



GAO

Accountability * Integrity * Reliability

United States General Accounting Office
Washington, DC 20548

January 5, 2001

The Honorable Henry A. Waxman
The Honorable Carolyn B. Maloney
House of Representatives

Subject: Decennial Censuses: Historical Data on Enumerator Productivity Are Limited

This letter responds to your request for data on enumerator productivity levels from the 1940 through the 2000 Censuses. In your respective capacities as the Ranking Minority Member, Committee on Government Reform, and the Ranking Minority Member, Subcommittee on the Census, you asked us to develop this information to better clarify the relationship between the Bureau of the Census' field data collection workload, and the time and labor force needed to complete it. These factors—used to calculate productivity—are some of the largest drivers of census costs, and the Bureau developed its budget for the 2000 Census using a model that contained key assumptions about expected workload and enumerator productivity.¹

We reviewed historical data to obtain information for the 1940 through 1990 Censuses; and, for the 2000 Census, we examined preliminary data contained in the Bureau's Cost and Progress database, a management information system the Bureau uses to track the status of the census. We also interviewed officials from the Bureau's 2000 Decennial Management Division and history office to obtain information on (1) how the data were developed, (2) their limitations, and (3) the extent of their comparability.

On December 11, 2000, we requested comments on a draft of this letter from the Secretary of Commerce. However, comments were not provided in time to be included in this letter. We performed our work between June and December 2000, in Washington, D.C., and Bureau headquarters in Suitland, MD, in accordance with generally accepted government auditing standards.

Results in Brief

We could not calculate productivity levels for the 2000 Census and most of the earlier censuses included in our review because data were largely unavailable, incomplete, or not comparable. With respect to enumerator productivity for the 2000 Census,

¹ 2000 Census: Analysis of Fiscal Year 2000 Amended Budget Request (GAO/GGD-99-291, Sept. 22, 1999).

Bureau officials said that they could not assure the reliability of this data until planned reliability assessments are completed in early 2001.

Moreover, definitional differences in how the Bureau counted the number of enumerators who worked on the census, and variations in census-taking methodologies, limited the comparability of productivity data from one census to the next. Given the importance of productivity information for validating assumptions the Bureau used to develop its budget for the 2000 Census, and the role that this information can play in planning for the 2010 Census, we recommend that the Secretary of Commerce ensure that the Bureau refines available productivity data as planned; identify the extent and nature of any anomalies, as well as the impact they have on data quality; and determine the extent to which the data can be compared by local census office type (rural versus urban).

Background

The decennial census is the nation's largest and most expensive data gathering program. For the 2000 Census, the Bureau estimated that it would need to collect information on more than 270 million residents living in about 119 million housing units, at a cost of at least \$6.8 billion. Although the majority of census questionnaires were to be returned by mail, the Bureau hired over half a million enumerators to follow up on the more than 42 million households that did not respond.

To develop its budget for the 2000 Census, the Bureau used a cost model that consisted of an extensive set of interrelated software spreadsheets. The model was built on cost data and workload history from the 1990 Census, and research conducted for the 2000 Census, to develop cost estimates of census operations for 2000.

According to the Bureau, in addition to being the foundation for the budget process, the 2000 Census cost model also provided planning and scheduling justifications within the Bureau and to external oversight groups, and was used to conduct "what-if" studies for alternative census scenarios. The 2000 cost model contains numerous formulas and assumptions developed by program managers that were generally based on either third-party evidence, such as independent studies, or senior management's judgment.

Two key assumptions in the cost model were expected workload and enumerator productivity. Past experience has shown that these variables are two of the largest drivers of census costs and in fact, as we noted in our September 1999 report,² the Bureau's request for an additional \$1.7 billion for fiscal year 2000 census operations was based largely on changes in assumptions related to increased workload and reduced employee productivity.

² GAO/GGD-99-291.

Historical Data on Enumerator Productivity Are Limited

For the 2000 Census, we could not calculate actual enumerator productivity levels because Bureau officials told us that they have not assessed, and therefore cannot assure, the reliability of productivity data for any of its field operations. For example, the Bureau is still refining data on the number of hours enumerators worked on collecting census information from nonresponding households. According to Bureau officials, a preliminary examination indicates that in some local census offices, a significant number of individuals on enumerator applicant lists were hired instead as crew leader assistants—a different position. In some instances, the position change was not reflected in the Bureau's personnel/payroll system. To the extent this occurred, Bureau officials said that it would overstate the number of hours that enumerators actually worked. Bureau officials also said that they are uncertain about the extent to which enumerators working on more than one census operation charged an operational code other than nonresponse follow-up. They noted, however, that while a certain level of mischarging occurs with any census operation, there is no evidence that this took place to a greater degree in the 2000 Census compared to past censuses or other Bureau surveys.

Bureau officials estimate that it will take them several months to review the Bureau's productivity data for potential problems and complete the data for release. They told us that to determine whether particular time charges are correct or in need of an adjustment, the Bureau will review productivity information from several computer databases, as well as from time and motion studies the Bureau conducted during nonresponse follow-up at selected sites.

According to Bureau officials, even when the Bureau completes its review, the productivity data will only be comparable at the national and regional census office levels. Bureau officials said that the data would allow for comparisons to be made among the Bureau's four types of local census offices—which differ by enumeration methods used and geographic makeup (urban versus rural)—but not among individual local census offices. The officials said that they expect to have reliable data on enumerator productivity available in the January to March 2001 time frame, at which time we plan to evaluate the data.

Some data on enumerator productivity levels, based on actual nonresponse follow-up cases completed per hour, were available for the 1990 and 1980 Censuses. According to data reported by the Bureau, enumerators completed 1.56 cases per hour for the 1990 Census nonresponse follow-up operation. During the 1980 Census nonresponse follow-up effort, enumerators completed 1.09 cases per hour at "centralized" offices and 1.26 cases per hour at "decentralized" offices. (Centralized offices were located in urban areas, and decentralized offices were located primarily in suburban and rural areas.) Combined national-level data were unavailable for the 1980 Census. As a result, productivity information for the two censuses are not comparable.

With respect to the earlier censuses (1940 through 1970), we could not calculate enumerator productivity because needed data on staffing levels and hours worked were unavailable, incomplete, or not comparable. For example, the Census of

Agriculture was conducted as part of the 1940 and 1950 Censuses, and available data include enumerators who worked on both operations—thus potentially overstating the number of enumerators working on the census of population and housing.

Moreover, available data on staffing levels did not count enumerators in the same way from census to census. For the 1950 and 1960 Censuses, staff figures are based on total enumerator positions planned for, but not necessarily filled, while staffing data from later censuses count the total number of enumerators actually hired. Further, for the 1940 through 1970 Censuses, available historical data lacked information on the number of hours enumerators worked. According to Bureau officials, this was because enumerators were paid on a piece-rate as opposed to an hourly basis and, as a result, the Bureau did not collect data on the number of hours enumerators worked.

More generally, the comparability of enumerator productivity data from one census to the next would be limited because of changes in the way the census was conducted. Most significantly, until the 1970 Census, enumerators visited each housing unit to obtain census information. Starting in 1970, questionnaires were mailed out to designated segments of the population for completion and mailed back, and the Bureau's field enumeration efforts largely consisted of following up with those housing units that did not mail back their questionnaires.

Conclusion

Data on enumerator productivity will be important for evaluating the validity of some of the key assumptions used in the Bureau's 2000 Census cost model, as well as for informing the Bureau's planning and budgeting processes for the 2010 Census. As previously noted, the Bureau cannot yet assure the reliability of data on staffing levels and the number of hours enumerators worked on the census. Because these two variables are used to calculate productivity, it will be important for the Bureau to assess their reliability as currently planned.

In addition, because of differences in the way the Bureau measured staffing levels and hours worked from census to census, such data could not be compared from one census to the next. As a result, the extent to which enumerator productivity had increased or decreased over time could not be accurately determined. As the Bureau plans for the 2010 Census, it will be important for the Bureau to develop a consistent set of key enumerator productivity indicators that could be used to compare enumerator productivity data from the 2000 Census and any future censuses that use the same basic census-taking method as was used in 2000.

Recommendations for Executive Action

To help ensure that the Bureau has the enumerator productivity data it needs to develop accurate budget, planning, and scheduling justifications for the 2010 Census, we recommend that the Secretary of Commerce ensure that the Bureau refines available productivity data from the 2000 Census as planned. At a minimum, the

Bureau should identify the extent and nature of any anomalies, the impact they have on data quality, and the extent to which the data can be compared by type of local census office. Where necessary, the Bureau should adjust the data accordingly and describe any actions taken in this regard. Moreover, to help ensure the comparability of data for the 2000 and future censuses, the Bureau should fully document how its enumerator productivity rates are calculated and report productivity data by type of local census office.

- - - - -

We are sending copies of this letter to Representative Dan Burton, Chairman of the Committee on Government Reform; Representative Dan Miller; the Honorable Norman Y. Mineta, Secretary of Commerce; and the Honorable Kenneth Prewitt, Director of the Bureau of the Census. Copies will be made available to others upon request. Robert Goldenkoff and Victoria E. Miller made major contributions to this letter. If you have any questions concerning this letter, please contact me at (202) 512-6806.

A handwritten signature in black ink, appearing to read "J. Christopher Mihm". The signature is written in a cursive style with a large initial "J" and a long horizontal stroke.

J. Christopher Mihm
Director
Strategic Issues

(410589)