GAO

August 1991

1990 CENSUS

Reported Net Undercount Obscured Magnitude of Error

144674

GAO

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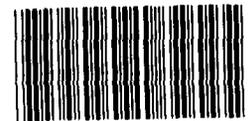
General Government Division

B-244895

August 22, 1991

The Honorable Thomas C. Sawyer, Chairman
The Honorable Thomas J. Ridge, Ranking Minority Member
Subcommittee on Census and Population
Committee on Post Office and Civil Service
House of Representatives

This report responds to your request that we provide additional information on the number of errors in the 1990 and 1980 censuses. We discussed census errors before the Subcommittee in our June 1991 testimony on the 1990 census Post Enumeration Survey (PES).[1] We testified in August 1990 how the challenges that confronted the Bureau of the Census in taking the 1990 census—such as escalating census costs, a declining level of public cooperation, and a shrinking workforce available for temporary census employment—may well become even more difficult to address in the future.[2] We believe that these challenges combined with the amount of error in the 1990 census underscore the need for a more effective and efficient approach to taking the census and reinforce the importance of census reform.

Results in Brief

We estimate that the 1990 census contained a minimum of 14.1 million gross errors and perhaps as many as 25.7 million errors, depending on how broadly census error is defined. In either case, these are substantially more errors than those indicated by the Bureau's widely reported 1990 census net undercount of about 5.3 million persons. A focus on the net undercount obscures the true magnitude of the error in the census because, while millions of persons were missed by the census, millions of other persons were improperly counted. Examining the amount of gross error, therefore, provides a more complete picture of the quality of the census.

The 14.1 million errors represent 5.7 percent of the census count of 248.7 million persons. The estimate is a minimum because it includes a gross estimate of 9.7 million missed persons and the approximately 4.4 million cases where PES results clearly identified persons as improperly

---

[1]1990 Census: Applying PES Results and Evaluations to the Adjustment Decision (GAO/T-GGD-91-49, June 27, 1991). The PES was a matching study in which the Bureau interviewed a sample of about 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.

[2]Decennial Census: Preliminary 1990 Lessons Learned Indicate Need to Rethink Census Approach (GAO/T-GGD-90-18, Aug. 8, 1990)

included in the census because they were either double-counted in their immediate geographic areas or were determined to have been fictitious. A more comprehensive definition of error that, in addition to the 4.4 million cases of improper counting, includes other types of census errors, such as placing persons in the wrong location, revealed that the census may have contained as many as 25.7 million errors, or about 10.3 percent of the count. Figure 1 illustrates that the census contained millions more errors than indicated by the net census error, whichever definition of error is used.

Figure 1: A Comparison of 1990 Census Gross Errors and Net Undercount Using Two Methods
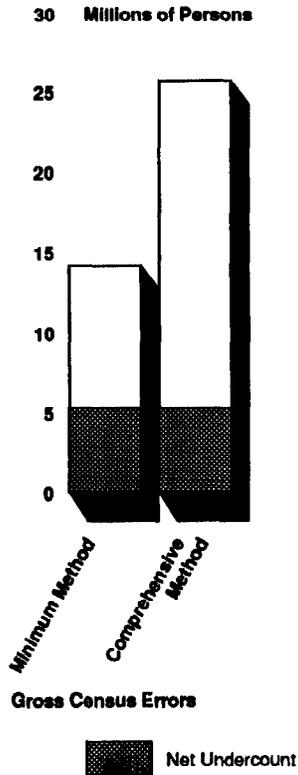
Source: GAO calculations based on Bureau of the Census data.

Finally, the 1990 census contained proportionally more errors than the 1980 census. The estimated minimum number of errors in the 1980 census (7.8 million) represented about 3.4 percent of the 1980 count in

contrast to 1990, when the minimum (14.1 million errors) represented about 5.7 percent of the count.

Objectives, Scope, and Methodology

Our objectives were to provide (1) a preliminary indication of the range of total errors in the 1990 census and (2) a general comparison of error levels in the 1990 and 1980 censuses.

To meet our first objective, we calculated the amount of error in the census by using the following equations:

- Net undercount = Estimated Population (from the PES or demographic analysis) − Census Count

- Gross Omissions = Estimated Population − Corrected Census Count (Census Count − Erroneous Inclusions)
  Erroneous inclusions are those persons improperly included in the census count. The census total is to some extent inflated because it contains double-counts and other errors. Subtracting these errors from the count provides a corrected census tally that can then be subtracted from the estimated population to provide the number of gross omissions.

- Gross Errors = Gross Omissions + Erroneous Inclusions
  The number of gross errors is determined by adding the total number of persons omitted to the total number included in error. Our estimates of gross error are expressed as a range because there is no universally accepted definition for erroneous inclusion. Different definitions of such inclusions can lead to significantly different estimates of gross error.

We based our calculations on preliminary PES data provided by the Bureau. We did not assess the accuracy of that data. The PES, although important to measuring error in the census, was a sample survey and, therefore, was subject to sampling and nonsampling error. However, the Bureau reported that the 1990 PES was of extremely high quality, especially at the national level. A Bureau evaluation of the amount of total error in the PES estimated the true net undercount as between 1.23 percent and 2.20 percent, within a 95-percent confidence interval. The specific PES point estimate calculated for the net undercount was 2.11 percent.

To meet our second objective, we examined attempts by the Bureau and independent researchers to determine the amount of total error in the 1980 census. We used 1980 census coverage data provided by the

Bureau. We did not assess the quality of these data. Only general comparisons between the 1980 and 1990 censuses are possible because of differences in the scope and quality of the two coverage evaluation programs.

The Bureau believes that the 1990 PES was of much higher quality and, therefore, more successful in detecting errors in the census than its 1980 matching study, called the post enumeration program (PEP). For example, unlike the PES, the 1980 PEP suffered from large amounts of missing data. As a result, the 1980 PEP did not provide definitive estimates for the net undercount or erroneous inclusions in the 1980 census. The estimate of the net undercount in the 1980 census is, therefore, based on demographic analysis—an independent population estimate derived primarily from administrative records such as births, deaths, and immigration.

We did our work in June and July 1991 in accordance with generally accepted government auditing standards.

Number of 1990 Census Errors Significant

For the 1990 census, both the PES and the Bureau's demographic analysis showed a net census undercount. The net undercount as estimated by the PES was about 2.1 percent of the resident census count of 248.7 million persons, or approximately 5.3 million persons. Based on demographic analysis, the net undercount was about 1.8 percent, or approximately 4.7 million persons.

However, the census contained millions of errors in addition to the reported net undercount. Moreover, a focus on the net undercount masks the extent and variety of these other census errors. Such errors, which are referred to as erroneous inclusions, encompass persons who were counted more than once, nonexistent persons included because census enumerators falsified data, persons included in the census but assigned to the wrong location, and other errors. Assessing the magnitude of these errors in addition to the net undercount provides a more complete indication of the quality of a census.

We estimate that, at a minimum, the number of gross errors in the 1990 census was about 14.1 million persons, or 5.7 percent of the census count. Our estimate is based on preliminary PES data that indicate that approximately 4.4 million persons were either double-counted in their

census block or in a nearby block or were fictitious.[3] In essence, such errors inflate the number of persons reported as being correctly counted in the census. Removing these 4.4 million errors from the census count for purposes of analysis shows that the actual number of persons missed was at least 9.7 million persons, rather than the reported net undercount of 5.3 million persons. However, this approach to defining erroneous inclusions provides only a minimal indication of the number of gross errors because it does not include other types of census errors.

A more comprehensive approach to defining census errors—an approach that considers other sources of census errors—revealed that there were about 10.2 million erroneous inclusions in the 1990 census. Table 1 compares the numbers of errors using the minimum and more comprehensive approaches.

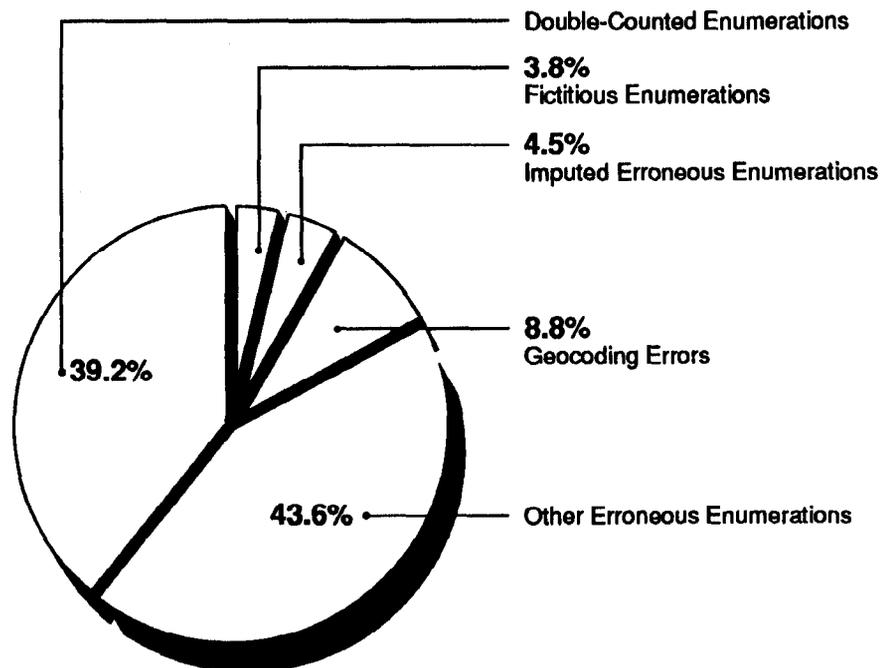**Table 1: 1990 Gross Census Errors Using Two Methods** (In millions of persons)

Method	Net undercount	Erroneous inclusions	Omissions	Gross errors	Gross errors as percent of census
Minimum	5.3	4.4	9.7	14.1	5.7%
Comprehensive	5.3	10.2	15.5	25.7	10.3%

Source: GAO calculations based on Bureau of the Census data.

The comprehensive method for calculating erroneous inclusions encompasses not only persons who were double-counted or determined to have been fictitious but also other sources of error such as persons placed at the wrong geographic location (e.g., those persons improperly counted at a vacation home rather than at their primary residence). Counting persons in their correct locations is important because, in addition to reapportioning the House of Representatives, census data are used to redraw congressional, state, and municipal legislative district lines and help determine the distribution of billions of dollars of federal funds. Figure 2 shows double-counts, fictitious enumerations, and three other types of erroneous inclusions as a percentage of all erroneous inclusions in the 1990 census.

---

[3]This estimate includes about 4 million double-counts and approximately 400,000 fictitious persons. The data presented here, although still based on preliminary PES results, are more current and complete than the data presented in our June testimony, in which we estimated that the number of double-counts alone could be 4.4 million.

Figure 2: 1990 Census Erroneous Inclusions as a Percentage of All Erroneous Inclusions

Double-Counted Enumerations

3.8%
Fictitious Enumerations

4.5%
Imputed Erroneous Enumerations

8.8%
Geocoding Errors

39.2%

43.6% — Other Erroneous Enumerations

Note 1: Percentages do not add up to 100 due to rounding.

Note 2: The figure excludes approximately 3 million cases, representing about 1.2 percent of the census count, that were unmatchable. Unmatchable cases are those where the census data were so incomplete (for example, missing names) that the Bureau could not attempt to match these records with the PES.

Source: GAO calculations based on Bureau of the Census 1990 PES, P-9 Evaluation (5/31/91).

In addition to double-counts and fictitious enumerations, census errors covered by the categories presented in figure 2 include the following:

- Imputed erroneous enumerations: cases in which the Bureau statistically assigned an erroneous enumeration status. Some persons, for whom the Bureau was not able to determine if they were counted in the census, were classified as unresolved. The Bureau statistically assigned a PES/ census match determination for unresolved cases. This statistical procedure, called imputation, is based on the results of similar cases where the Bureau was able to make a match determination. According to the Bureau, about 80 percent of the unresolved cases were imputed as correctly enumerated and 20 percent as erroneously enumerated.
- Geocoding errors: persons counted in housing units that were recorded by the Bureau in the wrong geographic location.

- Other erroneous enumerations: persons incorrectly included because they died before the census, were born after the census, or should have been counted elsewhere on the basis of census residency rules.

Number of Gross Errors in 1990 Significantly Greater Than in 1980

The 1990 census contained millions more errors than the 1980 census. Table 2 shows that, using the minimal definition of census error, the 1980 census contained at least 7.8 million errors in a count of about 226.5 million persons.

Table 2: 1980 Gross Census Errors Using Minimum Method (In millions of persons)

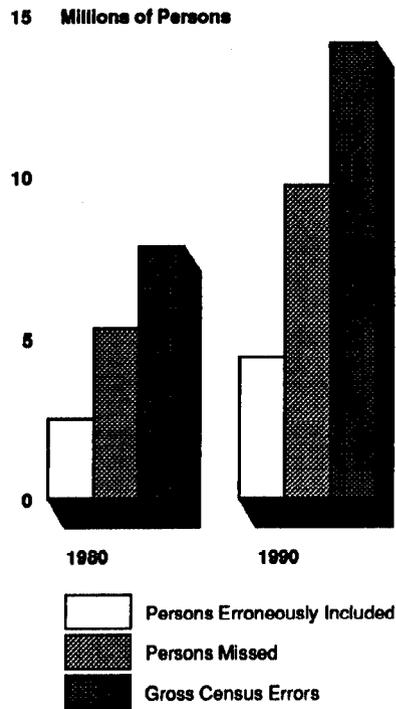| Net undercount	Erroneous inclusions	Omissions	Gross errors	Gross errors as percent of census
2.8ª	2.5ᵇ	5.3	7.8	3.4%

ªThe current demographic estimate for the 1980 net undercount is used because the 1980 PEP did not provide a single estimate of the net undercount. The Bureau continually changes its estimate of the undercount of prior censuses as more and better data become available.

ᵇThis figure includes only those persons double-counted in their immediate areas. It does not include fictitious enumerations because the PEP did not provide a reliable estimate of such errors in the 1980 census.

Source: GAO calculations based on Bureau of the Census data.

Comparisons between censuses must be made with caution because of differences in coverage evaluation programs. However, despite limitations in the data, the 1990 census appears to have contained many more errors than the 1980 census. Figure 3 compares the number of errors in the 1980 and 1990 censuses using the minimal definition of error. The best available data are used to estimate net undercount—for 1980, the demographic estimate as of May 1991 and for 1990, the PES. Significant differences in the structure of the 1980 PEP and the 1990 PES preclude comparison using a more comprehensive definition of census error.

Figure 3: Minimum Number of Errors in 1980 and 1990 Censuses

15   **Millions of Persons**

10

5

0

1980   1990

☐ Persons Erroneously Included

▨ Persons Missed

■ Gross Census Errors

Source: GAO calculations based on Bureau of the Census data.

Overall, the minimal percentage of gross error in the 1990 census was 5.7 percent of the census count. In contrast, the minimal percentage of gross error in 1980 was 3.4 percent of the count. If the 1980 census had an error rate similar to the 1990 census, it would have contained over 5 million additional errors.

Agency Comments

We discussed our findings with responsible Bureau of the Census officials. They generally agreed with the facts and analyses presented and suggested several technical clarifications. We incorporated the clarifications where appropriate.

As arranged with the Subcommittee, we are also sending copies of this report to the Subcommittee on Government Information and Regulation, Senate Committee on Governmental Affairs; other appropriate congressional committees; the Secretary of Commerce; the Director of the Bureau of the Census; and the Director of the Office of Management and

Budget. Copies also will be made available to other interested parties upon request.

The major contributors to this report are listed in the appendix. If you have any questions concerning this report, please contact me on (202) 275-8676.

L. Nye Stevens
Director, Government Business
  Operations Issues

Major Contributors to This Report

General Government Division, Washington, D.C.

William M. Hunt, Assistant Director, Government Information and Statistical Policy Issues
J. Christopher Mihm, Evaluator-In-Charge
Jack Kaufman, Adviser
Timothy A. Bober, Evaluator
Maria Z. Oliver, Evaluator

Ordering Information

The first five copies of each GAO report 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.

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

Official Business
Penalty for Private Use $300