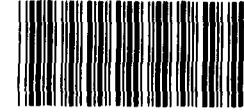


GAO

Testimony



144180

For Release
on Delivery
Expected at
1:30 p.m. EDT
Wednesday
June 19, 1991

1990 Census: Final Preparations for a
Possible Adjustment

Statement of
L. Nye Stevens, Director
Government Business Operations Issues
General Government Division

Before the
Subcommittee on Government Information and
Regulation
Committee on Governmental Affairs
United States Senate



051681 / 144180

1990 CENSUS: FINAL PREPARATIONS FOR A POSSIBLE ADJUSTMENT

SUMMARY OF STATEMENT BY
L. NYE STEVENS
DIRECTOR, GOVERNMENT BUSINESS
OPERATIONS ISSUES

The Census Bureau's recent release of 1990 census coverage estimates based on the Post Enumeration Survey (PES) and demographic analysis demonstrate that millions of persons were missed by the 1990 census. For 1990, the net undercount as estimated by the PES was about 2.1 percent, or approximately 5.3 million persons; and the net undercount based on demographic analysis was about 1.8 percent, or approximately 4.7 million persons. Moreover, on the basis of demographic analysis, the 1990 census was the first census not to reduce the net undercount over the preceding census and had the highest black/nonblack differential undercount since the Bureau began estimating coverage with the 1940 census.

The PES and its evaluations have been driven by the schedule necessary to meet the July 15 deadline for an adjustment decision. For the most part, the Bureau has been able to accommodate delays and problems in the PES and other census-related operations to stay on schedule. Bureau officials are now confident that should the Secretary decide to adjust census counts, the Bureau will be able to complete its activities by the deadline.

The major delay encountered by the Bureau toward the end of the PES was in "smoothing" the final PES adjustment factors and population estimates. Smoothing is a statistical technique that was used to reduce variation in the direct PES estimates of coverage error that occurred due to sampling. The Bureau delayed the announcement of the smoothed PES results for over 3 weeks--from the week of May 20 to the week of June 10--until it resolved concerns it had about its smoothing method.

The time schedule for an adjustment may have left insufficient time for analysis of the Bureau's 20 evaluations of PES. Although the Bureau was able to calculate the numbers and produce the data tables reporting the results of its evaluations, little time was available for analysis and interpretation of some of these results before they were presented to Bureau and Commerce Department decisionmakers.

At this point, GAO is unable to assess the quality of the PES, because it has not had time to assess the results of the Bureau's evaluations of the PES. The quality of PES data will influence the degree to which the Secretary can have confidence in the PES when deciding on adjustment. In the final analysis, Bureau and Commerce Department decisionmakers will need to use available data and their informed judgement when deciding upon the technical quality of the PES.

Mr. Chairman and Members of the Subcommittee:

We are pleased to be here today to discuss the status of the 1990 Post Enumeration Survey (PES)--a central methodology the Secretary of Commerce will use to decide whether or not to adjust census counts. The PES is a matching study in which the Bureau of the Census interviewed a sample of 165,000 households several months after the census. The results of these interviews were compared to census questionnaires to determine if each person was correctly counted, missed, or double-counted in the census.

In our previous reports and statements on the PES we noted the difficult task the Bureau has faced to complete a high-quality PES and its associated evaluations within an extremely tight schedule.¹ With the July 15 deadline for an adjustment decision from the Secretary of Commerce fast approaching, I would like to take this opportunity to build on that theme by discussing the final hurdles that the Bureau has encountered and other issues.

My comments are based on our ongoing work at the request of this Subcommittee and the House Subcommittee on Census and Population to monitor 1990 adjustment related matters. I must emphasize, however, that the Bureau is just beginning to provide us with data tables from the voluminous results from its 20 PES

¹See for example, 1990 Census Adjustment: Estimating Census Accuracy - A Complex Task (GAO/GGD-91-42, Mar. 1991) and Preparations for a Possible Census Adjustment (GAO/T-GGD-91-18, Mar. 19, 1991).

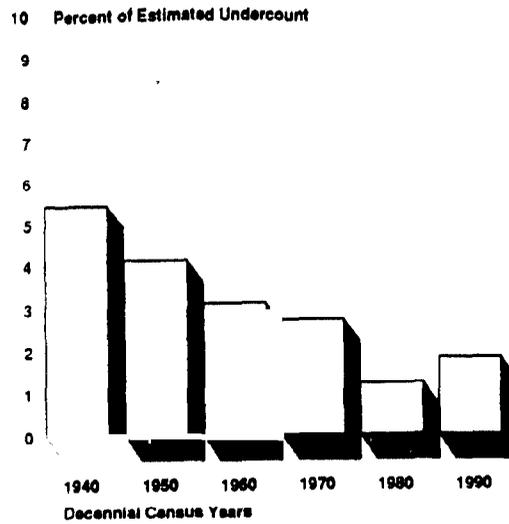
evaluation studies. As a result, we are not able today to assess in any detail the quality of the PES as indicated by the evaluations.

PES AND DEMOGRAPHIC ANALYSIS CONFIRM
SIZABLE ERRORS IN THE 1990 CENSUS COUNTS

The Bureau's release last week of 1990 census coverage estimates based on the PES and demographic analysis--an independent estimate of the population derived largely from administrative data such as birth and death records--demonstrated that millions of persons were missed by the 1990 census. For 1990, the net undercount as estimated by the PES was about 2.1 percent, or approximately 5.3 million persons; and the net undercount based on demographic analysis was about 1.8 percent, or approximately 4.7 million persons.

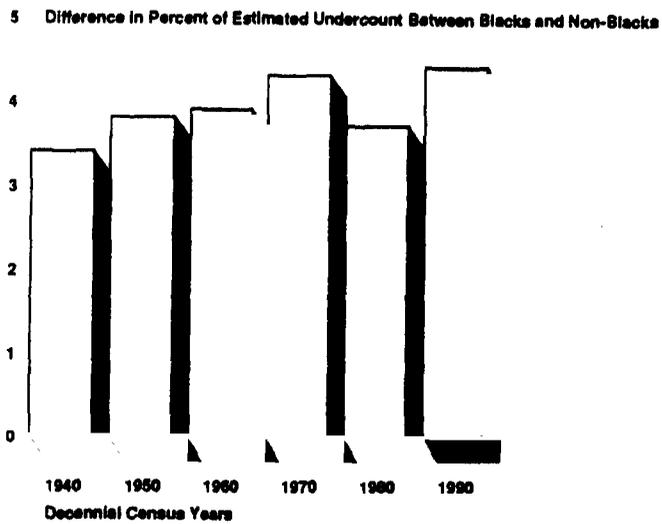
Demographic analysis is important because it provides both an independent estimate of the population and a consistent historical series of estimated undercounts for censuses between 1940 and 1990. For example, as shown in figure 1, on the basis of demographic analysis estimates of census coverage back to the 1940 census, the 1990 census was the first census not to reduce the net undercount over the preceding census. Moreover, as shown in figure 2, the 1990 census had the highest black/nonblack differential undercount since the Bureau began estimating coverage with the 1940 census.

Figure 1: 1990 Census Coverage Failed to Improve on Preceding Census



Source: Census Bureau estimates of the 1990 resident population based on demographic analysis as of May 1991.

Figure 2: 1990 Census Had Highest Differential Undercount Since Bureau Began Estimating Census Coverage



Source: Census Bureau estimates of the 1990 resident population based on demographic analysis as of May 1991.

The demographic analysis and PES data show only the estimated net error in the census. In reality, the gross number of errors in the census counts is much higher because some persons are missed by the census, while others are counted more than once.

Therefore, the net undercount data discussed above present only a partial picture of the amount of error in the census. As another measure of the quality of the 1990 census, we plan, once data are available, to compare the level of gross errors in the 1990 census as measured by the PES with the number of gross errors in the 1980 census.

We believe that the lack of improvement in reducing the undercount and its differential nature underscores the importance of the census reform effort that is now beginning. Regardless of the decision about adjustment the Secretary makes, we believe that the PES and demographic analysis results convincingly demonstrate the need for a more effective and efficient approach to taking the census.²

THE PES SCHEDULE: TIGHT TIME FRAMES
CONTINUE TO CHALLENGE BUREAU

The July 15 deadline for an adjustment decision required the Bureau to implement an extremely tight schedule for the PES and

²For additional information on the need for, and status of, census reform efforts, see Decennial Census: Preliminary 1990 Lessons Learned Indicate Need to Rethink Census Approach (GAO/T-GGD-90-18, Aug. 8, 1990); and Census Reform Needs Attention Now (GAO/T-GGD-91-13, Mar. 12, 1991).

its related evaluations. For the most part, the Bureau has been able to accommodate delays and problems in the PES and other census related operations to stay on schedule. But as the process came closer to the deadlines for getting results to decisionmakers at the Bureau and the Department, less cushion and flexibility were available. Bureau officials began looking for extra hours, rather than days, in the schedule to address last-minute problems. Bureau officials are now confident that should the Secretary decide to adjust census counts, the Bureau will be able to complete its activities by the July 15 deadline.

Production of "Smoothed" PES Results Delayed

The major delay encountered by the Bureau toward the end of the PES was in "smoothing" the final PES adjustment factors and population estimates. I want to stress, Mr. Chairman, that the Bureau's smoothing operation is an area that we have not had time to examine in depth. In general, smoothing is a statistical process used to reduce sampling variability. In the case of the PES, smoothing was used to reduce variation in the direct PES estimates of coverage error that occurred due to sampling.

These direct PES estimates for some population subgroups were based on relatively small numbers of sampled persons. The smoothing process, in essence, borrowed strength from other PES data to fit a regression model to predict a true undercount

value. By averaging the directly observed estimate of undercount and the model's predicted estimate of undercount, the final PES adjustment factors were theoretically made stronger, and some of the outlying observed values were moved to conform more closely to the general trend.

According to Bureau officials, a concern arose when the Bureau produced the smoothed PES results. The Bureau's calculation of variances for the smoothed and unsmoothed estimates should have been, but were not, consistent at regional levels. This concern delayed the announcement of the final smoothed numbers for over 3 weeks (from the week of May 20 to the week of June 10).

To its credit, even under the extreme time constraints the Bureau sought input to resolve concerns with smoothing from outside experts, such as the Special Advisory Panel that was formed to advise the Secretary on adjustment. However, since this is a crucial and sensitive process in the production of PES adjustment factors and population estimates, we believe that now that major PES technical activities have been completed, the Bureau should promptly release data on the model used and evaluations of its alternatives for review by the wider professional statistical community. By releasing information on the smoothing process, the Bureau would contribute to better public understanding of the challenge and complexity of adjusting the census, whichever decision the Secretary may make.

Time Constraints and PES Evaluations

In a larger sense, the time constraints faced in completing PES operations and evaluations in time for the July 15 deadline raise concerns about whether the quality of the PES data was jeopardized by the accelerated schedule. The results of the PES evaluation projects that are now becoming available should provide a general measure of PES data quality, although establishing any direct causal link between the rushed schedule and measurement errors will not be possible.

We believe the real loss due to an accelerated schedule was in the extent of analysis of the PES evaluations. As we stated in our March 19 testimony, careful and thorough evaluations are essential to measuring the amount of error in the PES and the degree to which the Secretary can have confidence in the results of the PES when making an adjustment decision. Under the tight time schedule, the Bureau was able to calculate the numbers and produce the tables reporting the results of its projects. However, little time was available for critically important analysis and interpretation of some of these results before they were presented to Bureau and Commerce decisionmakers.

INDICATORS OF THE QUALITY OF THE PES

I noted earlier in my statement, Mr. Chairman, that both the PES and the demographic estimates clearly show that millions of persons were missed by the census. The key question now is whether adjustment would bring the counts closer to the truth, especially at subnational geographic levels. The Bureau's evaluations of the PES, in particular its evaluation of overall error in the PES, provide important data that help answer that question. However, in the final analysis, a determination of the overall quality of the PES is not subject to absolute certainty and cannot be derived by producing the proper formula or mathematical equation. Bureau and Commerce Department decisionmakers will need to carefully consider the assumptions of the Bureau's proposed adjustment technique.

For some issues, the conclusions to be drawn from the data are much less ambiguous than for others. One question which has fairly clear answers is the effect of missing data in the PES. In our March testimony, we discussed the importance of minimizing the amount of missing data. We noted that in 1980 the high level of missing data and unresolved persons--those for whom the Bureau was unable to determine whether there was a match between the census and the PES--was an important factor in the recommendation of Bureau officials that 1980 post enumeration program estimates were not sufficiently accurate to adjust the census. We also

said in March that while the nationwide percentage of unresolved persons in the 1990 PES was relatively low compared to 1980, we were in the process of examining the detailed data on the distribution of these cases throughout the country.

Today, we can report that our analysis of missing data rates across various levels of aggregation such as census organization and geography and by the type of area (for example, central city) showed no major unexplained variations in these rates.³

Furthermore, the results of the Bureau's evaluation projects on missing data indicate that the model the Bureau used to impute missing data was very reasonable when compared to alternative models and simulations. The model used by the Bureau and 22 of the 23 alternative models it ran provided similar results.

Although the Bureau's efforts to address the missing data problem appear to have been successful, this does not necessarily mean that all concerns about the quality of the PES have been resolved and the census can be adjusted with confidence. Other sources of error affect the quality of the PES. Nevertheless, on a major source of error, the Bureau took the aggressive actions needed to minimize the problem.

³There were some district offices reporting high rates of unresolved cases where we traced the problem to a keying error in the computer files--cases that had been removed from the PES sample remained on the files and were coded as unresolved.

One area where the Bureau's efforts were not rewarded with similar success was in the measurement of correlation bias at subnational levels. As a result decisionmakers will need to use their judgement about the technical adequacy of the data. Correlation bias in the PES estimates results from persons being systematically missed by both the census and the PES. Since such persons are not counted by either the census or the PES, the PES estimates of population tend to be biased downward from the true population that would be calculated if all persons were counted. Estimating correlation bias is important because it is one of the major sources of error the Bureau examines as part of its assessment of overall error in the PES.

Estimating correlation bias at subnational levels has been a longstanding technical problem for the Bureau. For the 1990 PES, the Bureau planned to use two methods to estimate such bias. Both methods were relatively new and untested for purposes of measuring correlation bias in the PES. The Bureau was able to get only one of these two methods to produce usable data for purposes of its total error study. We believe that given the difficulty and uncertainty involved in measuring correlation bias, having two working methods would have helped decisionmakers assess the magnitude and effects of correlation bias in the PES. The Bureau's estimate of error due to correlation bias probably will be one of the most uncertain of all the error components in the PES.

- - - - -

In summary, Mr. Chairman, all measures of coverage error indicate that the 1990 census missed a greater percentage of the U.S. population than the 1980 census, the first time in modern census history that the coverage rate did not improve over the previous census. Furthermore, the differential undercount between the undercount of blacks and the undercount of nonblacks was greater than at any time since the Bureau began measuring the differential in 1940. However, at this point, we are unable to assess the quality of the PES. The quality of that data will influence the degree to which the Secretary can have confidence in the PES when deciding on adjustment.

This concludes my prepared statement. My colleagues and I will be pleased to answer any questions.

Copies of GAO testimonies and reports cited in this statement are available upon request. The first five copies of any GAO report or testimony are free. Additional copies are \$2 each. Orders should be sent to the following address, accompanied by a check or money order made out to the Superintendent of Documents, when necessary. Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

U.S. General Accounting Office
P.O. Box 6015
Gaithersburg, MD 20877

Orders may also be placed by calling (202) 275-6241.