

December 2009 COPAFS Meeting Report

Patricia C. Becker

The fourth and final COPAFS meeting of the year took place on Friday, December 4. A water leak had closed the BLS conference center, forcing the meeting into a second-floor small conference room. Turning to Ed Spar's introductory overview, there was little to update on the budget front, since most of the government (Census Bureau excepted) was still running on a Continuing Resolution.

The Bureau of Labor Statistics is considering a major redesign of the Consumer Expenditure Survey, a task not undertaken since 1982. COPAFS has been asked to host a workshop on the topic. The redesign process is expected to be lengthy.

As reported in the discussion at a previous meeting, the Census Bureau has been planning to release the components of census coverage -- i.e. persons missed and duplications -- only at the national and regional levels. Discussion at the recent 2010 Census Advisory Committee meeting (see Ken Hodges' report on the APDU website) has made the case for reporting these measures at the state level. Results are to be reported by the end of 2010, and COPAFS will conduct a workshop on the topic.

Another issue under discussion, arising at the 2010 Advisory Committee meeting, is the release date for the first ACS 5-year estimates. These data are currently scheduled to be released late in 2010, with weights based on the 2000 census and using 2000 geographic areas. There is concern with the inconsistencies between this data set and the 2010 census data which will be available only a few months later. Stay tuned for the results of this discussion.

Finally, the COPAFS website is being redesigned. It hasn't happened yet, but keep watching!

Data.gov Initiative Paul Bogg, OMB

On January 21, 2009, newly inaugurated President Obama sent a memorandum to the heads of executive departments and agencies. The subject was "Transparency and Open Government." "Information maintained by the Federal Government is a national asset," said the President. The idea is to disclose federal government information (or data) "rapidly in forms that the public can readily find and use." This led to the March 2009 launch of data.gov.

There are more than 24,000 "dot.gov" websites. A user often needs to understand the federal government's organizational structure in order to find data of interest. Further, data are not always downloadable. Data.gov's is designed, or intended, to make all of this easier, and to encourage innovative applications by enabling access to data in formats that can be analyzed. The system defines "raw data" as machine-readable structure datasets that can be used for multiple purposes, and uses a concept labeled "mashed-up" to describe combining data from different sources on the fly. It appears that Bugg and other statistical policy staff in OMB are not happy with these definitions, but haven't been able to make changes thus far. Eventually there will be federal register notices or opportunities for the public to comment via a notice on data.gov itself. Part of the plan, in the long run, is to include data from sources other than the federal government, such as states.

Nothing so far changes the current situation, in which individual agencies are responsible for producing data, guaranteeing confidentiality, and providing access mechanisms via the Internet. It does not appear that the OMB staff promoting data.gov (personnel outside of the Office of Statistical and

Science Policy) understand data in the same way that data users do. More details are available in the powerpoint presentation.

National Survey of Residential Care Facilities (NSRCF) Lauren Harris-Kojetin, NCHS

This survey has two major partners: the Department of Health and Human Services (NCHS and other agencies) and the Department of Veterans Affairs. The survey is needed because the aging population has increased the need for long-term care services, little known about these facilities, and no other survey covers them well. current survey provides the information.

The number of residential care facilities has grown rapidly, and now nearly equals the number of nursing homes nationwide. The difference between them is that the residential care facilities do not provide skilled nursing services. Many different terms are used to label them, including assisted living, congregate care, board and care, family care, and personal care. The residential care facilities provide room and board (at least two meals a day), help with personal care, and around the clock, on-site, supervision. Facilities included in the sampling frame must have at least four beds; facilities licensed to serve the mentally ill, or mentally retarded/developmentally disabled populations, are excluded.

The survey has two parts: a facility questionnaire and a resident questionnaire. The survey starts with telephone screening, and then a CAPI interview with the facility administrator. A sample of residents is then selected from within the facility, and the questionnaire is conducted with staff members knowledgeable about the selected residents, such as nurses aides or floor supervisors. The residents themselves are not interviewed. The overall goal is to complete interviews with 2,250 facilities and for 8,450 residents. The survey is scheduled to begin with training this coming March, with data collection to be completed by October and findings released in early 2012.

Much effort has gone into outreach activities. NCHS has worked closely with the Center for Excellence in Assisted Living (CEAL), an association of associations related to the industry. This work has provided insights into contact materials, and promotes participation. A joint letter sent in advance communicates industry support for the survey.

Who Creates Jobs? Small vs. Large vs. Young Ron Jarmin, Census Bureau

There is a long-standing debate about the extent to which small businesses are responsible for job creation. Ron Jarmin reported on research in this area, completed in collaboration with John Haltiwanger of the University of Maryland and Javier Meranda of the Census Bureau, and funded in part by the Kauffman Foundation.

There are two camps: some who think that most new jobs are created by small business, and others who say it's not so. Jarmin *et al's* work suggests that there is some truth on each side, and further that the age of the firm is a significant factor. Firm "births" -- which are almost by definition small businesses -- contribute strongly to job growth. Many young business don't make it, but when they do, they generate dynamic growth. Large firms, which are mostly older, also contribute to job growth. In fact, when controlled for age, there is actually a positive relationship between firm size and job growth. In other words, older small firms appear to be the least likely to grow new jobs.

Thus, in reality it's pretty complicated. It's not really possible to pick the winners and losers just using statistics. In conclusion, Jarmin suggested that we need a more nuanced view of small businesses

and their contribution to job creation. Age matters as well as size. Another issue is the quality of jobs-- we need to look beyond simple counting to the kinds of jobs that are being produced by younger firms, the kinds of workers in these jobs, and the long-term labor market outcomes.

Local Employment Dynamics: Synthetic Data for OnTheMap Version 4 **Jeremy Wu, Census Bureau**

OnTheMap is a "dynamic mapping and reporting tool" for the Census Bureau's Local Employment Dynamics (LED) data. In this presentation, Wu discussed the system, along with the integrated, synthetic data which underlie the product. OnTheMap debuted in 2006, with data for 14 states for the 2002-03 period. The product has grown continually. Currently, it covers 47 states and includes data from 2002 to 2008. Another new release next December will add 2009 data and may include additional states.

OnTheMap allows the user to select the geographic area of where workers live, or where they work, and report variables such as age, earnings, and cross-state flows. The base unit is the census block, and the product features innovative disclosure protection. We've tried it out looking at our own ZIP code (48034). Entering this into the box, and then clicking on it in the list that appears, brings up a map of the area. The *base map* tab allows the user to select the features shown on the map. The *analysis* tab allows selection of the desired data, and leads to the *results* button which shows the answers. To access the system, go to <http://lehmap4.did.census.gov/themap4/>, or click on *Local Employment Dynamics* in the Business and Industry section of the census home page (census.gov).

Turning to the data, Wu provided an overview of the history of sampling, which was first used by the Census Bureau in 1937 to measure unemployment, and made its debut in the decennial census in 1940. It has not been smooth sailing ever since. Computers have made possible release of a flood of administrative data, while surveys are increasingly troubled by lower response rates, higher labor costs, and confidentiality concerns. Can we both have access to microdata and maintain confidentiality protection?

The LED approach provides both with a complex design that involves record linkage, noise infusion, imputation, synthetic data modeling, and measures of goodness and quality. Censuses, surveys, and administrative records are combined together in the Longitudinal Employer-Household Dynamics Program (LEHD), producing data for LED, Local Transportation Dynamics, and possibly other systems. LED brings together data on the worker and the employer. Just how this is being done was not explained in any detail.

Wu described OnTheMap and the data behind it as the latest development in the evolution dating back to the first discussions of sampling in 1895. It took a very long time for sampling to be accepted; he hopes that this innovation will become accepted and widely used in a shorter time frame. "Is this," Wu asks, "the dawning of another new field of statistical knowledge?"

Conclusion

Slides for three of the four presentations are on the COPAFS web site, www.copafs.org. The next meeting is scheduled for March 5. We plan to attend.

APDU members wishing further information on the COPAFS meeting may contact Patty Becker pbecker@umich.edu or Joan Turek joan.turek@HHS.gov.