

GAO

Testimony

For Release
on Delivery
Expected at
10:00 a.m. EST
Thursday,
February 21,
1991

Components of the 1990 Census Count

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



050631 / 143203

COMPONENTS OF THE 1990 CENSUS COUNT

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

The 1990 census population count came from three broad sources: (1) data that individuals and households provided on themselves by mailing back a questionnaire or directly to a census enumerator; (2) data gathered from nonhousehold sources, such as administrative records or neighbors; and (3) data generated through statistical procedures such as imputation.

At this point, data are not available that show clearly how much each of the three sources contributed to the census count. Nevertheless, it is clear that as in previous censuses, the vast majority of census population data in 1990 was provided by households themselves. For the resident population of 248.7 million in 1990, preliminary data indicate that about 166.9 million persons were enumerated through questionnaires returned to the Census Bureau by mail. About 10.9 million persons were enumerated in areas where the mail-back method was not used. An additional 58.8 million persons were counted during census follow-up efforts, mostly as part of households that provided data on themselves to a census enumerator, although some persons were counted on the basis of data provided by others, such as neighbors. The balance of the population, about 12 million persons, was enumerated in group quarters, such as nursing homes and homeless shelters or through coverage improvement efforts. (The components do not add to 248.7 due to rounding.)

Two census programs--the overseas count and the parolee and probationer coverage improvement program, which together added about 1.4 million persons to the census count used for apportionment, relied heavily on administrative records as a basic tool to take the census. In addition, although final data are not available, the Bureau may have gathered census data from surrogates for about 7.3 million persons.

After census field efforts were completed, the Bureau could not determine the number of persons who lived in some housing units, or, in other cases, if the units were occupied. The Bureau uses completed census questionnaires from neighboring units to statistically impute results for uncertain units. In 1990, the Bureau imputed about 54,000 persons, many fewer than the 762,000 imputed in 1980.

An evaluation of the comparative quality of directly gathered, nonhousehold sources, and statistically generated data can provide insight into the best mix of methodologies to improve the cost effectiveness of future censuses. For example, the Bureau should explore if and how sampling can be used to improve the census.

Mr. Chairman and Members of the Subcommittee:

I am pleased to be here today to discuss the primary components of the 1990 census count. The 1990 census population count can be viewed in the broadest sense as deriving from three distinct methodologies: (1) data that individuals and households provide on themselves, for example by mailing back a questionnaire or providing the information to a census enumerator; (2) data gathered from nonhousehold sources, such as administrative records or neighbors and landlords; and (3) data that were generated through statistical procedures such as imputation. The Bureau's use of nonhousehold sources and imputation to help build the census population count in 1990 was consistent with the Bureau's practice in previous censuses.

My comments are based on our ongoing evaluation--which we are doing at the request of this Subcommittee and the Senate Governmental Affairs Committee--of the primary census programs that contributed to the count. Unfortunately, my comments today are tentative because the Bureau has thus far not developed complete information on what the various census programs contributed to the count. Much of the data the Bureau has developed are preliminary and are from census field office management information systems or unedited census files. Final state and local level data for some programs are not expected to be available for at least a year. Only when more definitive data

become available will it be possible to evaluate the relative merits of the Bureau's various coverage improvement programs.

Although data are not available that show clearly how much each of the three methodologies contributed to the census count, it is clear that as in previous censuses the vast majority of 1990 census population data was provided by households themselves. For the resident population of about 248.7 million in 1990,¹ preliminary data indicated that about 166.9 million persons, or about 67 percent of the resident population, were enumerated through questionnaires returned by mail to the Bureau. About 10.9 million persons, or about 4.4 percent of the count, in the most rural areas of the nation returned their questionnaires to enumerators. An additional 6.6 million persons, 2.7 percent of the count, were enumerated by the Bureau in group quarters such as nursing homes and homeless shelters.

Approximately 58.8 million additional persons, representing 23.6 percent of the resident population, were counted during census follow-up efforts, mostly as part of households that provided

¹The resident population includes all persons enumerated as residing in the United States. As we discuss on pages 3 through 4, the Bureau included about 923,000 American citizens residing overseas into the counts used for apportioning seats in the House of Representatives. The total enumerated population for apportionment purposes was about 249.6 million. These population counts are subject to possible correction for undercount or overcount. The Department of Commerce is considering whether to correct these counts and will publish corrected counts, if any, not later than July 15, 1991.

data on themselves to census enumerators, although some were counted by using information provided by surrogate sources such as neighbors and landlords. The Bureau's preliminary data show that the remaining resident population--about 5.4 million persons, or about 2.2 percent of the count--was included in the census count as a result of the Bureau's post follow-up coverage improvement efforts. (The components do not add to 248.7 million due to rounding.)

Although the 1990 census was consistent with previous censuses in that most data was collected from households that provided information on themselves, 1990 also was consistent with the past in that a percentage of census population data came from nonhousehold sources and statistical procedures. While the percentage of such data is small, it can have a definite impact on census results as was seen in 1980; I will discuss this impact shortly.

CENSUS POPULATION DATA FROM NONHOUSEHOLD SOURCES

One of the major methodological changes the Bureau made for the 1990 census was the increased use of administrative records as part of the census enumeration for reapportionment. The Bureau's procedures for enumerating group quarters allowed enumerators to use administrative records in certain specified situations. However, two programs--the overseas count and the Parolee and

Probationer Coverage Improvement Program which together added about 1.4 million persons to the census counts used for reapportionment, relied heavily on administrative records as a basic tool to take the census.

The Overseas Count

Selected groups of Americans living abroad have been counted in the census since 1900. However, the 1990 census was only the second time that the Bureau included the overseas count in the census figures used to reapportion seats in the House of Representatives. For the 1970 census, during the Vietnam War, the Bureau attempted to enumerate all American citizens living overseas and include them in apportionment counts.

For 1990, the Bureau decided to again include the overseas census count in the state population totals used for reapportionment. However, due to the administrative complexities in enumerating Americans residing abroad independently, the Bureau decided to include only military and federal employees and their dependents in the 1990 overseas count. Contrary to 1970, when the Bureau required American citizens living overseas to complete a census questionnaire, the Bureau in 1990 relied on federal agencies to provide it with census counts.

About 98 percent of the approximately 923,000 persons included in the overseas count were enumerated by the Department of Defense (DOD). DOD used personnel records to enumerate military personnel and their dependents. Civilian personnel were asked to complete and return a DOD questionnaire to DOD. DOD tabulated these questionnaires and supplied the Bureau with the results of the combined military and DOD civilian count. The Bureau did not independently confirm either the accuracy of the count that DOD and other agencies provided or the administrative records used to generate the count. Since the Bureau required that each agency certify the accuracy of its count, the Bureau believed an independent check of the counts was not necessary.

Parolee and Probationer Coverage Improvement

A second program where the Bureau used administrative records to assist with the enumeration was the 1990 Parolee and Probationer Coverage Improvement Program. The Bureau decided late in the census planning cycle to implement a special enumeration effort for parolees and probationers--of which the Bureau estimated there to be more than 2.6 million--to reduce the census undercount, particularly for minorities. As a result of the late development of the program, the Bureau was unable to test the program during the 1988 census dress rehearsal.

We are assessing the parolee and probationer program as part of the Subcommittee's request that we examine the primary components of the census count. At this point, we can report that the program contained two phases with distinctly different methodologies. For the first phase, the Bureau, in cooperation with the states, asked parole and probation officers to distribute a special census questionnaire to each individual in their client caseload. The parolees and probationers were asked to complete the form and return it to their parole or probation officer for return to the Bureau. The Bureau attempted to match each individual with the census questionnaire for his or her address or the questionnaire of nearby addresses. If the records did not match, the parolee or probationer was assumed to have been missed by the census and was added to the count.

According to the Bureau, about 384,000 parolees or probationers, or only about 15 percent of the estimated parolee and probationer population, responded. Because of the low response rate, the Bureau decided in August 1990 to shift the reporting burden away from parolees and probationers. For this second phase of the program, the Bureau targeted certain primarily urban areas with high expected concentrations of parolees and probationers and asked parole and probation officials to certify that the addresses on their client lists were accurate as of Census Day, April 1, 1990.

If the officials could certify the accuracy of the addresses, the Bureau attempted to match the individuals with the census forms from their reported addresses or surrounding addresses. In cases where the administrative record and the census form did not match, the individual was added to the census without a field follow-up to determine the accuracy of the administrative record. In cases where the officials could not certify the accuracy of the listed addresses, Bureau field staff attempted a telephone follow-up if a telephone number was available. Data on the number of cases that received such a follow-up were not available.

Overall, although the data are preliminary, the Bureau reported that the program identified over 1.3 million parolees and probationers and added about 448,000 persons to the count as a result of both phases of the program. Precise information on the relative contribution of each phase will not be available, even when the total number of additions to the count is final, because the Bureau did not distinguish between the two phases when it processed the program's questionnaires.

The use of administrative records for the parolee and probationer program can be compared with the way records were used as part of the Nonhousehold Sources Program in 1980. Like the parolee and probationer program, the Nonhousehold Sources Program was a coverage improvement program that was aimed at reducing the

disproportionate undercount of minorities. In 1980, the Bureau obtained lists from the Immigration and Naturalization Service and state and local government sources, such as lists of issued driver's licenses for certain areas. The lists contained names, basic demographic information, and the most likely permanent address for each person. Names on the lists were matched with the census questionnaire from the address indicated on the list.² The Bureau did a follow-up on persons who did not match to see if they should be included in the census count.

The 1980 nonhousehold sources program follow-up effort was a marked contrast to the 1990 parolee and probationer program, where unmatched persons were added to the census without follow-up. According to Bureau officials, parolees and probationers who did not match in 1990 did not receive follow-up because parolee and probation officers had certified that the reported address was accurate as of Census Day--which was not the case with the lists used in 1980. However, in the cases where the 1990 census and the lists did not match, the accuracy of the lists remained

²The Bureau did not use the nonhousehold sources program in the 1990 census. In 1980, the Bureau anticipated that about 10 percent of the persons found on the lists would be added to the census count; however, only 1.9 percent of persons on the lists, or about 127,000 persons, were added to the count. The Bureau reported that it spent about \$77 per person added to the count from the nonhousehold sources program. For more information on the nonhousehold sources program, see 1990 Census: Comparison of Coverage Improvement Programs for 1980-1990 (GAO/GGD-90-8, Nov. 1989, pp. 14-15.)

uncertain because the Bureau did not consistently attempt to confirm the information it received.

Census Data from Surrogate Sources

A second source of nonhousehold census data is from surrogates, such as neighbors, mail carriers, and building managers. In our testimony before this Subcommittee in July 1990, we noted that the Bureau instructed its enumerators to collect census data from surrogate sources in cases where the enumerators were repeatedly unable to locate a resident of the housing unit.³ We also noted that such "last resort" data introduces a potential source of error into the census.

Although not all last resort data come from surrogates, the Bureau's still preliminary and incomplete data indicate that it collected last resort data on about 3.2 percent of the nation's occupied housing units. Data on the number of persons enumerated through last resort procedures is not yet available from the Bureau. However, a rough calculation is possible. According to the Bureau, about 2.5 persons reside in the average household. Thus, if the average size of a household enumerated through last resort procedures is the same as that of all households, about

³Status of 1990 Census Follow-up and Evaluation Efforts (GAO/T-GGD-90-58, July 23, 1990.)

7.3 million persons may have been enumerated through last resort procedures.

As we reported in July, the amount of last resort data varied significantly among the Bureau's district offices, with urban offices generally having the highest rates. On the basis of the Bureau's more recent data, 51 district offices, all of them in urban areas, collected last resort data on occupied housing units at more than twice the national average, and 14 offices collected last resort data on 10 percent or more of all of the occupied housing units in their areas.

The amount and distribution of last resort data from the 1980 census are not available. As a result, it is not possible to compare the Bureau's 1990 performance in gathering complete information from household members with its 1980 performance.

STATISTICALLY GENERATED POPULATION DATA

After census field data collection efforts are completed, there is always a residual number of housing units where the Bureau was unable to determine the number of persons who lived in the unit or, in other cases, if the unit was even occupied. In such situations, the Bureau imputes population counts. Imputation, an accepted practice in survey research, is the assignment of information for unreported items on a questionnaire. The Bureau

has applied statistical procedures that use completed census questionnaires from neighboring units to complete the questionnaire for the unit where the Bureau could not gather data.

In 1970, the census count included about 4.9 million persons who were added on the basis of various statistical procedures, including imputations. For example, the 1970 count included about 900,000 imputed persons from housing units that the Bureau found to be occupied but could not determine a population count. Additional persons were included in the census count through a variety of other statistical procedures, such as sampling. For example, the 1970 National Vacancy Check--a recheck of housing units listed as vacant during census follow-up efforts--was done on a sample basis. The Bureau rechecked a sample of 13,546 vacant units and concluded that nationwide, about 11.4 percent of the housing units originally classified as vacant were really occupied.⁴ As a result, about 1.1 million persons were added to the count.

Sampling also was used to add to the 1970 count after a Postal Service records check of all housing units enumerated in the rural portions of 16 southern states--the area the Bureau

⁴In 1980, when the Bureau revisited all housing units originally identified as vacant, about 10 percent of the units were reclassified as occupied. In 1990 the conversion rate was about 7.6 percent.

believed was subject to the highest undercount. About 500,000 persons were added to the census count as a result of this sampling procedure. Additional persons were included in the census from other statistical procedures.

Due to concerns about the legality of sampling, the Bureau did not use sampling techniques as part of the 1980 census but did impute about 762,000 persons into the census count. However, even this small percentage of the total census count, about 0.3 percent, had an important influence on census results. The number of imputed persons was not evenly distributed across the nation and resulted in the shifting of a congressional seat from Indiana to Florida.

For 1990, the Bureau strived to reduce the number of imputations in part due to the controversy from the shifting of the congressional seat due to imputation in 1980. As a result, the preliminary data show that about 54,000 imputed persons were included in the count. If this number holds, imputation will not cause a congressional seat to shift between states as it did in 1980.

CONSIDERATIONS FOR PLANNING THE 2000 CENSUS

I have discussed how the Bureau employs a variety of methodologies--data from households, data from nonhousehold

sources including administrative records and surrogates, and data that are statistically generated--to enumerate a dynamic population. We believe the experiences from 1990 point to the following possible opportunities to reform the census-taking process:

- The increased use of administrative records in 1990 to help develop the census count represents a new approach to census-taking on the part of the Bureau. The use of administrative records provides an example of the Bureau's willingness to rely on other organizations to directly assist with the basic census data collection. We urge the Bureau to systematically explore how and when it can legally and effectively develop shared responsibilities for census data collection and use administrative records in future censuses. The Bureau also should assess its responsibilities for ensuring the accuracy and completeness of data it receives from others. Of course, any benefits from the use of administrative records in taking the census must be balanced against concerns about protecting individual rights to privacy.

- The Bureau plans to study the quality of imputed data in the 1990 census. We believe an important part of this study should be a thorough comparison of the quality of imputed data with the quality of data from neighbors and other

sources, particularly the cases where those sources were unable to provide even basic demographic data. Such an analysis could assist the Bureau in determining when to impute data and help resolve concerns about whether imputation is appropriate. The Post Enumeration Survey, a primary vehicle the Bureau is using to assess the accuracy of the census, may provide data to assist in such an evaluation.

-- More broadly, we believe that the Bureau should rigorously assess the extent to which sampling procedures can and should be used to assist with basic census data collection.

- - - - -

In closing Mr. Chairman, we believe that experiences from the 1990 census indicate a number of areas where the Bureau could make potentially significant improvements to the census-taking process. However, work is needed now, while we have the opportunity to make changes in the basic census design.

This concludes my prepared statement, Mr. Chairman. My colleagues and I would be pleased to respond to questions.