



UNITED STATES DEPARTMENT OF COMMERCE
The Inspector General
Washington, D.C. 20230

October 25, 2012

The Honorable Tom Coburn
United States Senate
172 Russell Senate Office Bldg.
Washington, DC 20510

Dear Senator Coburn:

This letter responds to your requests during the July 18, 2012, hearing—*Census: Planning Ahead for 2020*—before the U.S. Senate Committee on Homeland Security and Governmental Affairs, Subcommittee on Federal Financial Management, Government Information, Federal Services, and International Security. You specifically requested that we answer two questions: (1) What can Congress do to help the Census Bureau make greater use of administrative records? and (2) How can Congress ease the Census Bureau's restrictions on sharing address information with state and local governments?

Data collected by federal agencies for program purposes (typically referred to as *administrative records*) contain personally identifiable information that, when combined with other records, can produce statistical information such as demographic, labor force, and socioeconomic indicators. Administrative records have the potential to decrease data collection costs, reduce respondent burden, and produce additional data that is not collected by questionnaires. However, as we discussed in our initial response, accumulating person-level information from multiple sources must be balanced against privacy and confidentiality concerns.¹ As a solution, we suggest that Congress consider playing a role in (1) compelling other federal agencies to provide their data to the Census Bureau and (2) easing the Bureau's restrictions on sharing address information with state and local governments.

(1) What can Congress do to help the Census Bureau make greater use of administrative records? Currently the Census Bureau has the legal authority to request data from other federal agencies for statistical activities. However, those agencies are not required to provide their administrative records to the Bureau. Below are some federal records that the Bureau could use (those marked with an asterisk are records collected and maintained at multiple locations, such as across states):

¹ OIG and the U.S. General Accounting Office (GAO) concur on this point; see *Record Linkage and Privacy: Issues in Creating New Federal Research and Statistical Information* (GAO-01-126SP, April 2001).



U.S. Department of Agriculture		
Food and Nutrition Service-----		Nutrition assistance program records*
U.S. Department of Education		
National Center for Education Statistics-----		School records*
U.S. Department of Health and Human Services		
Administration for Children and Families-----		Family assistance program records*
Centers for Medicare & Medicaid Services-----		Medicaid records
Centers for Medicare & Medicaid Services-----		Medicare records
National Center for Health Statistics-----		Birth records*
National Center for Health Statistics-----		Death records*
U.S. Department of Housing and Urban Development		
Public and Indian Housing-----		Subsidized housing records
Public and Indian Housing-----		Public housing records
U.S. Department of Labor		
Employment and Training Administration-----		Unemployment insurance records*
U.S. Department of Treasury		
Internal Revenue Service-----		Information and tax return records
U.S. Department of Veterans Affairs-----		Veterans' benefits records
Selective Service System-----		Draft registration records
Social Security Administration-----		Social Security records

This list is not an exhaustive accounting of all potential record systems. The Census Bureau may be aware of additional records or may discount the potential use of some of the record systems listed above for data quality or access reasons. We have provided these examples to illustrate the vast amount of information that is collected, but not easily accessed, for statistical purposes.

To obtain records, the Census Bureau currently must enter into negotiations with the agency record holder because each data set presents its own legal constraints on data access and use. To resolve these constraints requires an often intensive, time-consuming process. For example, last decade the Bureau's Longitudinal Employer-Household Dynamics Program (LEHD) requested W-2 information from the Internal Revenue Service (IRS) for its linkage to survey records. Negotiations between the two agencies failed when IRS could not reconcile its tax regulations with LEHD's statistical objectives for using tax data. As a result, LEHD instead developed a partnership with state agencies to acquire comparable data from the unemployment insurance program. Incrementally, LEHD accumulated nationwide statistical coverage—after a decade of negotiations—to produce estimates of local employment dynamics. One former Bureau policy chief, in a peer-reviewed research journal, cited this example of how privacy and confidentiality conflicts among agencies can impede the effective use of

administrative records as the basis for national statistical databases.² Such barriers prevent federal agencies from establishing an agile and robust statistical system that meets the quickly evolving demands of policy makers, businesses, and the general public.

In addition, a number of federal laws affect data sharing, as well as the protection and use of identifiable administrative data.³ These laws, among others, can range from encouraging the use of administrative records to enhance statistics and reduce respondent burden to limiting use and requiring stringent confidentiality requirements. The laws can be broad, covering all personal information collected by all agencies, or specific to one agency or set of information, such as Census Bureau data or IRS tax records. For example, according to Bureau officials, a recent request for Department of Education information raised questions as to whether that department can provide the information due to the privacy requirements mandated by the Family Educational Rights and Privacy Act of 1974.⁴

To reduce this burden and enhance statistical operations, Congress could consider a legislative requirement for federal agencies to provide unclassified information to the Census Bureau for the sole purpose of conducting statistical operations. Providing the Bureau greater access to administrative records, while mitigating privacy and confidentiality concerns, could result in significant cost savings. For example, using administrative records to reduce the costs of the decennial nonresponse operations (\$1.6 billion for the 2010 Census) is worth exploring for the 2020 decennial.

(2) How can Congress ease the Census Bureau’s restrictions on sharing address information with state and local governments? Although tribal, state, county, and local governments share address information with the Census Bureau, Title 13 forbids the Bureau from reciprocating with those partners—except for a few very narrow exceptions, such as the once-a-decade address-updating program. The Census Address List Improvement Act of 1994 did amend Title 13 to provide a limited exception to these restrictions for local governments, allowing two-way sharing of address information between the Bureau and officials designated by local government units. However, the program has several requirements, including an appeals process through which address-sharing disputes among the bureau and participants may be resolved.

The 2010 decennial’s program occurred from January 2007 through March 2010; although it may require legislative action, a more informal method of two-way sharing of address lists earlier in the decade could improve address updating and mapping as well as the cost-effectiveness of the Census Bureau’s demographic censuses and surveys throughout the decade. Address and map information is already widely available on the Internet and through data vendors. Thus, extremely restrictive statutory requirements regarding the protection and use of addresses have become antiquated. These laws mire individual federal agencies in inefficiency

² Gates, G. W. 2011. How Uncertainty About Privacy and Confidentiality Is Hampering Efforts to More Effectively Use Administrative Records in Producing U.S. National Statistics. *Journal of Privacy and Confidentiality*. 3(2): 3–40.

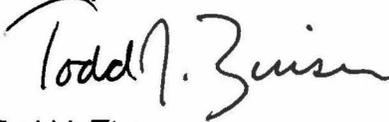
³ Pertinent laws include the Privacy Act of 1974, the Confidential Information Protection and Statistical Efficiency Act of 2002, the Freedom of Information Act, Title 13 of the United States Code (i.e., the Census Act), and Title 26 of the United States Code (i.e., the Tax Code).

⁴ 20 U.S.C. § 1232g; 34 C.F.R. Part 99.

as they collect, process, and maintain nonstandardized and duplicative records. While we have not calculated the cost agencies incur due to these constraints, it is evident that government-wide duplication of effort is expensive. By excluding the Bureau's address list from Title 13 restrictions, Congress would allow the Bureau to collaborate much more effectively with federal and local governments to improve these records. In addition, by continuing to encourage the Bureau's use of administrative records, Congress would help bring further improvements to the address list—as well as reduce the number of costly visits to housing units that do not return the decennial questionnaire.

This discussion of enhancing the Census Bureau's use of administrative records and data sharing supplements two responses to questions for the record that we provided following the July 18, 2012, hearing (see enclosure). Thank you for soliciting our further input on how the Bureau can execute a 2020 decennial that reduces costs while maintaining or improving accuracy. We are available to meet with you and your staff to discuss these issues in more detail. Please feel free to contact me at (202) 482-4661, or Ann Eilers, Principal Assistant Inspector General for Audit and Evaluation, at (202) 482-2754, for any further assistance.

Sincerely,

A handwritten signature in black ink that reads "Todd J. Zinser". The signature is written in a cursive style with a large, stylized "T" and "Z".

Todd J. Zinser

Enclosure

cc: The Honorable Thomas R. Carper, Chairman
The Honorable Scott Brown

Enclosure

Post-Hearing Questions for the Record
Submitted to the Commerce Inspector General Todd Zinser
from Dr. Tom Coburn
“Census: Planning Ahead for 2020”
July 20, 2012

1. **In your testimony, you discussed how the Census Bureau could improve its efficiency and accomplish its mission better if it had access to other federal agencies’ data and information about citizens. Can you please describe what type of information sharing between agencies currently exists and what is needed? Please describe any legislative or regulatory changes that would need to be made to allow the Census Bureau to access other agencies’ data and information.**

Information Sharing Between Agencies and the Census Bureau

Congress has authorized the Department of Commerce and, ultimately, the Census Bureau to acquire and use information available from any sources by purchase or otherwise—from states, counties, cities, or other units of government, or their instrumentalities, or from private persons and agencies—for the conduct of statistical activities instead of direct inquiries.¹ Congress has provided additional legislative support for these activities with the Bureau’s specific exemption in the Privacy Act of 1974 (as amended).² The Bureau actively uses these authorities to collect information and minimize respondent burden.

The Census Bureau receives billions of records each year from a variety of other federal agencies to conduct its own demographic and economic statistical programs, as well as statistical activities it conducts for other federal agencies,³ through the Statistical Administrative Records System (StARS). The Bureau began building this system in the mid-1990s, and agencies currently providing data for this system of records include the Departments of Agriculture, Education, Health and Human Services, Homeland Security, Housing and Urban Development, Labor, Treasury, and Veterans Affairs, as well as the Office of Personnel Management, the Social Security Administration, the Selective Service System, and the U.S. Postal Service.⁴ In addition, StARS includes data from state governments and private entities.

Most federal agencies, state governments, and private organizations collect and maintain basic information on the individuals and businesses they serve. Their records typically include the type of contact information (e.g., names, addresses, and telephone numbers) that could be used to improve the efficiency and effectiveness of statistical programs such as the decennial census. In addition, organizations maintain a wide array of other information that they have individually determined to support their mission. The types of information

¹ 13 U.S.C. § 6.

² 5 U.S.C. § 552a(b)(4).

³ Censuses, surveys, and other statistical studies as authorized by 13 U.S.C. §§ 1 – 402 and 15 U.S.C. § 1525.

⁴ As defined in its Privacy Act System of Records notice *COMMERCE/CENSUS-8, Statistical Administrative Records System* (74 FR 12834-02).

vary by source agency and program within those agencies. Therefore, the records the Bureau maintains in StARS may contain the following types of information:

- *Demographic information* such as date of birth, sex, race, ethnicity, household and family characteristics, education, marital status, tribal affiliation, and military veteran status;
- *Geographical and respondent contact information* such as name, address, telephone number, and geographic codes;
- *Mortality information* such as cause of death and hospitalization information;
- *Health information* such as type of provider, services provided, cost of services, and quality indicators;
- *Economic information* such as housing characteristics, income, occupation, employment and unemployment information, health insurance coverage, federal program participation, assets, and wealth; and
- *Business information* such as business name, revenues, number of employees, and industry codes.

Data Use Constraints

The Census Bureau's receipt of information from supplying agencies has required years of negotiations. Each data-providing agency, guided by its own statutes and policies regarding acceptable disclosure and use, places constraints on the information it supplies. When the Bureau uses records from multiple agencies to conduct a specific operation, the interplay between these legal constraints can be quite complex and limiting. As a result, the Bureau has only used administrative records in a limited fashion in prior decennial census operations. In some cases any presentation of data that describes the source of the record would violate a confidentiality provision. The merger of information from multiple data sources, such as the Internal Revenue Service and the Bureau, while removing information that describes a specific record's source, protects confidentiality. According to the Bureau, in a recent request for information from the Department of Education, there is a question as to whether that department can even provide information to the Census Bureau due to the privacy requirements mandated by the Family Educational Rights and Privacy Act.⁵

The Census Bureau has its own constraints on the return of information to other federal agencies. Agency information that it merges with data collected under the authority of Title 13 cannot be returned to that agency in an identifiable form, even if that entity is another federal statistical agency.⁶ Legal exceptions have been required to authorize the sharing of information for statistical purposes. One exception was permitted with the Census Address

⁵ 20 U.S.C. § 1232g; 34 C.F.R. Part 99.

⁶ 13 U.S.C. § 9.

List Improvement Act of 1994,⁷ which allowed the Bureau to share address information with local government entities for the conduct of the decennial census and other statistical activities, provided they signed a confidentiality agreement.

Another important example of a Congressional mandate to share data is the Confidential Information Protection and Statistical Efficiency Act of 2002.⁸ This Act provided a baseline of confidentiality protections for statistical information in Subtitle A; it also authorized the Census Bureau, the Bureau of Economic Analysis, and the Bureau of Labor Statistics to share business data for the efficient and consistent representation of businesses and the national economy in Subtitle B. Much of the business data held by the Census Bureau comes from the Internal Revenue Service, and the Bureau adds value through processing, coding, and merger with its own economic data collections.

All agencies operate under confidentiality and data sharing constraints that have been developed and promulgated after more than a century of federal laws and Supreme Court decisions. The United States has grappled with the often contradictory priorities of the government's need for information, to effectively and efficiently provide appropriate representation and services to its populace, and an individual's right to privacy and protection from harm. The conduct of a census or a statistical survey, by its very nature, places burden on respondents and impacts their privacy. However, this need for statistics versus the burden placed on respondents is balanced by legislative controls such as the Paperwork Reduction Act.

Considerations that Could Improve the Census Bureau's Efficient Production of Information

Agencies are only custodians of personal and business information that is owned by the people of the United States, and the use of that information is directed by Congress. The federal statistical system uses those records to create important economic indicators, as well as demographic and socio-economic statistics, to help inform the apportionment of representation, determine policy, and direct resources. In addition, these statistics inform the populace and provide businesses with the information they need to make sound decisions that eventually fuel our economy. Given the staggering cost growth experienced over prior decennials, the Census Bureau needs a new method of conducting the decennial census to reduce costs while maintaining or improving accuracy.

It is in the government's interest to foster a thriving statistical system even in a constrained budget environment, and the Census Bureau is a key component of our national statistical system. The government's ability to enhance the use, or reuse, of information that people and businesses have already provided reduces the burden placed on them to respond to questionnaires and reduces intrusion of their privacy. Agencies have already incurred the cost of collecting and processing the information they house. Reusing existing data minimizes the cost incurred by the government to collect data for statistical activities.

⁷ P.L. 103-430.

⁸ P.L. 107-347, Title V.

We are not recommending that the Census Bureau conduct a cost-effective decennial census solely through the use of existing data. It is important that individuals have the capacity to inform the government where they live and provide the basic characteristics required for a functioning representative democracy. However, if individuals are not directly contacted in the decennial census process, or they do not to provide their information, they can still be enumerated through information the government already holds.⁹ In addition, existing information can help the Bureau develop more cost-effective and efficient census and survey data collections. There are potential ways to improve the statistical system, and reduce costs, while minimizing the impact on individual privacy:

1. *Congress should consider whether address information could be shared more freely and continuously among agencies, including state and local governments, for the purpose of statistical operations.* Address and map information is widely available on the internet and through data vendors. These data alone do not impact the privacy of individuals. However, the confidentiality constraints under which the Census Bureau operates do not allow the sharing of this type of information with other agencies and leads to inefficient and perhaps ineffective government-wide operations. This new approach might require legislation and the further development and promulgation of address and mapping standards.
2. *Congress should consider a legislative requirement for federal agencies to provide unclassified information to the Bureau for the sole purpose of producing relevant, high-quality statistics.* Even though the Census Bureau has the legal authority to request data from other federal agencies for statistical activities, those agencies are not required to provide the information to the Bureau. Furthermore, when agencies do provide information, it comes with legal constraints on data access and use and often requires labor intensive negotiations. New legislation requiring federal agencies to provide the Bureau their information could reduce this burden and enhance statistical operations—including, but not limited to, the decennial census and the long-form replacement American Community Survey.
3. *Congress should consider legislation that transfers the custodial requirements from the source agency to the Census Bureau when that agency provides its information.* Agencies are very diligent about data confidentiality and appropriate use of information they collect, maintain, and eventually provide to the Census Bureau. Serving as custodians of the public interest and national records, these agencies continue to encounter the burden of overseeing and reporting on the Bureau's use of the data they have provided. This burden remains with agencies, despite the fact that the Bureau's confidentiality protections are more restrictive than the protections applicable to most federal unclassified information. New legislation would help eliminate this burden.

⁹ As provided by 13 U.S.C § 6.

2. **Do you think that the Census Bureau’s strategy for mapping is efficient? Could the Census Bureau improve its efficiency by using commercially available maps? Please provide any recommendations for how the Census Bureau could improve its mapping strategies.**

OIG has not specifically conducted an analysis of the Census Bureau’s strategy for the acquisition and use of commercially available products either to replace or improve its own maps. Our evaluation efforts have focused on addresses, the critical means for locating and enumerating the population in the decennial census. As required by Congressional apportionment and redistricting, the Bureau maps link each address and other living quarter identifiers to a specific spot on a map through a process called geocoding. These addresses and map also provide tallies for political and statistical geographic entities.

In May 2012, we released our evaluation of the impact that various address-updating operations (including U.S. Postal Service address updates) have on the Census Bureau’s combined master address file (MAF) and topologically integrated geographic encoding and referencing (TIGER) map database (combined, the MAF–TIGER database, or MTdb). In this report, we identified trends that introduced error in its production process.¹⁰ In addition, we reviewed the process by which Bureau field staff and local governments provided address, map, and boundary updates in the 2010 decennial.

We found that the Census Bureau’s efforts to assess MTdb quality were unsuccessful in 2010. In addition, the Bureau did not realize its goal of updating address and map information from tribal, state, county, and local government partners. It must meet both of these goals to implement a 2020 decennial address-canvassing operation with reduced costs. Furthermore, as of June 2011, the MAF contained 3.5 million ungeocoded records—a number likely to rise, as it did during the 2010 decennial. Without maintenance of the MTdb by continuous geocoding throughout the decade, the Bureau will again have to rely on an expensive end-of-decade address-canvassing operation. Finally, the MAF updating process of accepting more recent address changes without adequate verification may result in a lower-quality address list.

We recommended that the Census Bureau:

1. Develop an MTdb measure for determining address list quality at a low level of geography that (a) provides a fair and equal opportunity for targeting selection, (b) drives selection and planning decisions, and (c) is well-documented and verifiable.
2. Work with the Department to determine the feasibility of improving methods of sharing MTdb information throughout the decade with governmental entities (partners) to create a uniform, national address list.

¹⁰ See U.S. Department of Commerce Office of Inspector General, May 2012. *High-Quality Maps and Accurate Addresses Are Needed to Achieve Census 2020 Cost-Saving Goals*, OIG-12-024-I. Washington, DC: Department of Commerce OIG.

3. Investigate and remedy the exclusion of 500,000 ungeocoded address records, which had been designated as valid U.S. Postal Service delivery addresses, from the 2010 decennial.
4. Conduct the necessary research, develop a proven methodology, and allocate the necessary funds to continuously reduce the number of ungeocoded records throughout the decade.
5. Develop and implement quality indicator tools, including use of administrative records,¹¹ to ensure that updates to the MAF are accurate.

Congress should continue to encourage the use of administrative records to improve the address list and reduce the number of visits to housing units that do not return the questionnaire. Although tribal, state, county, and local governments share address information with the Census Bureau, Title 13 forbids the Bureau from reciprocating with those partners—except for a few very narrow exceptions, such as the once-a-decade address-updating program. The Census Address List Improvement Act of 1994 did amend Title 13 to provide a limited exception to these restrictions for local governments, allowing two-way sharing of address information between the Bureau and officials designated by local government units. However, the program has several requirements, including an invitation and review phase, as well as an appeals process. The 2010 decennial's program occurred from January 2007 through March 2010. Although it may require legislative action, a more informal method of two-way sharing of address lists earlier in the decade could improve address updating and geocoding as well as the cost-effectiveness of the Bureau's demographic censuses and surveys throughout the decade.

¹¹ The Census Bureau receives administrative records from a variety of sources. Bureau use of these records in the address updating and verification process could substantially reduce the size and scope of expensive field operations in the 2020 decennial. See the first consideration, at the end of our first response, for a more detailed discussion of the potential legislative issues inherent in these uses of administrative records.