
September 29, 2014

Katherine R. Smith,
Executive Director
Council of Professional Associations
on Federal Statistics
This cover of a Government Accountability Office (GAO) report is not exceptional. As a legislative branch watchdog of executive branch performance, it often identifies administrative problems. Inconsistent data is one such problem.
Federal Statistical System

• About 100 federal agencies collect data
• Thirteen designated (by OMB) as official statistical agencies
• Bureau of Labor Statistics, Census, and Bureau of Economic Analysis produce economic statistics
• But so do dozens of others
• There are also numerous sets of administrative data on econ. phenomena

13 officially designated agencies:
Economic Research Service (ERS/USDA)
National Agricultural Statistics Service (NASS)
Bureau of Economic Analysis (BEA)
Census Bureau
National Center for Education Statistics (NCES)
Energy Information Administration (EIA)
National Center for Health Statistics (NCHS)
Statistics Of Income Division of the Internal Revenue Service
Bureau of Justice Statistics (BJS)
Bureau of Labor Statistics (BLS)
National Center for Science and Engineering Statistics (NSF)
Office of Research, Evaluation, and Statistics, Social Security Administration
Bureau of Transportation Statistics (BTS)
Decentralization is an outgrowth of how statistics was funded very early in U.S. history.

Smokestacks are not just between agencies, but also within many agencies with multiple functions.

Although highly decentralized, there is a great deal of interdependence among agencies. BEA, specifically, relies on statistical input from a variety of agencies in order to construct its indicators. Thus, when BLS announced it would be unable to continue its export price index program, BEA was thinking it was “up the creek” since these are important to the National Income and Product Accounts.
Census money income is defined as income received on a regular basis (exclusive of certain money receipts such as capital gains) before payments for personal income taxes, social security, union dues, Medicare deductions, etc. The Census Bureau determined that respondents report income earned from wages or salaries much better than other sources of income.

Each of these surveys differs from the others in some ways, such as the length and detail of its questionnaire, the number of households included (sample size), and the methodology used to collect and process the data. The Small Area Income and Poverty Estimates program actually uses statistical models to produce income and poverty estimates by combining survey results with administrative records. Thus it is not surprising that their results differ somewhat.

But, Census also derives alternative income measures that systematically remove or add various income components such as deducting payroll taxes and federal and state income taxes and including the value of specific noncash benefits, food stamps, school lunches, housing subsidies, health insurance programs, and return on home equity.

Says Census itself: “As a result of this multiplicity of sources, it is important to understand that different surveys and methods, which are designed to meet different needs, also produce different results.” Census “guidance” directs users to the descriptions of each survey and its methodology – Not a user friendly approach.
But there are many other statistical estimates of “Income.” For example:

IRS income is from tax returns – It is very different from Census data, especially for county income.

USDA conducts a special survey of farm households.

BEA personal income, estimated largely from administrative sources, is the income received by persons from participation in production, from government and business transfer payments, and from government interest.

If one goes to Data.Gov...they are shown a total of 1,853 datasets on income of some sort or another.
How Wealthy are Americans?

- Survey of Income and Program Participation (SIPP)
- Federal Reserve Board Survey of Consumer Finance
- Federal Reserve Board Flow of Funds Data
- Panel Study of Income Dynamics (PSID)
- Bureau of Economic Analysis (BEA)
- Internal Revenue Service (IRS)
- Economic Research Service farm household wealth

There may be more and more distinctly different measures of wealth across the statistical system and other national surveys than for income.
Health insurance coverage: Important as a gauge of the initial effect of the ACA. Three different federal agencies estimate. Earlier this month (September 2014) Census released estimates derived from ACS, and NCHS released estimates based on its National Health Information Survey. It was planned that the releases would occur on the same day, an idea hailed as representing better coordination between the agencies. But neither release referred to the other or explained why (for good reasons) they were different.

Office of Immigration Statistics (OIS) data come from administrative data on citizenship, visas, border enforcement actions, deportations, etc., while Census figures on the “foreign-born” U.S. population are survey data. Neither tells the whole story, but “unauthorized” immigration is an arithmetic function of the two.
Differences in Business Classification

- 30% of U.S. establishments are classified as belonging to different industries by BLS and Census
- Census uses IRS tax filing data to develop its establishment list. BLS has access to neither IRS lists, nor the Census list developed with IRS input.

Also want to note here as we introduce business statistics, a different kind of taxonomic problem that adds to the confusion.

Definitions from Census:
- Establishment - A single physical location where business is conducted or where services or industrial operations are performed.
- Enterprise - An enterprise is a business organization consisting of one or more domestic establishments that were specified under common ownership or control. The enterprise and the establishment are the same for single-establishment firms. Each multi-establishment company forms one enterprise - the enterprise employment and annual payroll are summed from the associated establishments.
- Firm - A firm is a business organization consisting of one or more domestic establishments in the same state and industry that were specified under common ownership or control. The firm and the establishment are the same for single-establishment firms. For each multi-establishment firm, establishments in the same industry within a state will be counted as one firm- the firm employment and annual payroll are summed from the associated establishments.

Does not make logical sense: (Some B=A, Some C=A, but no B = C). Furthermore, all are referred to as businesses throughout Census, BLS and BEA reports.
Data on jobs and employment are available from different sources, principally the monthly CPS and the monthly Current Employment Statistics payroll survey. However, the measures can vary because the CPS surveys households that report the number of employed people in the household, while the payroll survey counts the number of jobs reported by business establishments.

Second, estimates for a similarly defined phenomenon may differ due to varying methods of data collection, including type of data collection (survey information or administrative records), as well as differences in question wording, unit of analysis, mode of data capture, or target population.

Also the fact that many survey based economic indicators are released on a schedule and are later revised as additional information comes in, making the estimate better, is VERY CONFUSING to casual users and media.
We rely on the agencies themselves, information specialists, and the media, to translate a statistical release for the general public and be able to compare its statistical information to other statistics on the same subject. A TOUGH JOB.

“Perhaps the largest risk associated with competing estimates is the impression of redundancy—the perception that having multiple data series cover the same information may be wasteful. In congressional battles over budget resources, perceived redundancy may lead to discussions of trade-offs.” - Dan Cork, CNSTAT
For example, picking one official estimate, or at least the denominator therefor, is done by OMB’s Statistical Policy Office by defining metropolitan statistical regions. OMB also gives guidance on the official measurement of poverty, and weighs in on aspects of race and ethnicity.

Concurrent releases – Best if joint rather than concurrent. Very good IF ACCOMPANIED by EXPLANATION of differences.
BLS has initiated an ambitious project to put its own data house in order – taxonomically and lexographically.

Bureau of Labor Statistics
Taxonomy/Lexicon Initiative

Mission

“Develop a BLS-wide, consistent taxonomy (hierarchy of categories or concepts) and lexicon (word list with synonyms) of BLS concepts in plain English geared towards the lay public, to be used in data query tools, document tagging, website search and navigation, and records management.”
• Acknowledge that federal estimates of like economic phenomena can conflict and help others capitalize on the differentiation.
• Insist on clarity from each source of official statistics. If an official web site does not explain a variable or uses jargon that is not comprehensible, ask questions!
• Seek reasons that explain differences in reported economic phenomena.
• Improve the value of information conveyed by illuminating the reasons for seeming discrepancies that may actually be measures of different things.
• Give feedback to agencies to help them identify specific needs for naming conventions, common definitions and compatible modes of data presentation,
Thank You!

Katherine Smith
kitty.smith@copafs.org
703-836-0404

Visit us on Twitter @COPAFS