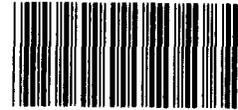


GAO

Testimony



144233

For Release
on Delivery
Expected at
9:30 a.m. EDT
Thursday
June 27, 1991

1990 Census: Applying PES Results and
Evaluations to the Adjustment Decision

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

Before the
Subcommittee on Census and Population
Committee on Post Office and Civil Service
House of Representatives



1990 CENSUS: APPLYING PES RESULTS AND EVALUATIONS
TO THE ADJUSTMENT DECISION

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

The Census Bureau's June 13 release of 1990 census coverage estimates based on the 1990 census Post Enumeration Survey (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 example, the net undercount as estimated by the PES was about 2.1 percent of the enumerated resident population of 248.7 million, or approximately 5.3 million persons. However, the scope and magnitude of the changes that would result from an adjustment are greater than simply adding the 5.3 million net undercount to the census because adjustment would correct both over- and undercounts.

Although the data are preliminary, various approaches to illustrating the number of gross errors in the census show that the 1990 census contained a substantial number of errors. Using a conservative approach that focuses just on double-counts and missed persons, GAO found the 1990 census contained about 14.1 million errors, including a gross census undercount of about 9.7 million persons. This is based on the Bureau's estimate that it double-counted about 4.4 million persons.

The Bureau's evaluation of the combined effect of sampling and nonsampling error in the PES found that the national undercount was between 1.23 percent and 2.20 percent at a 95 percent confidence level. Although this assessment of total error confirmed an undercount at the national level, a key question now is the amount of error in PES over- and undercount estimates at subnational levels.

The need for precision is especially important because the Bureau's procedure for carrying down PES adjustment factors to lower geographic levels makes the same adjustment to large numbers of people over wide geographic areas with similar demographic characteristics. The levels of sampling variation measured by the evaluations of the PES were generally much higher than anticipated by the original design of the PES. Because of this, GAO believes that the Secretary must closely examine whether the results of the PES are precise enough to make an adjustment at lower geographic levels given the amounts of sampling and nonsampling error identified by the PES evaluations.

Mr. Chairman and Members of the Subcommittee:

We are pleased to be here today to discuss the results of the 1990 census Post Enumeration Survey (PES)--a central methodology the Secretary of Commerce is using to decide whether or not to adjust the 1990 census counts. My comments today supplement my statement last week before the Senate Subcommittee on Government Information and Regulation and are based on our ongoing work to monitor 1990 adjustment-related matters.¹

PES AND DEMOGRAPHIC ANALYSIS DOCUMENT
SIGNIFICANT ERRORS IN THE 1990 CENSUS COUNTS

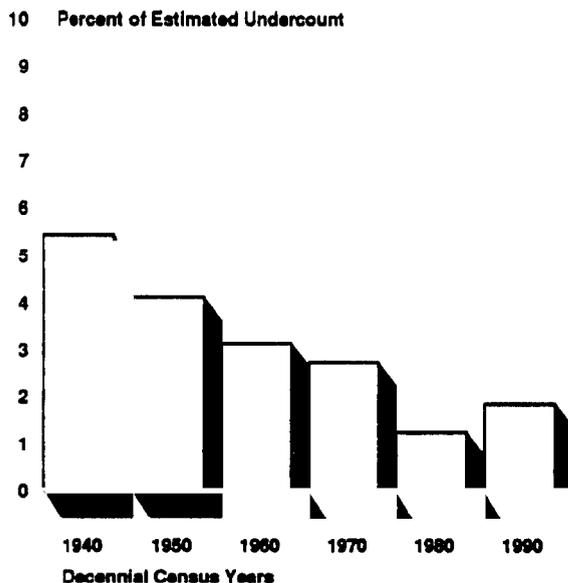
The Bureau's June 13 release of 1990 census coverage estimates based on the PES and demographic analysis--an independent estimate of the population derived largely from administrative data such as birth and death records--demonstrated that millions of persons were missed by the 1990 census. The net undercount as estimated by the PES was about 2.1 percent of the enumerated resident population of 248.7 million, 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

¹1990 Census: Final Preparations for a Possible Adjustment
(GAO/T-GGD-91-26, June 19, 1991).

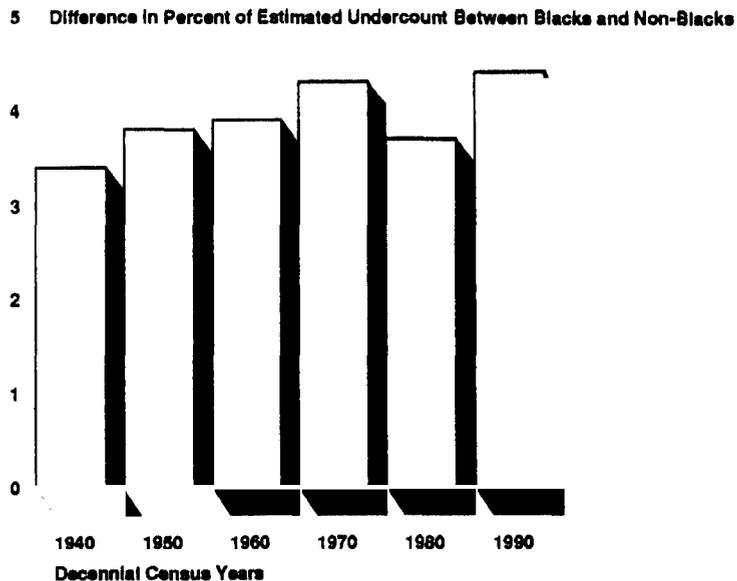
1940 and 1990. For example, as shown in figure 1, on the basis of demographic analysis estimates of census coverage back to the 1940 census, the 1990 census was the first census not to reduce the net undercount over the preceding census. Moreover, as shown in figure 2, the 1990 census had the highest 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.

Although demographic analysis is important for comparing net coverage among censuses, the PES performs a vital role in identifying the amount of double-counts and other census errors. This is important because should the Secretary of Commerce decide to adjust the census, counts would be corrected for both over- and undercounts. The scope and magnitude of the changes that would result from an adjustment are greater than simply adding to the count the 5.3 million persons included in the net undercount.

Although the data are preliminary, various approaches to illustrating the number of gross errors in the census show that

the 1990 census contained a substantial number of errors. Using the PES and the Bureau's independent estimate that it double-counted 4.4 million persons, and focusing just on double-counts and missed persons, and we estimate that the 1990 census contained about 14.1 million errors.² The gross census undercount using this method was about 9.7 million persons. Combining the number of double-counts with the undercount provides only a minimum indication of the number of gross errors because it does not include other types of census errors. For example, persons were placed at the wrong location or infants were incorrectly included in the census because they were born after April 1, 1990.

A more comprehensive approach that can be used to estimate gross census error includes additional census errors--such as persons counted in the wrong location--as well as undercounts and double-counts. Using this method, on the basis of preliminary PES data, there were about 25.9 million gross errors in the 1990 census, or about 10.4 percent of the resident population.

Comparisons between the 1990 census and the 1980 census must be made with caution due to differences in the two censuses' coverage evaluation programs and resulting data quality.

²The number of gross errors can be calculated as follows: gross errors = gross undercount (net undercount + erroneous inclusions) + erroneous inclusions. Differing definitions of "erroneous inclusion" can significantly affect the number of gross errors.

However, the amount of gross error in the 1990 census appears to be significantly greater than that in the 1980 census, using either the conservative or more comprehensive method for determining gross error. In both cases, despite problems with direct comparability of data, the 1990 census appears to have had at least 50 percent more errors than it had in 1980.

The lack of improvement in reducing the net undercount, its differential nature, and the amount of gross census error reinforce the importance of the census reform effort that is now beginning. We believe that the PES and demographic analysis results convincingly demonstrate the need for a more effective and efficient approach to taking the census, regardless of which decision on adjustment the Secretary makes.³

BUREAU EVALUATIONS OF PES FURTHER CONFIRM UNDERCOUNT
BUT CRITICAL QUESTIONS FOR ADJUSTMENT REMAIN

As you know, Mr. Chairman, we have long been concerned about the extremely tight time schedule for the PES.⁴ We believe the accelerated schedule limited the Bureau's ability to analyze the

³For 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).

⁴See, for example, Critical Issues for Census Adjustment: Completing the Post Enumeration Survey on Time While Protecting Data Quality (GAO/T-GGD-90-15, Jan. 30, 1990); and 1990 Census Adjustment: Estimating Census Accuracy--A Complex Task (GAO/GGD-91-42, Mar. 1991).

results of its 20 evaluations of the PES. As we said in our March 19 testimony before this Subcommittee and the Senate Subcommittee on Government Information and Regulation, 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 reliability and useability of the PES when making an adjustment decision.⁵

Under the tight time schedule, the Bureau was able to calculate the numbers and produce the tables reporting the results of its evaluations. 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. Careful analysis is important because a determination of the overall quality and usefulness of the PES is not subject to absolute certainty and cannot be derived by producing a single formula or mathematical equation.

A close examination of available data is important to determining the usefulness of the PES results for the purposes of making an adjustment, especially at lower geographic levels. For example, the Bureau's evaluation of the combined effect of sampling and nonsampling error in the PES found that the national undercount was between 1.23 percent and 2.20 percent at a 95 percent

⁵Preparations for a Possible Census Adjustment (GAO/T-GGD-91-18, Mar. 19, 1991).

confidence level. Although this assessment of total error confirmed an undercount at the national level, a key question now is the amount of error in PES over- and undercount estimates at subnational levels.

For example, we believe the amount of sampling error, or variability, deserves attention by the Secretary because it was a consistently high source of uncertainty in PES over- and undercount estimates. The PES estimates are based on samples and therefore are subject to random variation. The levels of sampling variation measured by the evaluations of the PES were generally much higher than anticipated by the original design of the PES. For example, even after smoothing to reduce sampling variability, PES over- and undercount estimates for 4 of the 13 evaluation groups did not show a statistically significant difference from the census count. In other words, due to the variability resulting from doing a sample, the Secretary cannot be sure whether 4 of the 13 population groups reviewed in the Bureau's evaluation of total error in the PES were overcounted by the census, undercounted, or if the census count was correct.

The need for precision is especially important because the Bureau's procedure for carrying down PES adjustment factors to lower geographic levels applies the same adjustment factors to large numbers of people over wide geographic areas with similar demographic characteristics. For example, the PES process

assumes that all black and nonblack Hispanic males aged 20 to 29, whether homeowners or renters, in central cities in Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, and Wyoming have the same net undercount rate and would be adjusted accordingly. We believe that the Secretary will need to closely examine whether the PES results are sufficiently precise to bring the census counts closer to the "true" population at all levels of geography.

- - - - -

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. 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. Finally, the number of gross errors in the 1990 census appear to be substantially greater than in 1980.

Nevertheless, in our view, the dependability of the PES as a tool for adjusting census counts remains open to question at this point. In the 3 weeks remaining before the deadline for an adjustment decision, the Secretary of Commerce must grapple with some difficult technical questions in deciding if adjustment

would improve the accuracy of the counts, particularly at lower geographical levels.

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

Copies of GAO reports and testimonies 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.