Big Changes at Economic Census Will Provide New Insights into US Economy

Every five years, the US Census Bureau conducts an economic census, providing official benchmark measures of American business and the economy. Despite being less well-known than the population census, the economic census—based on a representative sample of approximately 4 million businesses from about 400 industries—is vital to understanding how the nonagricultural sectors of the US economy are performing.

Statistics from the economic census are used by policymakers and trade and business associations, as well as individual business owners. They are inputs to key measures of the US economy such as the gross domestic product (GDP), product accounts (NIPAs) and the Producer Price Index (PPI). And like the population census, the economic census data are used to construct up-to-date frames for ongoing sample surveys.

The 2017 economic census leadership team endorsed a number of innovative updates, each introducing a new set of statistical and production challenges. For example, in 2017, data collection will be primarily via the web, instead of mailing out paper questionnaires (as done in all previous censuses). Standard unit response rates will be released for the first time with the 2017 economic census, as will imputation rates for key statistics. Variability estimates for selected sample-based statistics will be published for the first time, as well. These last two sets of measures provide unprecedented transparency on the data quality for this large and widely used program.

And beginning in 2017, the economic census will use the North American Product Classification System (NAPCS) to produce economy-wide product tabulations. NAPCS is a comprehensive, market- or demand-based, hierarchical classification system for products (goods and services) developed and used by Canada, Mexico, and the United States (see *www.census.gov/eos/www/napcs/more.html*). Using this new classification system helps standardize vital economic statistics that measure the growth of output, prices, productivity, and trade.

How does NAPCS affect data users? There is no short answer. In 2017, the economic census will collect data on more than 8,000 products. In some cases, products are sold in one or two industries. In others, the same product is sold in a variety of industries. Think, for example, of a

flu shot, which can be administered in a doctor's office (the services sector) or grocery store (retail trade sector), might be stored in a warehouse (wholesale trade sector) and will be shipped to different areas (transportation sector). Under the previous product collection system, that flu shot would have been classified using a different code in at least four sectors. Under NAPCS, it has one code.

Of course, just because the US Census Bureau wants to collect this information doesn't mean every business reports it. Respondents can report data from a long, prespecified list of potential products in a given industry—some lists contain more than 50 potential products—and can write in descriptions of other products not prespecified.

"We find a lot of businesses are easily able to report big numbers such as total sales, total payroll, payroll unemployment, and so on—numbers that they need themselves," says Katherine Thompson, methodology director of complex survey methods and analysis in the economic directorate of the US Census Bureau. "But when you get to product breakdowns—where we give businesses a huge questionnaire, lots of definitions and ask them to make sure that the total sale of all their products is the same as the total sales that they reported earlier—the response drops off," she says. "Ideally, we would like to get answers from the businesses themselves. However, we have all kinds of outside constraints such as timing (deadlines) and budget, and it is often much more economical to *not* burden a business that provided all of these 'big numbers,' but failed to provide products." So, the economic census conducted extensive research into how to account for missing data.

The other quality that makes economic data useful is precision. How confident are you that something is being precisely measured? Look at a gas station, says Thompson. Most of its sales, unsurprisingly, come from gasoline—and the scale of those sales enables a high degree of measurement precision. "But they also sell sundry items like candy bars. That's a very small percentage of their sales, and so the precision will be much poorer for that reason. We want to provide people with a sense of how precise all our data measurements are."

Figuring all this out was not easy. Simulations had to be run, problems simplified and then scaled back up in complexity to match the comprehensiveness of the full census. The result is an intellectual achievement with enormous practical use. "There are two stories here. The first is that this is an exciting development in the collection of high-profile economic data. The economic census has standardized the collection of inputs and outputs across North America," says Thompson. "And we are giving the people who use these data measures of how precise that data is. That's a big story. The second is how we got to this point through combining principled statistical analyses and methods to account for practical issues like missing data and computing time; in other words, how we got it done."

JSM Talk:
http://ww2.amstat.org/meetings/jsm/2018/onlineprogram/ActivityDetails.cfm?SessionID=215435

For details, contact: Katherine J. Thompson

Email: *Katherine.J.Thompson@census.gov*

Tel: (301) 763-4941

About JSM 2018

JSM 2018 is the largest gathering of statisticians and data scientists in the world, taking place July 28–August 2, 2018, in Vancouver. Occurring annually since 1974, JSM is a joint effort of the American Statistical Association, International Biometric Society (ENAR and WNAR), Institute of Mathematical Statistics, Statistical Society of Canada, International Chinese Statistical Association, International Indian Statistical Association, Korean International Statistical Society, International Society for Bayesian Analysis, Royal Statistical Society and International Statistical Institute. JSM activities include oral presentations, panel sessions, poster presentations, professional development courses, an exhibit hall, a career service, society and section business meetings, committee meetings, social activities and networking opportunities. http://ww2.amstat.org/meetings/jsm/2018/index.cfm

About the American Statistical Association

The ASA is the world's largest community of statisticians and the oldest continuously operating professional science society in the United States. Its members serve in industry, government and academia in more than 90 countries, advancing research and promoting sound statistical practice to inform public policy and improve human welfare. For additional information, please visit the ASA website at *www.amstat.org*.