
Issues Related to Using Administrative Records in Census 2000

Robert W. Marx, Bureau of the Census

At the first of two sessions on Administrative Records this morning, we heard lots of good information about why the Census Bureau should use these records and learned--or relearned--many of the reasons why life is never as simple as we think it ought to be.

I am sure these "pros" and "cons" were well understood by all experts several years ago. Nevertheless, a distinguished panel of the National Research Council (NRC), chaired by one of our own, Norm Bradburn, issued a report about two years ago entitled "Counting People in an Information Age." That report reinforced work already planned for the 1995 Census Test and concluded with the following recommendation:

- "1. The panel urges the Census Bureau to adopt a proactive policy to expand its uses of administrative records, and it urges other executive branch agencies and Congress to give their support to such a policy. Doing so would be in keeping with the Census Bureau's tradition of innovation and adaptation to the technical and social environment in which it carries out its mission as fact-finder for the nation."
- "2. Any proactive policy has some risks associated with it, but the panel believes that the risks are justified by the potential benefits. Clinging to the twentieth century census model in a twenty-first century data collection environment could well prove to be even more risky."

Why would this panel make such a strong recommendation?

First, the mission of statistical agencies is to meet the information needs of government and society as effectively and efficiently as possible, using all available sources of information. The primary sources of statisti-

cal information about people are censuses, statistical surveys, and a large and diverse set of administrative records systems that have been created primarily for nonstatistical purposes.

Second, cost. A major reason for considering an administrative records census is its potential to be less expensive than a revised traditional methodology. Tom Jabine noted another very good reason in the previous session when he pointed out the obvious conclusions from the 1995 Census Test results--the traditional procedures, even the enhanced ones, still miss lots of folks, and administrative records find some of them!

Jerry Gates provided an excellent overview of the major issues in his opening remarks. Let me quote again from the panel's report to summarize the issues as they saw them:

- "1. *Privacy*: Is the American public prepared to accept the use and linkage of administrative records for the purposes of a census?"
2. *Coverage*: Can administrative records deliver the level and distribution of overall coverage needed from a census?"
3. *Geography*: Can administrative records allocate individuals accurately by place of residence as required by a census?"
4. *Content*: How much of the traditional content can we expect an administrative records census to deliver, and with what quality?"
5. *Cost*: How would the cost of an administrative records census compare with other alternatives? Approximately \$25/case to do a followup visit; and the Congress has told us to reduce the cost of the census."

I would add a sixth, which the panel actually discussed in great detail--namely, *Access*: Will the holders of the records provide access for the Census Bureau to use the records?

Starting with the "added" issue first, and reinforcing, from a slightly different perspective, why the Census Bureau believes it must be proactive in its pursuit of making administrative records work, I quote from Title 13, U. S. Code:

“. . . to the maximum extent possible. . . the Secretary shall acquire and use information available from any source. . . instead of conducting direct inquiries.”

I know the next thing I say will be a great surprise to many of you, because you probably believe, as I once did, that such things could never happen in a great country, such as the United States. But just because a law passed by Congress says we should do something, and that we have not only the right but the responsibility to be given access to the information held by others that would make it work, does NOT mean that we, in fact, can have access.

Sometimes, this is because Congress, unbelievable as this may seem, passes another law, pertaining to the agency holding the records, that says to them, "Thou shalt not give these records to anyone." This is a real dilemma for everyone.

Sometimes, it is because the law is somewhat ambiguous, and, as good bureaucrats, and in keeping with the advice of their even more cautious lawyers, the holders of the records want to be very cautious--better safe than sorry. This was the case with the U. S. Postal Service two years ago regarding their providing the Delivery Sequence File for our use. Finally, Congress had to step in and direct them to do what they could have done voluntarily. This is a real waste of taxpayer dollars and the time of Congress that has plenty of other things it needs to get resolved.

Speaking of issues for Congress, and the consistency of the guidance they provide on many issues, I give you one more quote from the NRC Panel:

“If the ideal is to be rejected on political grounds, it might as well be rejected before significant investment is made.”

To pursue the "access" issue just a little longer, there is a third category of agencies that are cautious sharers--for limited purposes and/or for testing. And, occasionally, we encounter an agency that just says, "Well, of course, here is what we have. Why didn't you ask sooner? It makes so much sense for us to all work together. It is the same taxpayer who is paying all our salaries, and this sharing makes so much common sense." A clearly enlightened group of folks!

Just to finish off the "access" discussion, I quote again from the panel's report:

“Many of the custodians of administrative records systems recognize that sharing their records for statistical uses would have benefits. Several officials said they would feel more comfortable sharing the records if specifically authorized to do so by their legislative authority. They were all conscious of the need to inform data subjects about how their data would be used and to inform the public about the benefits and risks associated with data sharing. They felt the need for some mechanism, such as an individual ombudsman, an expanded role for OMB, or a data protection board, "to develop and oversee policies on data sharing and to provide a balance between the interests of data providers and data users.”

The "content" and "coverage" issues raised by the panel are items for which there are empirical processes to arrive at an answer--and we have been working on that in our 1995 Census Test, in our planning for our 1996 Community Census, and in our planning for our 1998 Dress Rehearsal.

The administrative records sets available in the United States clearly are not "perfect" for use in a decennial census environment--but, then, neither are our current data collection procedures or the cooperation of the public with those procedures. The comparison of

results needs to be benchmarked not against “perfect” but against “no response”; is some information better than no information?

On this issue of “who” has “what” in their records, and “how” it ought to be, the panel made another recommendation:

“The Office of Management and Budget should review identifiers, especially addresses, and demographic data items currently included in major administrative records systems with a view to promoting standardization and facilitation of statistical uses of information about individuals both in these records systems and in new ones that may be developed.”

Clearly, the panel could not have anticipated the political changes that were on the horizon and the less than enthusiastic reception that might await a request to impose greater reporting burdens on program participants or managers. But those ideas could change again, and, perhaps, a view might develop that those who benefit from government programs also have a responsibility to help reduce the cost of government, in general, by making release of their information a condition of participation.

The final issue is probably the thorniest, and I would not want to deprive the other panelists of a juicy subject to discuss, so I will say only a few words about “Privacy.” I quote again from the panel:

“We believe that this is necessary to proceed with public debate about the ethical, legal, and policy issues associated with statistical uses of administrative records. The Census Bureau has had contacts with privacy advocates in connection with previous censuses and has informed us that it plans future discussions that will focus specifically on statistical uses of administrative records in the 1995 Census Test. However, there has been some reluctance on the part of the agencies involved to enter a public debate for fear that calling attention to these questions might lead to discontinuance of important existing activities,

such as the use of income tax returns and Social Security data in the Census Bureau’s intercensal population estimates programs. We believe it is better to face these questions directly. Discussions are likely to be more productive if they focus on specific uses of administrative records, such as improvement of coverage in the 2000 census and the implications of an administrative records census in 2010, rather than on broad philosophical questions.”

Enough said. This is mostly an emotional issue of perception, slightly legal, and almost totally nonempirical.

Finally, there is the issue of, “So what should we be doing now?” Again, the panel had some advice:

“We have recommended that the Census Bureau follow a two-track approach to expanded uses of administrative records. The 2000 census track would identify and test new uses of administrative records considered feasible for use in the 2000 census. Possible uses include coverage improvement, content improvement, evaluation of the census, and measures to improve operational efficiency. The long-range track would develop and test procedures for a possible administrative records census in 2010 or beyond and uses of administrative records in other demographic data programs, such as current population estimates and current sample surveys, including any new surveys that might be started as part of a continuous measurement program.”

The Census Bureau has taken that advice very seriously. On Track 1, I spoke yesterday morning about our Plan for Census 2000 and noted that it could be summed up in five simple words: partnership, simplicity, technology, and statistical methods.

The second of the four strategies is to “Keep it simple.” Assuming that legal and technical issues can be overcome, the Census Bureau intends to use lists of people and addresses compiled for Federal, State, and tribal programs to improve the completeness and reduce

the cost of Census 2000 while reducing the reporting burden on the public. Three different uses of administrative records are contemplated:

1. To supplement the count in nonresponding households,
2. To enhance the quality check procedure (also known as Integrated Coverage Measurement or ICM), and
3. To supply missing items from households that responded.

The first use (nonresponse use) is primarily a cost savings use. If the Census Bureau can use information from administrative records to derive the census count for five percent of the nonresponding households, it can reduce the cost of the nonresponse followup operation by approximately \$72 million. This use of administrative records requires no contact with the household, but it does require a high degree of confidence in those records used.

Two other uses of administrative records are planned for Census 2000 that are not included in this discussion--the use of U. S. Postal Service information and local/tribal address files to update the Master Address File, and the use of various administrative records to compile lists of special living locations, such as hospitals and prisons. These are address-based lists that do not require the same kind of research as administrative records about individuals.

The anticipated benefit of the second use (ICM use) is an improvement in the completeness of the census. In this use, administrative records will help the Census Bureau identify people missed by the census that would

not otherwise be identified in the ICM procedures.

The third use (missing content use) would allow the Census Bureau to improve the quality of the data by completing some data items left blank or incomplete by respondents, items that would not otherwise be collected. We are investigating whether data contained on national administrative records are comparable to sample data collected in the census and the extent to which use of these data for missing responses would improve data quality over existing imputation methods.

Our ultimate decision about how much we can use administrative records for these three purposes will be made once we have the additional information from the 1995, 1996, and 1998 tests.

Also, as Tom Jabine noted in the previous session, even these planned uses are fairly "marginal"--5 percent of 30 percent nonrespondents or 1.5 percent of households to fill in missing people--and some fraction of the 0.6 percent of households in the ICM program.

Unfortunately, as Terriann Lowenthal noted in the Census 2000 session yesterday morning, this is NOT the kind of answer Congress likes to hear--they want "certainty" in agency plans, certainty in cost estimates, (and an easy target to shoot at).

We also plan to pursue Track 2 as recommended by the panel. This will be part of our Research and Experimental Program in Census 2000. We will also be continuing our discussion with OMB and the agencies that hold the various records we believe (or "know" at some point) will be of most value to Census 2000.

The path is likely to be rocky, but, then, most great changes do not come easily; they require lots of hard work.