

CRS Report for Congress

Received through the CRS Web

Census 2000: The Sampling Debate

Jennifer D. Williams
Analyst in American National Government
Government Division

Summary

Plans by the Bureau of the Census to incorporate data from two new sample surveys into the 2000 decennial census count have had a mixed congressional reception. Three sampling bills in the 105th Congress (H.R. 1220, H.R. 1178, and H.R. 776) have been referred to committee, without further action. Sampling has been debated chiefly in the appropriations process (H.R. 1469, H.R. 1871 [P.L. 105-18; 111 Stat. 158], and H.R. 2267/S. 1022 [P.L. 105-119; 111 Stat. 2440]). The bureau now is a defendant in two anti-sampling suits brought under P.L. 105-119. The law also established a Census Monitoring Board and directed the bureau to prepare for a traditional headcount in 2000, not just to continue with its sampling plans. This report will be updated to reflect further legislative action. For a related CRS product, see: *Census 2000: Sampling as an Appropriations Issue in the 105th Congress*.

Background

In 2000, the Bureau of the Census, Department of Commerce, will conduct the 22nd consecutive decennial enumeration of the U.S. population. The count is a constitutional requirement for reapportioning the House of Representatives. Census data also are used for within-state legislative redistricting, allocating certain federal funds to states and localities, administering various federal programs, and performing many kinds of research.

Two problems that underlie ongoing congressional concern about the bureau's preparations for Census 2000 are coverage (whom the census counts accurately and whom it misses or counts incorrectly) and costs. Of equal concern, however, are the bureau's plans to deal with these problems by incorporating the results of two new sample surveys into the 2000 enumeration.

Although the 1990 census counted about 98.4% of the total population, this census was the first since 1940 in which coverage did not improve over that in the previous enumeration. Further, the recurrent problem of the "differential undercount" (in which the census inadvertently omits higher proportions of minority groups than of the majority population) persisted in 1990. Estimates indicate that the 1990 census missed 4.4% of blacks, 5.0% of Hispanics, 2.3% of Asians and Pacific Islanders, and 4.5% of American

Indians, compared with 0.7% of non-Hispanic whites. The difference between the black and nonblack estimated undercounts was the largest since 1940.¹

Despite repeated legislative and judicial attempts to require statistical “adjustment” (alteration based on sample survey data) of the 1980 and 1990 census figures, the bureau never officially has made such an adjustment.² The stated purpose of these efforts was to improve census accuracy. While maintaining that adjustment would not have met this goal unequivocally for the past two censuses, the bureau has continued working to develop adequate methods for coverage improvement and cost control in Census 2000.

The 1990 census cost a record-high \$2.6 billion. In constant 1990 dollars, the cost of enumeration rose from about \$11.00 per housing unit in 1970 to \$20.00 in 1980 and \$25.00 in 1990. The 1970 through 1990 censuses used essentially the same approach: mailed-out, mailed-back questionnaires, with multiple followup visits by enumerators to households (the occupants of housing units, one household per unit) whose members did not return their questionnaires. The percentage of questionnaires filled out and returned from all occupied housing units dropped from 87% in 1970 to 83% in 1980 and 74% in 1990.³ Declining public cooperation with enumeration has meant greater expenditures for nonresponse followup, estimated at \$490 million to \$560 million for the 1990 census. The bureau has projected that a 2000 census incorporating the results of two proposed sample surveys might cost almost \$4.0 billion, but a census without these surveys might cost an additional \$675 million to \$800 million.⁴

Congress, through P.L. 102-135; 105 Stat. 635, directed the Secretary of Commerce to contract with the National Academy of Sciences (NAS) to study alternative census methods and their costs. One resulting NAS report noted that “sampling for nonresponse followup could produce major cost savings in 2000.” The report also endorsed sampling at a second point in the census process, stating that the “differential undercount cannot be reduced to acceptable levels at acceptable costs without the use of integrated coverage measurement and the statistical methods associated with it.”⁵ (“Integrated coverage measurement” is the bureau’s term for its proposed correction of coverage errors in the 2000 census.)

¹ U.S. General Accounting Office, *Decennial Census: 1990 Results Show Need for Fundamental Reform*, GAO report GAO/GGD-92-94 (Washington: June 9, 1992), pp. 19-21; U.S. Bureau of the Census, *Report to Congress: The Plan for Census 2000* (Washington: July 1997, reissued Aug. 1997), p. 4.

² For a history of the adjustment controversy, see: U.S. Library of Congress, Congressional Research Service, *Decennial Census Coverage: The Adjustment Issue*, by Jennifer D. Williams, CRS Report 94-89 GOV (Washington: July 23, 1996).

³ U.S. General Accounting Office, *Decennial Census: 1990 Results Show Need for Fundamental Reform*, pp. 23-24, 36.

⁴ National Research Council, Committee on National Statistics, Panel to Evaluate Alternative Census Methods, *Counting People in the Information Age* (Washington: National Academy Press, 1994), pp. 97-98; U.S. Bureau of the Census, *Report to Congress*, pp. 37, 39.

⁵ National Research Council, *Counting People in the Information Age*, p. 206.

On February 28, 1996, the bureau unveiled its Census 2000 strategy,⁶ which reaffirms NAS's recommendations for sampling. The bureau's original sampling plans, with later modifications, and the congressional response to these plans are discussed below.

The Sampling Plans

Nonresponse Followup. The bureau analyzed various options for nonresponse followup. An option highlighted on February 28, 1996, and in later congressional testimony, stated that the bureau would interview by telephone as many nonresponding households as possible, then would resort to the most expensive type of followup: personal visits to the remaining unenumerated housing units. These visits were to continue until the bureau obtained responses for 90% of the housing unit addresses in a given *county*. At 90%, the bureau would complete nonresponse followup for the county by surveying a random sample of housing unit addresses for which questionnaire responses were lacking. Enumerators would visit these addresses and obtain information that the bureau would use to estimate data for the other nonresponding households.⁷

On September 12, 1996, the bureau announced that, before starting to sample, it would attain a 90% response rate at the *census tract* level, "to provide more equitable and accurate representation of...neighborhoods."⁸ (A census tract is a statistical subdivision of a county. The typical tract has 2,500 to 8,000 persons with fairly similar economic characteristics.)⁹ On March 11, 1997, the bureau announced a further modification. It will not reach a 90% response rate in every census tract before sampling the remaining 10%. Instead, it has chosen an option called "direct sampling,"¹⁰ in which random sampling will begin immediately, or "directly," after the mail-back/phone-in phase¹¹ of the census ends. At this point, the bureau will calculate the response rate for each tract and use the rate to determine what fraction of nonresponding households to sample. For example, if a tract has a 60% (.60) mail-back/phone-in response rate and a 40% (.40) nonresponse rate, the bureau will sample three-out-of-four (75%, or .75) of the nonresponding households (.75 times .40) to obtain a 90% (.90) total response rate for the tract: .75 times .40 equals .30;

⁶ U.S. Bureau of the Census, *The Plan for Census 2000* (Washington: Feb. 1996, reissued April 1996).

⁷ *Ibid.*, pp. III-34, III-36; U.S. Congress, House Committee on Government Reform and Oversight, *Census 2000: The Challenge of the Count*, hearing, 104th Cong., 2nd sess., June 6, 1996 (Washington: GPO, 1996), p. 27.

⁸ U.S. Bureau of the Census, "Sampling and Estimation Update," distributed at the 2000 Census Advisory Committee Meeting (Washington: Sept. 26, 1996).

⁹ U.S. Bureau of the Census, *1990 Census of Population and Housing, Guide, Part B, Glossary*, 1990 CPH-R-1B (Washington: GPO, 1993), p. 10.

¹⁰ U.S. Bureau of the Census, "Commerce Department's Census Bureau Releases Refinements to Plans for Census 2000," press release CB97-C.01 (Washington: March 11, 1997), p. 1; U.S. Bureau of the Census, *Report to Congress*, p. 28.

¹¹ The bureau plans to mail two waves of census questionnaires; a mailed notice or reminder will precede each wave. Questionnaires also will be available in multiple public locations, such as shopping malls, community centers, and schools, for persons who did not receive or who misplaced their questionnaires. Anyone who would rather answer the questionnaire by telephone than by mail will be able to do so by calling a toll-free number. (U.S. Bureau of the Census, *Report to Congress*, pp. 21-22.)

.30 plus .60 equals .90. Alternatively, if the mail-back/phone-in response rate for a tract is 80%, with a 20% nonresponse rate, the sample will comprise one-in-two (.50 times .20) nonresponding households: .50 times .20 equals .10; .10 plus .80 equals .90. By this means, the bureau still expects to have contact (through mailed-back/phoned-in responses and enumerators' interviews for nonresponse followup) with 90% of all households in every census tract. Using data from the sample, the bureau will estimate information for the nonresponding households not sampled.

Integrated Coverage Measurement. The bureau will conduct a second sample survey, for integrated coverage measurement (ICM).¹² The ICM survey will allow the bureau to estimate the extent to which the initial Census 2000 results (mailed-back/phoned-in responses and enumerators' interviews) correctly included, missed, or erroneously counted housing units and persons. By applying these estimates to the initial census figures, the bureau expects to correct coverage errors.

“Dual system estimation” (DSE), which the bureau used to evaluate the quality of the 1980 and 1990 censuses, is the method it has chosen to produce estimates for ICM. The DSE method involves comparison of two independent data sets for the same geographic areas. For ICM, the geographic areas will be a sample of blocks (the smallest units of census geography). One independent data set will be persons in the ICM sample; the other will be persons included in the initial phases of the census. Comparing these groups will enable the bureau to determine, for the sample blocks, the proportion of housing units and persons included or missed in each group. The bureau will resolve discrepancies between the initial census results and responses to the original ICM interviews, usually through followup ICM interviews, to arrive at the correct information.

The ICM sample will consist of about 750,000 housing unit addresses in 25,000 blocks. To derive the sample, the bureau first will classify the nation's seven million blocks in groups called “strata,” then will select blocks at random from every stratum. The strata will be based on the similar characteristics of blocks, as determined in the 1990 census, and will not cross state boundaries. One stratum in a state, for example, might comprise every block in large central cities with a 1990 census population of at least 30% black renters and at least 10% Hispanic renters. The bureau will design strata and “poststrata” (discussed below) both to be representative of state populations and to correct coverage errors for hard-to-count groups and areas. Examples of some groups counted less effectively than others in the 1990 census are blacks and Hispanics, American Indians on reservations, renters (particularly in rural areas), and children under age 18.

After completing the ICM interviews, the bureau will sort the initial 2000 census population into poststrata, or groups of persons with similar chances of having been counted in the initial census phases. The poststrata will classify persons according to state geographic subdivision (such as rural or urban), housing unit owner or renter status, sex, age, race, and Hispanic ethnicity. DSE “estimation factors,” applied to the initial census results, will provide estimates of the total number of housing units and persons in each poststratum. Summing across poststrata will produce estimates of state totals, and the sum of state totals will yield national totals. The bureau will use the state-level estimation factors by poststrata to obtain housing unit and population estimates for all blocks in the United States.

¹² *Ibid.*, pp. 29-32.

The Congressional Response

Congress's interest in better census coverage and its determination to control spiraling census costs are matched by congressional reservations about sampling for nonresponse followup and for ICM. These issues have carried over from the 104th Congress to the 105th. Especially in the House, but in the Senate as well, sampling tends to have Democratic support and Republican opposition. The press has attributed this difference to the perception that sampling will benefit Democrats in House reapportionment and within-state legislative redistricting after the 2000 census.¹³

On September 18, 1996, the bureau's House oversight committee, Government Reform and Oversight, adopted a nonbinding report that stated: "The Bureau should not use sampling methods to complete or adjust the actual enumeration" in 2000.¹⁴ The report questioned the objectivity of the sampling design, the constitutionality of sampling and its possibly negative effect on public confidence in the census, the quality of small-area data obtained partly from sampling, and the operational feasibility of completing two complex sample surveys by the statutory deadline for producing official census numbers. In dissenting opinions, 18 committee members asserted that the report presented no counterproposal to correct persistent census coverage errors and control census costs. The report, with dissenting views, set the terms for further debate.

Thus far in the 105th Congress, three sampling bills have been referred to the House Government Reform and Oversight Committee, without further action. (1) H.R. 1220 (Petri), introduced March 21, 1997, would prohibit using sampling or other statistical procedures to determine state population for House reapportionment. Mr. Petri introduced H.R. 3589, with the same prohibition, in the 104th Congress. (2) H.R. 1178 (Maloney), introduced March 20, 1997, would amend Title 13, *U.S. Code*, the census statute, to clarify that sampling may be used to improve decennial census accuracy. (3) H.R. 776 (Meek), introduced February 13, 1997, would require the bureau to contact, by mail or in person, 90% of the households in a census tract before sampling the other 10% for nonresponse followup. Mrs. Meek introduced H.R. 3558 in the 104th Congress. (As already discussed, the bureau plans to arrive at this 90% response rate by direct sampling.)

The main forum for the sampling debate in the 105th Congress has been the appropriations process,¹⁵ beginning with FY1997 supplemental appropriations legislation for disaster relief. President Clinton vetoed the first such bill approved by Congress, H.R. 1469, in part because it contained language to ban the use of sampling in any census to

¹³ "Census Funds Dispute," *Washington Post*, July 11, 1997, sec. A, p. 3; John Mercurio, "Appropriators Freeze Census Spending on Sampling," *Roll Call*, July 14, 1997, pp. 10, 32; Juliana Gruenwald, "Statistics Stir the Passions of Parties' Boosters," *Congressional Quarterly Weekly Report*, June 21, 1997, pp. 1434-1436.

¹⁴ U.S. Congress, House Committee on Government Reform and Oversight, *Sampling and Statistical Adjustment in the Decennial Census: Fundamental Flaws*, 104th Cong., 2nd sess., H. Rept. 104-821 (Washington: GPO, 1996), p. 17.

¹⁵ For more information about the appropriations controversy, see: U.S. Library of Congress, Congressional Research Service, *Census 2000: Sampling as an Appropriations Issue in the 105th Congress*, by Jennifer D. Williams, CRS Report 98-321GOV (Washington: March 31, 1998).

determine the reapportionment population.¹⁶ A second bill passed by Congress, H.R. 1871, stipulated only that, within 30 days of enactment, the Commerce Department give Congress a detailed report¹⁷ on its proposed methods for conducting the 2000 census, including estimates of sampling errors. The President signed this legislation on June 12, 1997 (P.L. 105-18; 111 Stat. 158).

Congress then resumed the sampling debate in FY1998 appropriations bills for Commerce, Justice, and State, the Judiciary, and Related Agencies, H.R. 2267 and S. 1022. As enacted (P.L. 105-119; 111 Stat. 2440), section 209 of this legislation provided for expedited judicial review of the bureau's sampling plans, to determine their legality and constitutionality. The bureau currently is a defendant in two suits brought under this section.¹⁸ Both suits seek to prevent the use of sampling in the census to derive the reapportionment population. Section 209 further instructed the bureau to prepare for a traditional headcount in 2000, not simply to continue with its sampling plans. This section also stated that in Census 2000 and "any dress rehearsal or simulation" in preparation for the census, "the number of persons enumerated without using statistical methods must be publicly available for all levels of census geography...." The data subject to this provision include the official state populations used for reapportionment and redistricting. The provision counters the bureau's proposed "one-number census," that is, a single set of official census reapportionment numbers (for release by the December 31, 2000, legal deadline), which would incorporate sampling for nonresponse followup and would be altered on the basis of ICM. Finally, Section 210 of the FY1998 appropriations legislation established a Census Monitoring Board to oversee all aspects of the coming enumeration and report periodically to Congress. The eight-member board (with two members appointed by the Senate majority leader, two by the Speaker of the House, and four by the President) will go out of existence on September 30, 2001.

The bureau's present operating conditions include simultaneous preparation for two types of census (one with, and one without, sampling); heightened scrutiny from the Census Monitoring Board and a new House Government and Reform census subcommittee, created on November 13, 1997;¹⁹ and the lack of a permanent director, after Martha Farnsworth Riche's resignation, effective January 31, 1998. James Holmes, head of the bureau's Atlanta regional office, was named acting bureau director on January 30, 1998.²⁰

¹⁶ U.S. Congress, Conference Committee, 1997, *Making Emergency Supplemental Appropriations for Recovery from Natural Disasters, and for Overseas Peacekeeping Efforts*, conference report to accompany H.R. 1469, H.Rept. 105-119, 105th Cong., 1st sess. (Washington: GPO, 1997), pp. 66-67.

¹⁷ U.S. Bureau of the Census, *Report to Congress*.

¹⁸ *Glavin v. Clinton*, no. 98-207-A (D.VA filed Feb. 12, 1998); *U.S. House of Representatives v. U.S. Department of Commerce*, no. 1:98CV00456 (D.DC filed Feb. 20, 1998).

¹⁹ *Congressional Record*, daily edition, vol. 143, Nov. 13, 1997, pp. H10790-H10793.

²⁰ U.S. Dept. of Commerce, "Commerce Secretary Daley Names Acting Census Bureau Director," press release issued Jan. 30, 1998, <http://www.census.gov/Press-Release/com98-01.html>, visited Feb. 23, 1998.