

## THE PANEL ON FUTURE CENSUS METHODS

**Benjamin King, Statistical Consultant**  
2769 NW 27<sup>th</sup> Avenue, Boca Raton, FL 33434-3693

KEY WORDS: 2010 Census; planning; administrative records; American Community Survey.

I know that not all of you have been following these developments day-by-day. Thus to help you keep things straight, I must explain that, in addition to the panels chaired by Keith Rust and by myself, there exists another Committee on National Statistics (CNSTAT) panel that deals with the Census. It is chaired by Janet Norwood, and its other distinguished members (in alphabetical order) are Robert Bell, Norman Bradburn, Lawrence Brown, William Eddy, Robert Hauser, Roderick Little, Ingram Olkin, and Bruce Petrie. Since Keith's panel is now retired and mine has only recently met for the first time, Janet's panel, whose job it is to monitor the preparations leading up to and the actual execution of Census 2000, is the most currently operational. It is, not surprisingly, called the *Panel to Review the 2000 Census*. We refer to it informally as the 2000 Panel. It follows that the new *Panel on Future Census Methods* that I am going to discuss today is called the 2010 Panel.

Our 2010 Panel held its initial and only meeting thus far on June 7-8 of this year, and it is expected to continue its activities until the spring of 2003. Borrowing from the proposal written by Andy White, the Deputy Director of CNSTAT, in July 1997 that covers the two existing panels, and also from statements on CNSTAT's current web page, I shall try to describe its formal tasks as foreseen by its planners. They fall into three phases:

Phase I, picking up where Keith's *Panel on Alternative Census Methodologies* left off, involves a review of the plans for experiments and other methodological studies to be built into the 2000 Census and recommendations for fine tuning if called for. In addition, plans for collecting and retaining data to be used in the design of the 2010 Census must be reviewed. In view, however, of the late start of our panel, most of this activity has already been undertaken by the current 2000 Panel. That does not, of course, preclude the 2010 Panel from chiming in with its observations and opinions as it gets rolling. Although it may already be too late for our new panel to contribute much to the research design for 2000, we certainly can review and critique the anticipated mechanism for feeding the results into the planning for 2010. There is a process currently in place at the Census Bureau of evaluation of past research, execution of new research, and planning that carries the

Bureau from the end of one decennial census to the next one. It is especially important that our panel familiarize itself with the research design aspects of that process as soon as possible.

The second phase was envisioned as running from April 1999 until March 2001 when the final results for reapportionment and redistricting will be available. Thus we are already in Phase II and are running parallel to and congruent with the activities of the 2000 Panel. A difference as I see it, however, is that the 2000 Panel, in performing its task of watching closely the operations of the 2000 Census, will be meeting much more frequently than the 2010 Panel. Although only meeting a couple of times a year during this phase, our panel will be eager to observe the outcomes of the tracking system and the experiments in the 2000 Census and to digest the findings of our sister panel in that regard. To the extent possible we shall make recommendations concerning the best methods of analyzing the data produced by those systems in order to maximize the value of that research for the planning of the 2010 Census.

Finally, in Phase III, extending from March 2001 until the end of our tenure in 2003, we shall shift into high gear and synthesize our observations and those of the 2000 Panel, producing a formal judgment concerning the overall accuracy of the 2000 Census, evaluating the results of the built-in research studies, and reporting on their implications for 2010. For example, it has been suggested that we convene a workshop, or special meeting, on administrative records involving the leading experts outside the panel as well as panel members—that is but one of a number of possible special activities that may result from the 2000 findings. Most important of all, however, we must make recommendations for the research and experimentation that should be included in the prior testing, the dress rehearsal, and the execution of the 2010 Census.

Before I continue, this is a good time to enumerate the other members of our panel. They are (in alphabetical order):

DAVID A. BINDER, Methodology Branch, Statistics Canada

MICK P. COUPER, Survey Research Center, University of Michigan

WILLIAM D. KALSBECK, Department of Biostatistics, University of North Carolina

SALLIE KELLER-MCNULTY, Los Alamos National Laboratory, Los Alamos, NM

MICHAEL M. MEYER, The Boeing Company,  
Seattle, WA

DARYL PREGIBON, AT&T Labs Research, Florham  
Park, NJ

KEITH F. RUST, Westat, Inc., Rockville, MD

JOSEPH J. SALVO, Population Division, New York  
City Department of City Planning

JOSEPH L. SCHAFER, Department of Statistics,  
Pennsylvania State University

ALLEN L. SCHIRM, Mathematica Policy Research,  
Inc., Washington, D.C.

JOSEPH SEDRANSK, Department of Statistics, Case  
Western Reserve University

DONALD YLVISAKER, Department of Statistics,  
University of California, Los Angeles

I am sure that many of you, perhaps most of you, are familiar with the accomplishments of these eminent persons in their respective fields of statistics--if not, let me say that their varied expertise covers especially well the full spectrum of skills necessary to handle our tasks; and you can appreciate their sacrifice in the form of time taken away from their teaching, research, and other scholarly endeavors, as well as family activities, by serving on this panel.

I must also mention Michael Cohen, our able study director, and of course the CNSTAT regulars: Connie Citro, Andy White, and Meyer Zitter, who will be on hand for consultation and counseling.

Our kickoff meeting in early June was attended by Ken Prewitt, the Director of the Bureau, who spoke to us about the current outlook for the 2000 Census and our charge for 2010.

As a way of explaining some of our future activities, I am going to paraphrase some of the things that he told us. [Ken was not under oath, so he is free to change his mind later.]

Dr. Prewitt began by saying that the full-cycle cost of the 2000 Census is estimated to be between \$6.5 billion and \$7 billion, which is quite a stunner compared to the \$2.6 billion for 1990. There are many valid reasons for this increase— inflation, technological change, increased labor-market costs, and going from third-class to first-class mailing, etc.--but a very important question that he invites our panel to consider is that of how much increase in accuracy we are actually achieving with that increase in expenditures. As the decennial census takes a bigger and bigger bite out of the federal budget, the public has a right to know what it is getting for that — or to turn the question around, “How much accuracy do we lose by cutting the budget?” There apparently has been little systematic research in trying to answer such questions and Ken would like our panel to tackle the issue. Frankly, I think that we shall be lucky if we can

get as far as prescribing the kind of research that ought to be done, but we all agree that it is a high-priority item.

A second, and not unexpected, element in our charge concerns the effect on Census operations of new technological developments— for example, the Internet. In the 2000 Census it will be possible to fill out the short-form questionnaire via your PC. The Bureau is prepared for this in terms of having a system to process these results, but I doubt that it has a very good forecast of the magnitude of the response by this means. With some 80 million using the net, the number could be one third of that although my personal guess is that it will be small this time around. What about 2010, however? Can any of you say how big a factor this will be ten years from now? The way that things are going at present, it may be huge--or maybe, like the automobiles that convert to helicopters at the press of a button that I, as a kid reading *Popular Mechanics*, envisioned arriving by 1960, it will not quite have materialized. At any rate, we as a panel have to consider this issue and be prepared to deal with a range of possible scenarios.

A third factor that Ken Prewitt sees as very relevant for our panel is the fact that this 2000 Census is going to be more seriously scrutinized by the public than any preceding decennial operation. This is partly a result of the intense focus by the Congress as well as other interest groups and the media that has been applied to census planning during the past few years. Another reason, however, is the strong effort by the Bureau itself to combat problems of civic disengagement (e.g., low response rates, etc.) by forming partnerships with advocacy groups, church organizations, racial/ethnic interest groups, local governments, and so forth. There are now 22,000 such partnerships and the number is growing fast as we approach Y2K. Many of these organizations are more than just bush-beaters and cheer leaders— they have their own expert staffs and constituencies who want to be sure that they are not undercounted if anything can be done about it. Thus the mechanism for estimating the degree of under- and overenumeration and adjusting for it (within the limits of the law) will be under more intense scrutiny than ever before, and it is imperative that our panel learn as much as possible about its effectiveness and accuracy. The post-enumeration survey, now called the *Accuracy and Coverage Evaluation Survey (ACE)*, will consist of 300,000 households, higher than the 150,000 used in 1990 which the Bureau deemed to be an inadequate sample size. You may recall that the originally planned *Integrated Coverage Measurement (ICM)* survey was to have 750,000 households, but the Supreme Court decision forbidding the use of sampling

for reapportionment made direct state-level estimates unnecessary and thus the sample size could be reduced without compromising its purpose. Here I'll quote Ken's remarks at our meeting:

"...we are going to come out of 2000, we hope, with a stronger statement than we possibly could have come out with in 1990 about the importance of an ACE--or at least we will be able to make a strong statement one way or the other. We will have the National Academy committee to help us make that statement.

So your planning has to take into account, obviously, what we are going to learn from 2000, both technically and publicly or politically. One of the things that we have learned is that the census is owned, and should be owned, by the American public, not by the statisticians, the methodologists, the Census Bureau professionals, the National Academy of Sciences. So whatever 2010 is going to be, it has to be a census that a lot of people buy into, hopefully early on."

And later:

"All I am trying to say is that we cannot do 2010 planning without being awfully sensitive to the lessons of 2000. Those lessons are going to be very widely recognized and internalized by stakeholders that happen to live in this city or work in this city [Washington, DC]. There are many, many stakeholders in state houses, mayoralty offices, non-government organizations, and so forth and so on. So the 2010 planning has to be constantly sensitive to the messages that are going to come out of 2000 if the 2010 planning is going to work."

These are inspirational and also awesome words. As one who was involved in the CNSTAT panel concerning the 1990 census that was dissolved in 1987, I must say that although we were aware (deep inside at least) that the census belonged to its stakeholders the whole atmosphere was one that was much more attuned toward helping the Bureau improve its methods, thinking of those methods themselves as something quite esoteric that the public cared little about as long as it got the results that it desired. The atmosphere is clearly quite different today and will continue to be so in 2010.

The Bureau of the Census has a 2010 Planning Staff consisting of eight professionals. That number will be increased to the low 20's during the next year and will continue to expand thereafter. During the past two years the planning staff has accomplished the following:

- It has established the research and experimentation (REX) steering committee which has led to the decisions concerning research and testing built into the 2000 Census
- It has played a key role in establishing the CNSTAT 2010 Panel about which I am speaking to you at this moment.

- It is working with the American Community Survey staff on the decennial planning and testing program for 2001-2005. It is hoped that integration of results from the ACS with the 2010 Census will bring great efficiencies and cost savings. [More on this later.]
- It has established goals and objectives for the 2010 communications and consultation program.
- It has established the *New Millennium Speaker Series* and drafted the report, "One Nation, Many Communities," bringing to bear external advice and studying long-term trends with respect to the environment in which the 2010 Census will take place.

The stated objectives of the 2010 Planning Staff are to achieve a 2010 Census that will:

- Maintain or improve upon overall coverage.
- Reduce the differential undercount.
- Simplify census operations and smooth costs throughout the cycle.
- Increase the engagement of the citizenry in the census process.
- Keep costs from increasing without diminishing accuracy.

In his presentation at our first meeting, Jay Keller, the Assistant Division Chief of 2010 Census Planning, discussed the following strategies for attaining the goals that I have just enumerated:

1. Integrate 2010 Census development and design with the implementation of the American Community Survey.
2. Start the full-scale planning and testing program two years earlier than the present cycle. This implies that it is underway as we speak.
3. Continue re-engineering the data collection operation.
4. Develop, test, and analyze potential improvements to all other operational components of the census.

Some of you, as I did a few weeks ago, may be wondering, "What is this *American Community Survey (ACS)*, anyway?" Chip Alexander, who has been heavily involved in all aspects of its development, gave us a briefing. It is quite a revolutionary idea. Basically, it will be a rolling sample that will survey about three percent or three million households a year (250,000 hh./month). Thus over five years it will have covered about 15 percent of all households, comparable to the present one-in-six sample for the long form. It

is designed to be a mail survey with 100 percent follow-up, and then a field follow-up of a sample of nonrespondents. The present plan is to achieve 60 percent by mail, and 25 percent of the remaining 40 percent by telephone, bringing the response up to 70 percent. Then the remaining 30 percent will be sampled at a rate of 1 in 3, hopefully achieving an overall response rate of 80 percent. The sample will not be clustered but there will be some oversampling of small governmental units. The plan has already been tried out in some demonstration sites, and 31 comparison sites are planned for 1999-2001, followed by full implementation in 2003. Assuming that all goes well with the ACS, the result may be the *elimination of the long form in the 2010 Census*. Wow!

This is really fascinating stuff, but it is impossible to go into much of the detail now. Suffice it to say that a successful ACS will bring improvements to the 2010 Census through the creation of a better address list via the new Community Address Updating System (CAUS); it will produce a better understanding of local communities and census support systems; there are also several implications for the possible use of administrative records. Just to pique your interest, however, as mine was, I give you the following quote from Chip's briefing to us:

"We are running those counties [he is referring to the 31 demonstration sites] at a 5 percent sample per year. If you look at the three years, 1999 through 2001--and we hope to get those data out in mid-2002--you will have approximately a 15 percent sample in those 31 sites to compare to the census long form, which is about a 17 percent sample."

"So the idea is that during that comparison period, we will be going down to the individual census tracts in these sites, comparing the ACS three-year average against the census for those. We will be working with people from the sites. One of the sites will be in an area that Joe Salvo is familiar with, the Bronx, in New York. Working with people like Joe who are familiar with the area, we will look at the differences and try to explain what the differences are. This will probably be our most important thing in terms of our understanding of what the ACS number means compared to what a census number means."

It is clear that the ACS alone will give our panel plenty to chew on before the end of our tenure in 2003.

Returning to Jay Keller's discussion of the 2010 Planning Staff, this is what he had to say about expectations from our panel:

"We would like you to propose critical research. We will, funding permitting, be starting a fairly large program of research and small-scale testing in 2001. What do we need to know? Are we looking at the data-collection issues in the right way? Do we have a complete list of what the possible refinements are? Are we missing anything on that list?"

"Are we going to have the right data in place, as you do

your work and we do our work together over the next few years, to be able to assess potential options for 2010? Or are there gaps, and how soon can we identify those gaps and make sure that we do not lose data that we might be able to preserve as we go through Census 2000?"

"How will we decide that administrative records are adequate for profiling nonresponding households or for other purposes? What standards do we need to set up for that? What do we have to do to develop those standards and then get the information to determine whether or not we have met the standards?"

"Finally, are we on the right testing schedule? Is it too slow, too fast? We would like your ideas about that."

The final report of the *Panel on Alternative Census Methodologies*, discussed today by Keith, lists the research projects and experiments as they were specified at the time that it was written. Here is the final list decided upon by the Bureau, copied exactly from the document that was distributed at our June meeting:

**In Census 2000 the Census Bureau will conduct four experiments to inform future planning:**

**Census 2000 Alternative Questionnaire (AQE2000) experiment** - The AQE2000 is designed to continue efforts to develop user friendly questionnaires that encourage response. This experiment examines a modified short form structure and presentation of residence rules, revised navigational instructions for the long form, and differently sequenced race and Hispanic origin questions.

**Administrative Records Census 2000 (AREX2000) experiment**- The AREX 2000 explores two designs for using administrative records as the primary means of data collection for the census. It also evaluates the potential for using administrative records to enumerate households that do not respond to the census. The results, including cost information, will help plan major components of future decennial censuses.

**Social Security Number, Privacy Attitudes, and Notification (SPAN) experiment** - The SPAN experiment examines public response to a request on the short form questionnaire for Social Security number information and two different notifications about the Census Bureau's proposed use of administrative records obtained from other government agencies. Effects on mail and item response will be assessed. A telephone survey will also be conducted - one before Census 2000 and another shortly after Census Day - to assess the public's attitudes on privacy and confidentiality issues.

**Response Mode and Incentive (RM&I) experiment** - The RM&I experiment provides alternative response modes for the decennial census. Respondents in the sample receive a calling card as an incentive that is activated for their use if they answer the census using the Internet; other households may respond using one of two telephone options (a live operator or a voice

recognition mode). The effect an incentive has on response will be examined as well as the impact an incentive may have on the census infrastructure in the future.

**In addition to conducting these experiments, the Census Bureau will undertake two other research activities as part of this program:**

**The use of an Employee Reliability Inventory (EU) file**, which measures interpersonal skills, will be reviewed to determine if it is a credible and feasible tool in selecting enumerators. Current plans are to use this selection aid for hiring enumerators for the nonresponse operation in selected areas.

**Error Modeling and Simulation Research** will be conducted to allow a comparison of the proposed alternative designs for the 2010 census. This research will look at accuracy of population counts and level of geography, total costs, feasibility, and value of the data.

I am not prepared today to give you much additional detail on these proposed projects and experiments, nor am I sure that all of the details on design and execution have been worked out by the Bureau itself. I merely enumerate them to show you what the 2000 Panel will be closely monitoring during the 2000 Census and what will be the target of the 2010 Panel's recommendations for analyzing the results. In closing, however, I do want to emphasize a feature of the plans for Census 2000 that has been a pet interest of mine ever since my participation on the *Panel on Decennial Census Methodology* prior to 1990. In those days as our panel thought about some of the proposed research, especially that aimed at better understanding and eventually improving the response process, we would try to visualize the analyses that one might do to answer some of the important questions, and we seemed to be constantly asking questions such as "How many callbacks were made to get this case or another?", "What was the average number of attempts before a proxy was used?"--you know, the kind of information that is necessary before you can really design a cost-effective system for sampling nonrespondents, or figure out how to allocate scarce resources to alternative approaches in gaining cooperation, and the like. Unfortunately, we learned that in the operations of the decennial census little detailed information of that type at the interviewer and household level is retained for possible future study. This finding led us in our 1985 report, "The Bicentennial Census," to make a brief recommendation:

**Recommendation 6.3.** We recommend that the Census Bureau keep machine-readable records on the follow-up history of individual households in the upcoming

pretests and for a sample of areas in the 1990 census, so that information for detailed analysis of the cost and error structures of conducting census follow-up operations on a sample basis will be available.

In citing this recommendation I am aware that sampling for follow-up is dead and no longer on the table (although I have not heard that anyone pounded a wooden stake through its heart), but it should be apparent that the retention of information that is proposed would be valuable for many other issues.

The recommendation was repeated and elaborated upon in the 1990 Panel's final report to the Bureau in 1988, "Priorities for the 1990 Census: Research, Evaluation, and Experimental (REX) Program", but somewhere along the line the Bureau decided that it could not fund the undertaking in 1990.

The proposal still lives on, however. In support of the 1988 recommendation, the *Panel on Alternative Census Methodologies*, as just reported by Keith, has strongly urged the Bureau to collect what is now referred to as *The Master Trace Sample*. Here is its Recommendation 5.1 as stated in the 1999 document, "Measuring a Changing Nation":

The panel recommends that a trace sample be collected in roughly 100 tracts throughout the United States and saved for research purposes. The trace sample would collect detailed process data on individual enumerations. In addition, similar information on integrated coverage measurement should be collected, on a sample basis if needed. It would be very useful if information could be collected, again on a sample basis, to support complete analysis of the census costs model, all aspects of the amount of duplication and the efforts to unduplicate, and information needed to support total error modeling of the 2000 census.

The Bureau of the Census has accepted this recommendation and says that it is going to collect a Master Trace Sample during the 2000 Census--to which I can only say, "Hallelujah!"

I am not sure that the Bureau has worked out the exact details of how it is going to obtain this trace sample nor do we know how successful the undertaking will be or how the data will be analyzed, but those are just some of the problems that the 2000 Panel and our 2010 Panel will have to deal with as time marches on.