

Statement of

Warren A. Brown
Senior Research Associate
Director, Program on Applied Demographics
Research Director, New York Census Research Data Center
Cornell Institute for Social and Economic Research
Cornell University

Before the
Subcommittee on Federalism and the Census
Oversight Hearing

Committee on Government Reform
U.S. House of Representatives

Hearing Topic: "Two Plus Two Should Never Equal Three: Getting Intercensal
Population Estimates Right the First Time."
Wednesday September 6, 2006
Room 2247 Rayburn House Office Building

Introduction

Chairman Turner and members of the House Government Reform Subcommittee on Federalism and the Census, my name is Warren Brown and I am a demographer at Cornell University in Ithaca, New York. Thank you for the opportunity to be part of this hearing on "Two Plus Two Should Never Equal Three: Getting Intercensal Population Estimates Right the First Time." I direct Cornell's Program on Applied Demographics and am Research Director for the New York Census Research Data Center, a consortium of research institutions in the New York metropolitan area and upstate New York. I graduated from Cornell University in 1982 with a Ph.D. in Sociology and minors in Demography and American History. I began my career in demography as a graduate assistant at Cornell working with Cooperative Extension on program planning. I learned to apply demographic information in planning programs ranging from nutrition education for low income families, to agricultural land preservation in rapidly suburbanizing communities, to urban gardening in New York City. I was introduced to methods for estimating population when I attended a short-course conducted by Don Starsinic and Fred Cavanaugh of the Census Bureau in 1978. I have been working with demographic estimates and projections ever since. I represent New York State in the Federal State Cooperative for Population Estimates, and am a past chair of its Steering Committee.

Observations on the Census Bureau's Program for Population Estimates

I appreciate the opportunity to share my observations regarding the Census Bureau's population estimates activities. These are largely based on my experience

working with staff in the Census Bureau's estimates program on behalf of New York State. My testimony covers four areas: 1) population estimates are important; 2) estimates are not perfect; 3) description of the current programs to make and correct estimates; and 4) ways to improve population estimates.

Population Estimates are Important

Title 13 of the U.S. Code—the law that authorizes and regulates the Census Bureau—requires that the Bureau produces population estimates for all states, counties and local units of general purpose government. The reason for producing estimates, given in Section 183 of Title 13, is “for the purpose of administering any law of the United States in which population or other population characteristics are used to determine the amount of benefit received by State, county, or local units of general purpose government...” The exact amount of money distributed by the federal government on the basis of the population estimates is hard to determine and estimates range from \$100 billion to \$200 billion. Suffice it to say, a considerable amount of money is allocated on the basis of these numbers.

The Census Bureau itself is one of the primary consumers of the estimates. The Bureau uses the post-census estimates as controls for its various surveys, including the Current Population Survey and American Community Survey. The estimates must be delivered within a relatively short time frame for use in the weighting and estimating stages of surveys, which in turn must be produced by their deadlines. Private sources of investment affecting states and local governments use the estimates as well in order to identify growing markets. Most of the private sector data vendors control their estimates of population size and consumer characteristics to the Census Bureau's estimates of counties and municipalities. The estimates are also grist for the media informing their publics as to the fortunes of states, regions and communities. Each time the Census Bureau releases a new set of estimates eager reporters contact those of us who labor in this area of demography, asking us why an area grew or declined. Then subsequent to the publication of the estimates, representatives of local governments reacting to the news articles, call attention to the estimates as a sign of local prosperity or discredit the estimates and question their accuracy.

Estimates are Not Perfect

Population estimates are not perfect and some error is inevitable. The goal is to minimize error and produce estimates that are as accurate, precise and equitable as possible. The decennial census is used as a starting point for estimates in the post-census interval and as a standard for evaluating estimates at the end of the decade once a new census is conducted. In general, estimates are not intended to overcome shortcomings in the decennial census, rather the objective is to produce estimates of population and characteristics that would be reported if an enumeration of the population, on the order of the previous census, were conducted. The Census Bureau itself is the source of some of the best information on the quality of their post-censal estimates. They produced a file of population estimates for April 1, 2000 based on the 1990 Census in order to evaluate the level of error in their methods. The estimates were low and the national estimate of population was 6.8 million persons below the Census 2000 count of resident population. That was a -2.4 percent error. Had all states, counties and municipalities been similarly

low then at least the estimates would have been equitable. They were not equitable, however. While the population estimates for all states were low and not one state was overestimated, there were inequities. On a percentage basis the states with the greatest under estimates were the District of Columbia at -9.3 percent; Nevada at -6.8 percent; Arizona at -5.4 percent; and Rhode Island at -5.1 percent. The Census Bureau's population estimates were less than 1 percentage point off for West Virginia, Michigan, Ohio, Alaska, and Kansas.

Demographers tend to focus on the accuracy and precision of estimates, while the issue of equity is the least pursued even though it is perhaps the most important dimension to quality. David Swanson [Swanson, David A., "Allocation Accuracy in Population Estimates: An Overlooked Criterion with Fiscal Implications," in *Small Area Population Estimates: Methods and Their Accuracy and New Metropolitan Area Definitions and Their Impact on the Private and Public Sector*, Series GE-41, No. 7 (Washington, DC, Government Printing Office, 1981), pp. 13-16] discusses this issue in the use of estimates to allocate monies. Swanson developed the "Index of Misallocation" as a summary measure of the shares based on an estimate that would need to be reallocated in order to match shares based on a census count. If we look at how monies would be distributed to the states on the basis of the Census Bureau's research estimates for April 1, 2000, we get a better sense of the inequities. The District of Columbia's share was 7.0 percentage points below what it should have been, while West Virginia's was 2.2 percentage points above its actual share based on Census 2000. Table 1 provides a summary of the necessary reallocation of shares to states, as expressed in the previous sentence.

Current Programs to Make and Correct Estimates

The Census Bureau has a number of programs in place to improve the quality of its estimates. The most effective has been the Federal State Cooperative for Population Estimates (FSCPE). Each state appoints a representative, typically the state demographer, to participate in the FSCPE. The Census Bureau represents the federal government. The National Governor's Conference and the Council of State Governments sponsored the 'First National Conference on Comparative Statistics' in 1966. Out of this conference the FSCPE began to evolve from an informal arrangement between the Census Bureau and state demographers and by 1973 became a formal organization. The by-laws of the FSCPE call for:

- promotion of cooperation between the states and the U.S. Census Bureau;
- preparation of a set of consistent and jointly prepared county and subcounty estimates with complete state coverage;
- assurance of highest quality estimates through the use of established methods, comprehensive data review and thorough testing;
- reduction of duplication in the production of population estimates and improvement of communication among the groups compiling population figures;
- improvement and advancement of techniques and methodologies and the encouragement of joint research efforts; and
- enhancement of the recognition of local demographic work.

During the 1980s the Census Bureau and the individual states jointly selected methods most appropriate for that state. Two or three methods were employed and the results were averaged in order to arrive at estimates for the state's counties. The selection of methods was determined by rigorous testing against the 1980 census. If a state had a unique data source that worked well for that state—such as California's driver license address change file—then it was used. Different methods using different data were the practice.

Under pressure to streamline the estimates process and produce a uniform methodology, the Census Bureau adopted a single method approach to estimates in the 1990s using primarily federal administrative records to estimate components of population change. The previously active participation of states in the production of estimates was reduced to an advisory and review role. The states do have an opportunity to review the Census Bureau's data on births, deaths, internal and international migration for the population estimates and data on building permits, mobile home placements and loss of housing for the housing unit estimates. The Census Bureau has established an online clearinghouse for state representatives (FSCPE members) that posts the issues raised by states in this review process and the responses of the Census Bureau. The Census Bureau began the clearinghouse with comments on the estimates being prepared for 2003. The clearinghouse covers comments for 2003 and 2004, and comments for 2005 have not yet been added. Slightly less than half of the states have participated in the clearinghouse which is a vital early step in pointing out problems with county and sub-county estimates.

Once estimates have been made official and publicly released, local governments have the opportunity to challenge the numbers. Frequently the state's FSCPE representative will contact the local government suggesting they challenge, or respond to an expressed interest by the local government in challenging the number. The Census Bureau's challenge process is far more collaborative than adversarial. The Census Bureau provides the local government with guideline materials, a workbook for preparing their challenge, and informs them that their state FSCPE representative is available to provide technical assistance in preparing their challenge. The challenge process uses an alternate methodology, the Housing Unit Method, to prepare estimates. The Housing Unit Method is widely used by states and local governments in preparing their own estimates, independent of the Census Bureau, and has the potential for producing good quality estimates. A variation of the Housing Unit Method is used by the Census Bureau to prepare its sub-county estimates. One of the reasons it is not used by the Census Bureau for preparing their county estimates is the problem in getting quality input data in a timely manner. Most of the inputs are from local administrative records. It is an appropriate method for locals to use in challenging the county and sub-county population estimates. County estimates are based on federal administrative records on internal migration (matched IRS tax returns and Medicare enrollments) and guesstimates of international migration; and sub-county estimates are based on the Census Bureau's survey of local governments regarding residential building permits, and weak estimates of mobile home placements and loss of housing units.

The number of successful challenges has increased each year. In 2001 there were 3 successful challenges; in 2002 there were 14; then 27 in 2003; and 38 successful challenges in 2004. So far in 2005 there have been 8 successful challenges but the books

are not yet closed on 2005. These represent an exceedingly small fraction of the units of local government for which estimates are prepared. Given the errors we cited in the estimates of state population earlier in our testimony, we have to conclude that the vast majority of local governments are unaware of errors in estimates of their population or are not taking advantage of the challenge process. This may be a resource issue and therefore the fairness of the challenge process itself may be subject to inequitable conditions. We also have to note that none of the challenges since Census 2000 have resulted in a lower estimate, although we suspect that in some cases the Census Bureau estimates for local areas may be too high.

But even for those units of local government that have successfully challenged the pay-offs are limited. At what point are the allocations of funds based upon the estimates being made? Do any of the allocation programs wait for the outcome of the challenge process or re-allocate monies once challenges have been settled and corrected numbers are released? The Census Bureau's surveys that use the estimates as controls in making their reports don't use the corrected estimates either. The Current Population Survey and American Community Survey use the initial estimates and do not revise their reports to reflect the corrected totals. New York City is preparing its third consecutive challenge to the Census Bureau's estimates. Although they were successful in challenging the 2003 and 2004 estimates, the Current Population Survey and American Community Survey for those years don't reflect the corrected data. Correcting the estimates after they are released does not necessarily correct all the downstream uses made of the initial estimates.

In Figure 1 the Census Bureau's estimates of population for New York State are displayed for the years 2000 through 2005 produced in Vintages 2003, 2004 and 2005. Focus on the differences between estimates for the same year produced in the different vintages. Note that the differences are very slight for July 1, 2000 and increasingly pronounced with each subsequent year. On a percentage basis the differences are still slight and by narrowing the range of the vertical scale to vary between 18.85 million and 19.35 million I have accentuated the differences. The greatest difference is between the Vintage 2004 and Vintage 2005 estimates for 2004. The numerical difference is 53,639 which is only a difference of 0.28 percent. What makes this small difference meaningful is the difference it makes in the trend line. It is the difference between continued growth and a downward decline. The Vintage 2005 estimates of county population for 2005 are going through the challenge and revision processes and are likely to make 2005 look slightly different in the Vintage 2006 estimates.

The estimate of national population does not change and so the challenge and revision process is a zero-sum game. The population increases of counties gained through the challenge process are subtracted from the balance of counties in the country. To simplify, if you are not successfully challenging then you are losing. The losses are small when spread across all the remaining counties, but that is why some counties, and their states, increase in population for a given estimate year with succeeding vintages and others decrease. Which are the correct estimates for New York State for 2003, 2004, and 2005? Which estimates should be used for the allocation of resources, as controls for surveys whose data are used in the allocation process, and as a signal to private investors and the public as to the trend line in growth, stability or decline?

The answer is obvious. Use the correct estimate and get it right the first time! How can that be done?

Improving Population Estimates

The way to correct the dilemma of which vintage estimate to use for a particular year is to get the estimate right the first time, less obvious is how to do that in a cost effective and timely manner given budgetary and reporting deadline constraints. The Census Bureau sponsored a conference titled, “U.S. Census Bureau Population Estimates: Meeting User Needs” that was held July 19, 2006. At the conference, Joseph Salvo of the New York City Department of City Planning, and I made a presentation titled, “Population Estimates and the Needs of Local Governments.” I am submitting a copy of that paper in addition to this testimony. Here I wish to summarize some of our points.

The current procedure used by the Census Bureau to estimate county population relies entirely on a single method, called the Administrative Records Component Method. Reliance on a single method using administrative data from federal agencies has the advantage of streamlining the estimates process and allowing the estimates to be produced in a timely manner. The disadvantage is that individual methods tend not to work equally well for all counties. A second disadvantage is that relying entirely on a few administrative record series leaves the method vulnerable to changes in those series that render the input data less consistent and accurate. The Administrative Records Component Method had relied on immigration statistics from the Immigration and Naturalization Service (INS) in order to allocate estimates of international migration to individual counties. Following the terrorist attacks on September 11, 2001 the role of INS shifted toward national security and away from immigration statistics. This left a void that required the Census Bureau to make “guesstimates” regarding the counties of destination for international migrants.

The use of multiple methods provides more stability and complementary strengths in working with a diverse array of counties. Prior to the emphasis on streamlining, the use of multiple methods and more sources of inputs had been championed by the Census Bureau itself. Evaluation research has shown that the averaging of methods produces better quality estimates than any single method alone. Among the methods that ought to be used for county estimates is the Housing Unit Method. The Housing Unit Method calculates the population in households as the product of housing units, occupancy rates and average household size. When persons in group quarters (i.e., prisons, nursing homes, dormitories and other facilities) are added to persons in households, an estimate can be created for the total population. In the Census Bureau’s guidelines for challenging estimates, the Housing Unit Method is the accepted alternative to the Administrative Records Component Method. Why not employ the Housing Unit Method, along with the Administrative Component Method to begin with, rather than in a challenge process after release of the official estimates?

Implementation of the Housing Unit Method and other methods as well, will require a closer working relationship between the Census Bureau, the state representatives of the Federal State Cooperative for Population Estimates, and local governments. This calls for a return to earlier practices when the states were full and equal partners with the Census Bureau in the estimates process. Calling for such a partnership is easy; however making it work will be a difficult undertaking. The Census

Bureau has excellent partnership programs, and in addition to the Federal State Cooperative for Population Estimates there are the State Data Centers and the Census Information Centers representing the interests of underserved communities. On the research side, the Census Bureau's Center for Economic Studies administers a network of Census Research Data Centers located at major research universities and similar research institutions. These Census Research Data Centers could serve to facilitate joint efforts between staff at the Census Bureau and external researchers to experiment with multiple methods for estimating population and to develop improved approaches. The Census Bureau's Longitudinal Employer Household Dynamics Program (LEHD) could serve as a model for collaboration between internal and external researchers to develop superior data products. The LEHD Program is providing timely and detailed information on labor markets that go well beyond what was previously available.

In summary, getting estimates right the first time (or at least with maximum accuracy, precision and equity) calls for a fuller partnership between the federal, state and local governments; use of multiple methods; and research on the integration of administrative records—federal, state and local—with sample survey information.

Thank you for inviting me to participate in this oversight hearing.

Table 1: Misallocation to States Based on Research Estimates for April 1, 2000

State or District	Pct Error
Alabama	1.0%
Alaska	1.6%
Arizona	-3.0%
Arkansas	-1.4%
California	1.4%
Colorado	-1.9%
Connecticut	-0.9%
Delaware	-0.6%
District of Columbia	-7.0%
Florida	-2.0%
Geogia	-1.1%
Hawaii	-0.3%
Idaho	0.4%
Illinois	0.4%
Indiana	0.6%
Iowa	0.7%
Kansas	1.5%
Kentucky	0.9%
Louisiana	0.3%
Maine	1.0%
Maryland	0.8%
Massachusetts	0.0%
Michigan	2.1%
Minnesota	0.3%
Mississippi	0.2%
Missouri	0.6%

State or District	Pct Error
Montana	0.7%
Nebraska	0.0%
Nevada	-4.5%
New Hampshire	0.5%
New Jersey	-0.2%
New Mexico	-1.7%
New York	-1.4%
North Carolina	-1.7%
North Dakota	0.6%
Ohio	1.7%
Oklahoma	0.3%
Oregon	-0.1%
Pennsylvania	0.0%
Rhode Island	-2.8%
South Carolina	0.0%
South Dakota	0.0%
Tennessee	-0.6%
Texas	-0.2%
Utah	-1.0%
Vermont	0.4%
Virginia	0.6%
Washington	0.8%
West Virginia	2.2%
Wisconsin	0.7%
Wyoming	-0.3%

Figure 1: Estimates Are Subject To Revision

