives, potentially redistrict state legislatures, and distribute federal funding, the Bureau is considering an adjustment of the 1990 census counts to correct these errors.

The objective of the 1980 coverage evaluation program was to measure the errors in the census and to possibly adjust the count. The Bureau's original plan included 1) a post enumeration survey (PES), 2) two Current Population Survey (CPS) samples, and 3) demographic analysis. However, the Bureau did not conduct the planned PES mainly because the resulting data would not have been available until December 1981, several months after the July 1981 deadline for data estimates for revenue sharing distribution. Consequently, the Bureau used 1) a Post Enumeration Program (PEP) which included the originally planned CPS samples and an added reinterview sample, and 2) demographic analysis.

The PEP and demographic analysis did not provide sufficiently accurate results to measure the national undercount for adjustment purposes because of various limitations. For example, the Bureau could not choose from among 12 sets of PEP results so it did not produce a final, definitive estimate of the census errors. Also, demographic analysis produced estimates only at the national level and could not produce estimates at state or substate levels. Because of such limitations on the accuracy of error estimates, the Bureau decided that an adjustment could add more error to the counts than it corrected.

In carrying out its overall objective to obtain an accurate and complete population count in the 1990 census, the Bureau plans to 1) conduct the most complete census possible and 2) concurrently prepare to adjust the counts by the legally mandated deadline of December 31, 1990, if an adjustment would improve the counts. Based on research and testing, the Bureau has selected a large-scale PES of 300,000 households and demographic analysis as the methodologies for coverage measurement.

To overcome the problems experienced in 1980, the Bureau has planned major improvements for 1990, such as the development of an automated matching system designed to accelerate matching PES and census records. The Bureau planned to decide by May 31, 1987, on the technical and operational feasibility of adjusting the 1990 census counts, but as of June 30, 1987, the Bureau had not made its decision.

The 1986 Test of Adjustment Related Operations (TARO) in Los Angeles was the only test of adjusting the census counts conducted before the Bureau's scheduled 1987

decision on the technical and operational feasibility of adjustment. The adjustment of the Los Angeles counts was not completed until almost 3 months after the scheduled deadline because of operational problems. Partially on the basis of this test, the Bureau tentatively has concluded that, given the present planned timing of the basic census and PES enumerations, it will not be able to complete an adjustment by the December 31, 1990, deadline. The Bureau is reviewing alternative schedules to identify changes which might enable it to meet this deadline.

As arranged with your office, unless you publicly announce its contents earlier, we plan no further distribution of this document until 30 days from the date of its issuance. At that time, we will send copies to the Senate Subcommittee on Federal Services, Post Office and Civil Service, Committee on Governmental Affairs; other appropriate congressional committees; the Secretary of Commerce; and the Director, Office of Management and Budget. Copies will also be made available to other interested parties upon request. If there are any questions about the information presented, please call me on 275-8387.

Frank W. Todd

for Gene L. Dodaro
Associate Director

Contents

	<u>Page</u>
APPENDIX	
I	SUMMARY OF THE 1980 COVERAGE EVALUATION PROGRAM 5
	Objective of Coverage Evaluation 5
	1980 Planned Coverage Evaluation 5
	Methodology
	Changes to the 1980 Coverage 6
	Evaluation Methodology
	Results of the 1980 Evaluations 7
	Limitations of the 1980 Methodology 8
	Bureau's Decision Not to Adjust 9
	1980 Counts
II	SUMMARY OF 1990 COVERAGE MEASUREMENT AND 11
	ADJUSTMENT ACTIVITIES
	Dual Strategy for the 1990 Census 11
	Planning for an Adjustment of 1990 Counts 11
	Coverage Measurement Methodologies 12
	Major Improvements Planned for 1990 12
	PES Testing Activities to Date 15
	Test of Adjustment Related Operations 15
	Bureau Concerns about the Timing of 16
	the Census and PES
	Costs for the 1990 PES 17

Abbreviations

CPS	Current Population Survey
PEP	Post Enumeration Program
PES	Post Enumeration Survey
SMSA	Standard Metropolitan Statistical Area

SUMMARY OF THE 1980 COVERAGE
EVALUATION PROGRAM

Objective of Coverage Evaluation

- The objective of the coverage evaluation program for the 1980 census was to measure the errors in the census counts (errors in the census counts are persons missed in the census and persons incorrectly counted¹) and to possibly adjust the count.

1980 Planned Coverage Evaluation Methodology

- The Bureau's originally planned methodology for the 1980 coverage evaluation program included 1) a post enumeration survey (PES), 2) two Current Population Survey (CPS) samples, and 3) demographic analysis.
- This methodology, particularly the PES, was to be used for estimating the errors in the census for national and subnational levels (e.g., regions, states, large cities).

Post Enumeration Survey

- The PES was a large-scale survey designed to compare independent interviews of a sample of 250,000 households to census records. The persons listed in the survey were to be matched with the census listings to estimate the number of persons missed in the census and persons incorrectly counted.
- The PES was to provide census error estimates for all states, 26 large cities--each with a population over 500,000--and their standard metropolitan statistical areas (SMSA)² and for six additional cities and their

¹A person or household is incorrectly counted if the household is counted in the wrong block, the individual is counted in the wrong household, the individual was born after or died before Census Day, or the individual is counted more than once.

²An SMSA was defined by the Bureau as a county containing at least one city with 50,000 inhabitants or more, or several economically and socially related contiguous counties with at least one city of 25,000 inhabitants or more.

SMSAs which had a Black or Hispanic population over 40 percent.

Current Population Survey Samples

- The April and August 1980 samples of 150,000 households from the Current Population Survey (CPS)³ were to be matched to the census to estimate the number of persons missed in the census.

Demographic Analysis

- Demographic analysis was to combine data sources independent of the census to estimate the true population. A comparison of the estimated population and the actual census count was to provide national estimates of census errors for age, sex, and race groups.
- Data sources were to include birth registration, death records, immigration and emigration estimates, medicare rolls, and previous censuses.

Changes to the 1980 Coverage Evaluation Methodology

- The Bureau changed the originally planned coverage evaluation methodology by dropping the large-scale PES and adding a reinterview sample. The 1980 coverage evaluation program thus consisted of 1) a Post Enumeration Program (PEP) using two CPS samples totaling 150,000 households and a reinterview sample of 110,000 households from the census and 2) demographic analysis.

PES Dropped

- In early April 1980, the Bureau decided to drop the PES mainly because the resulting data would not have been available until December 1981, after the July 1981 deadline for data estimates for revenue sharing distribution. In comments on a GAO report, Procedures to Adjust 1980 Census Counts Have Limitations (GGD-81-28, Dec. 24, 1980), the Bureau stated that the decision to drop the PES would result in producing coverage error

³The CPS is a monthly statistical series to collect current data about the U.S. labor force for use in producing monthly unemployment data.

data sooner, cost savings, reduced respondent burden, and the possibility of more accurate results. The Bureau stated that the limitations cited in the GAO report may have been more severe in the PES.

- The Bureau believed that it would save about \$8 million by dropping the PES.
- The Bureau acknowledged that dropping the PES would result in a precision loss in the estimates of census errors. The lost precision would be greater below the largest state and city levels. However, the Bureau stated that the small losses in precision would be more than offset by the gains in timing and cost savings.

Reinterview Sample Added

- In part to compensate for dropping the PES, the Bureau added a reinterview sample of 110,000 households. Households from the census were reinterviewed to provide estimates of the number of persons counted incorrectly.
- The PEP combined the results of the CPS-census match with the reinterview sample results to estimate the net census error rates for the nation, states, and selected large cities.

Results of the 1980 Evaluations

- The Bureau was unable to provide precise data on the errors in the 1980 census.

Post Enumeration Program

- The PEP produced net undercount rates for the nation by age, sex, race, and Spanish-origin categories, and for four regions, 50 states, and 16 cities, without detail about age, sex, race, and Spanish-origin.
- The Bureau initially produced 29 different sets of PEP estimates to reflect alternative treatments of missing data. Discarding extreme estimates, the Bureau narrowed the PEP estimates down to 12 sets. The Bureau could not choose among the 12 sets so it did not produce a final, definitive estimate of the census errors.
- Preliminary PEP estimates of census coverage ranged from an overcount of 1.0 percent to an undercount of 2.0

percent for the total population in 1980; for the Black population, from an undercount of 0.7 percent to an undercount of 7.2 percent; for the non-Black Hispanic population, from an overcount of 0.2 percent to an undercount of 7.6 percent.

- The Bureau estimated that the 1980 PEP cost about \$14 million.

Demographic Analysis

- Preliminary demographic analysis data published in February 1982, showed an overcount of 0.4 percent for the total population in 1980. After incorporating new data and revisions, the Bureau's latest analysis, as of December 1985, shows a net undercount of 1.0 percent for the legally resident population.

Limitations of the 1980 Methodology

- According to the Bureau, the PEP and demographic analysis evaluations did not provide sufficiently accurate results to measure the national census errors for adjustment purposes, although the Bureau does consider the evaluations useful for census planning.

Post Enumeration Program

- The PEP was an experimental program designed to provide general estimates of the 1980 census errors. The PEP estimates were not accurate enough for use in adjusting the census counts.
- According to the Bureau, the CPS and the census tend to miss the same people, so the estimate of census error may be understated.
- High noninterview rates and missing data in the CPS and reinterview samples resulted in unresolved cases during the matching operation. The Bureau stated that the significant number of unresolved cases was a major problem with the PEP.
- Problems in locating addresses obtained during the CPS and reinterview slowed the PEP processing. According to the Bureau, many CPS interviews had insufficient information to be matched to the census.

- Clerical matching was time-consuming, inconsistent, and error-prone. Clerks had difficulty matching both CPS and reinterview cases with insufficient information to the census.

Demographic Analysis

- Demographic analysis produces estimates only at the national level and cannot produce estimates at state or substate levels.
- Initial demographic analysis did not include an allowance for the number of undocumented aliens in the population due to the lack of an accurate measure of the number and distribution of undocumented aliens in the country. Based upon subsequent research, the Bureau estimated that, of those undocumented aliens in the country, about 2 million were counted in the 1980 census. Revised demographic analysis included alternative estimates of the population and undercount assuming the total population includes 1) only legal residents, 2) legal residents and 2 million undocumented aliens, 3) legal residents and 3 million undocumented aliens, or 4) legal residents and 4 million undocumented aliens.
- Demographic analysis cannot produce accurate estimates of the Hispanic population, because many administrative records, such as birth certificates, do not always record whether a person is Hispanic.

Bureau's Decision Not to Adjust 1980 Counts

- Because of the limitations on the accuracy of the census error estimates, the Bureau decided that an adjustment could add more error to the counts than it would correct, especially at the subnational level.
- The Bureau stated that there was no statistically defensible method of distributing the national level undercount to subnational levels. The Bureau also stated that, in planning for the 1980 census, it conducted little research on distributing census error estimates to substate areas and incorporating the adjustment into the census tabulations.
- In the December 16, 1980, edition of the Federal Register, the Bureau announced its decision not to adjust the 1980 census population totals unless directed to do

so by the courts. This decision was based on extensive consultations with Congress, public officials, census advisory groups, professional organizations, and other interested parties.

- According to the Bureau's Office of the Chief Counsel, 36 lawsuits were filed against the Bureau regarding adjustment of the 1980 census. These lawsuits allege that undercounts prevent equal political representation and equitable fund distribution. As of June 26, 1987, 22 lawsuits were pending. Of the other cases, 1 case was decided in the Bureau's favor, and 13 cases were dismissed.

SUMMARY OF 1990 COVERAGE MEASUREMENT
AND ADJUSTMENT ACTIVITIES

Dual Strategy for the 1990 Census

- The Bureau will attempt to 1) conduct the most complete census possible and 2) concurrently prepare to adjust the census counts by December 31, 1990, if an adjustment will improve them.

Planning for an Adjustment of 1990 Counts

- Because the 1980 coverage evaluation program did not provide results accurate enough to adjust the census, the Bureau established a research program on adjustment for the 1990 census.
- The program included a number of preparations that the Bureau believed necessary to implement an adjustment. These included 1) establishing methodologies to accurately measure census coverage errors, 2) developing acceptable statistical techniques to estimate coverage for small geographic areas and for a variety of population and housing characteristics, and 3) establishing and publishing technical standards for evaluating the quality of the adjusted and unadjusted data.

Decisions on Adjustment

- The Bureau planned to decide by May 31, 1987, on the technical and operational feasibility of adjusting the 1990 census counts, but as of June 30, 1987, the Bureau had not made its decision. This decision will represent the Bureau's judgment about whether a large-scale adjustment program should be conducted. It will not be a decision about whether 1990 census counts should be adjusted.
- In December 1990, the Bureau plans to decide whether to release the adjusted data as the official census results. As the basis for this decision, the Bureau plans to implement the adjustment and compare the adjusted and unadjusted data to the standards. The Bureau plans to then release the adjusted data if an adjustment would improve the census counts.

Coverage Measurement Methodologies

- Based on research and testing, the Bureau has selected the following methodologies for the coverage measurement program: 1) a post enumeration survey (PES) and 2) demographic analysis. The census error estimates from the post enumeration survey and demographic analysis would be the basis for adjusting the 1990 census counts.

Post Enumeration Survey

- The planned PES of about 300,000 households will be the primary source for detailed data about census errors. The PES will provide data on people missed in the census and people counted incorrectly, on subgroups of the population, and for substate areas.
- In the PES, the Bureau plans to interview a sample of households independent of the census. The Bureau will then match each person counted in the sample to the census records to determine whether that person was counted correctly or missed in the census.

Demographic Analysis

- As in 1980, demographic analysis will be a method of coverage measurement. The Bureau will develop demographic estimates of the size and distributions of 1990 census errors at the national level. These national estimates will be combined with the PES estimates to measure the errors in the census.
- Preliminary demographic estimates by age, sex, and race (White, Black, Other) are scheduled to be available by October 1990, for adjustment purposes. These initial estimates will represent updates of the 1980 demographic estimates.

Major Improvements Planned for 1990

Independent Block Sampling for the PES

- For the 1990 PES, the Bureau plans to use an independent sample. In 1980, the Bureau used data from two CPS samples which were collected to provide statistical data about the U.S. labor force. Also, the 1990 sample will include 300,000 households, while, in 1980, the CPS samples and the reinterview sample included 150,000 and

110,000 households, respectively. The Bureau has estimated that the use of an independent sample and the increased sample size will increase the cost for the 1990 coverage measurement program.

- The Bureau plans to sample census blocks, and PES staff will then list and interview every household in the sampled blocks. The Bureau anticipates that block sampling will accelerate matching and reduce matching errors which should alleviate the delays similar to those experienced in the 1980 PEP processing.

PES Operations to be More Integrated with the Census

- To meet the dual goal of taking the most complete census while preparing for a possible adjustment, the Bureau plans to more fully integrate the PES and census operations. PES operations will occur concurrently with the census field enumeration and processing operations. The Bureau anticipates that integration will allow for earlier access to census materials and sufficient time for matching and PES followup which are essential to completing an adjustment by December 31, 1990.
- The Bureau plans to conduct PES interviewing in the summer of 1990, after the field offices complete the basic census field work. In 1980, interviewing for a planned PES would have been conducted in October through December 1980, according to the planned schedule. However, the Bureau dropped the 1980 PES because the results would not have been timely.
- The Bureau's goal will be to complete basic census field work in as many offices as possible by July 1990, so PES interviewing can begin. In 1980, the field offices were still enumerating people in late summer and early fall.

Automated Matching System

- The Bureau is developing an automated matching system to improve the speed, accuracy, and consistency of matching PES records to census records. In 1980, the clerical matching was time-consuming and error-prone. Although the Bureau plans to automate the matching process, some cases will still require clerical intervention. The computer will also assist in the clerical matching. The Bureau is concerned that it may be difficult to develop a large-scale automated matching system based on the 1985 and 1986 pretests.

- Automated matching will be dependent upon the Bureau's planned automation advances for the 1990 census, including 1) an automated address file, 2) early conversion of questionnaire data into computer-readable format, and 3) key-entry of the names for persons in the PES sample blocks and nearby blocks into computer files.
- The Bureau has established a professional matching staff to handle the most difficult cases which cannot be matched by the computer or clerks. To assure the accuracy and consistency of the clerical matching, this special matching group will also perform quality control reviews of the clerks' work.

Plans to Improve Data Quality

- For the 1990 PES, the Bureau plans to have a 3-week interviewing period, with an added week if problems arise. In 1980, the CPS interviewing was conducted during a 1-week period. The Bureau anticipates that the longer interview period will reduce the rates of noninterviews and missing data below those experienced in the 1980 PEP.
- During the initial PES interviews, the Bureau plans to obtain more complete information, asking questions that were included in the PEP followup in 1980. The Bureau anticipates that having this additional data will result in fewer followups.
- PES followup will be conducted during September and October 1990, which is only 2 or 3 months after the initial interviewing. In 1980, followup was conducted during January through April 1981, or 5 to 8 months after the August CPS sample and 9 to 12 months after the April CPS sample. The Bureau anticipates that the shorter time lapse between the PES interviewing and followup will reduce the rates of noninterviews and missing data below those experienced in the 1980 PEP.

Undocumented Alien Estimate

- The 1990 demographic analysis will use CPS data about the foreign-born population to estimate the number of undocumented aliens. Unlike 1980 estimates, the preliminary 1990 demographic estimates will incorporate data for undocumented aliens in the total population.

PES Testing Activities to Date1985 PES in Tampa

- A major objective of the 1985 PES in Tampa, Florida was to collect data to develop a computer and clerical matching system for use in later test censuses and in the 1990 Coverage Measurement Survey.
- In a report on the Tampa PES, the Bureau stated that the automated matching system was fast and accurate. According to the Bureau, the automated system greatly exceeded expectations for both match rate and accuracy.

1986 Rural PES

- A major objective of the 1986 PES in Mississippi was to gain experience with automated matching in rural areas and to test the feasibility of such matching.
- The Bureau concluded that the preliminary computer matching results for the rural PES indicate that automated rural matching is feasible. The Bureau further concluded that these interim results are encouraging and that the rural PES planned for 1987 will provide more experience in automated rural matching.

Test of Adjustment Related OperationsObjectives

- The objective of the 1986 Test of Adjustment Related Operations (TARO) in Los Angeles, California, was to test the feasibility of adjusting the census by December 31, 1986, to simulate the December 31, 1990, deadline. As a test of the operational feasibility of the Bureau's dual strategy, TARO was to collect data on the timing and costs of census adjustment operations.
- The 1986 TARO was the only test of adjusting census counts before the Bureau's scheduled 1987 decision on the feasibility of adjustment. But as of June 30, 1987, the Bureau had not made a decision.

Adjustment Not Timely

- The Bureau did not complete the TARO adjustment by its December 31, 1986, deadline. According to a Bureau official, the adjustment of the Los Angeles counts was not completed until March 20, 1987, or almost 3 months past the original deadline.
- Delays in searching the entire census file by computer to match the PES records adversely affected the availability of materials necessary for subsequent clerical matching and PES followup. Final results from the matching and followup were 2 weeks late. For 1990, the Bureau plans to limit the search area to only those blocks adjacent to the PES sample block.
- According to the Bureau, major delays also occurred in estimating the census errors and adjusting the census files. The Bureau attributed these delays, in part, to problems with untested software and keying errors in preparing a computer file of the matching results. Preparation of the PES results necessary for estimation was initially scheduled for 1 day, but subsequently took about 6 weeks to complete.

Adjusting for Undercount

- TARO produced the following estimates of the undercount rates: 1) 9.8 percent for Hispanics, 2) 7.3 percent for Asians, and 3) 6.2 percent for other races (non-Hispanic and non-Asian).
- TARO did not produce a separate undercount rate for Blacks. The Los Angeles PES site included only a small number of Blacks. Originally the Bureau planned to use two pretest sites. However, one site, which had the larger black population, was cancelled because the mail response rate was lower than expected and the resulting nonresponse followup workload would have caused the costs to exceed the allotted budget.
- The Bureau adjusted the census files for the undercount.

Bureau Concerns about the Timing of the Census and PES

- Partially on the basis of the TARO test, the Bureau has tentatively concluded that it will not be able to complete an adjustment by the December 31 deadline. On

the basis of examining the present planned schedules for census and PES operations, the Bureau recognizes that it cannot meet its dual goal to conduct a complete census and adjust the counts by December 31, 1990.

- Review of standards and ultimate decisionmaking will further lengthen the process of coverage measurement and census adjustment.
- Two major constraints on timing include the availability of 1) address and census data files needed to initiate matching operations and 2) edited census data needed to estimate census errors. According to the Bureau, the present time schedule for these items might not allow for adjusted counts by December 31, 1990.
- The Bureau is reviewing the current and alternative timing schedules to identify changes to the census operations which might enable adjusting the counts by the December deadline. The formal decision on the timing of major census and PES data collection and processing activities will be announced with the Bureau's decision on the operational and technical feasibility of adjustment.

Costs for the 1990 PES

- As of June 30, 1987, the Bureau had not provided its latest cost estimates for the 1990 PES because this matter is under consideration in the fiscal year 1989 budget cycle. However, as of August 1986, the Bureau's estimated cost of the large-scale 1990 PES ranged from about \$29 million to \$36 million.

(017005)



Requests for copies of GAO reports should be sent to:

U.S. General Accounting Office
Post Office Box 6015
Gaithersburg, Maryland 20877

Telephone 202-275-6241

The first five copies of each report are free. Additional copies are \$2.00 each.

There is a 25% discount on orders for 100 or more copies mailed to a single address.

Orders must be prepaid by cash or by check or money order made out to the Superintendent of Documents.

United States
General Accounting Office
Washington, D.C. 20548

First-Class Mail
Postage & Fees Paid
GAO
Permit No. G100

Official Business
Penalty for Private Use \$300

Address Correction Requested
