

THE

WITH

TROUBLE

DATA

**Declining response rates weaken government
surveys and the policies they help inform**

By SHAWN ZELLER





it does each year like clockwork, the Census Bureau released a batch of data Sept. 10 about Americans' well-being. It showed the median income holding steady at \$63,179. The percentage of people living in poverty dropped half a percentage point to 11.8 percent. The share of those without health insurance rose to 8.5 percent from 7.9 percent a year earlier.

News outlets treated the data as gospel, focusing mainly on the rise in the uninsured rate and blaming it on President Donald Trump's lack of interest in promoting the government-subsidized policies created by the 2010 health care law. Interest groups, too, latched onto the numbers, bemoaning stubbornly high poverty numbers for children, for example, while others touted employment statistics that showed most Americans who want jobs have them, using the figures as a case for Trump's tax cuts and his penchant for deregulation.

It's a given, on both sides of the political aisle, that figures like those released by the Census Bureau drive public policy.

But policymakers should take the good news, and the bad, with a dose of skepticism. Deep within its report — the Census Bureau's Current Population Survey and the annual supplement on which the latest data were based — the government surveyors reveal an unsettling truth: The combined non-response rate was 32.4 percent. It has doubled in the past eight years.

That, though, is only part of the problem. In addition to those who ignored the survey completely were those who filled out only part of it. In those cases, the bureau will, essentially, make an educated guess. For some questions, it had to do that in nearly three-quarters of cases.

It's not only the Current Population Survey that Americans are ignoring. Across government, surveyors are having trouble reaching people like they used to. And they are, as a result, gathering questionable data on topics important to policymakers, from the number of people using opioids to the number refusing to vaccinate their children.

It's a function of Americans' migration to electronic communication tools, like text messages and email, and away from the telephone. Once surveyors' reliable tool, Americans' phones are now besieged by robocall scams and marketing pitches, some presented as official surveys. Armed with caller ID and spam blockers, fewer people even bother to answer when an unfamiliar number comes up.

When response rates drop, surveyors worry about bias creeping into their results. Are some Americans — the old, white and well-to-do — more likely to respond to phone calls than others?

"There's a huge credibility issue there," says Don A. Dillman, deputy director for research at Washington State University's Social and Economic Sciences Research Center. "If a bunch of people don't respond, it's a huge problem."

That's because surveys, and the data they produce, are crucial guides to policymakers. "Regardless of what side of an issue one is on, you have to deal with an established set of facts," Dillman adds.

Four years ago, in a paper for the National Bureau of Economic Research, University of Chicago public policy professor Bruce D. Meyer and two colleagues warned that the declining response rates represented a "cri-

sis,” as they would drive down trust in government data and make the data less valuable to policymakers at a time when their desire to measure the results of government programs is keen.

“Nothing has happened to alleviate my concerns” in the intervening years, Meyer says.

Meyer is now advising the Census Bureau, which is trying to shore up its results in advance of the 2020 decennial count. He’s focused on figuring out how to improve the government’s assessment of Americans’ incomes, which are thought to be badly under-reported to surveyors.

“We have a lot more evidence that things are worse.” He adds: “Congress should care immensely.”

Meyer, like many surveyors, views the replacement of some survey-collected data with administrative data — for example from the Internal Revenue Service on the taxes it collects or the Social Security Administration and the Agriculture Department on the benefits programs they manage — as the best answer to the problem.

But the Census Bureau is, at best, dipping its toe into a changed methodology, planning to use administrative data next year to confirm that homes that haven’t responded to its survey are vacant.

It is reluctant to move more quickly for several reasons. Administrative datasets often don’t line up neatly with survey questions. Any significant change in methodology will blow up long-established trend lines and ripple outward, affecting other data trends. And there is simply bureaucratic reluctance: The agency’s employees grew up on survey analysis and are loath to abandon it.

Money, of course, is an impediment to change too. The shift is bound to be expensive, says Meyer: “It’s really a question of short-run investment that would have a long-term payoff.”

Congress, as yet, isn’t willing to make that investment. But it is growing more attuned to the problem.

In 2016, it passed legislation to create a blue-ribbon panel aimed at bolstering the data on which policymakers rely. The Commission on Evidence-Based Policymaking issued recommendations 18 months later, focused on facilitating a shift away from surveys to administrative data.



George Frey/Getty Images

NOT IMMUNE: The response rate to an immunization survey conducted by the CDC has plummeted 17 percentage points in the last two decades. Above, a nurse administers a measles, mumps and rubella virus vaccine to a patient in Provo, Utah, in April.

Playing Catch-Up

Ironically, as corporations run roughshod over privacy concerns and use “big data” ever more effectively to advertise to consumers, the government must cope with privacy laws — such as those barring the sharing of health and tax information — that impede using data collected for one purpose for another.

“We are better at getting information on people when they don’t expect it, but worse when they do,” laments Ben Harris, chief economist with Results for America, a group that tries to help government agencies harness the power of data. He spoke in April at an event at the conservative American Enterprise Institute think tank where panelists tried to answer the question: “Are survey data still useful?”

To assuage privacy concerns, the Commission on Evidence-Based Policymaking’s report recommends creating a National Secure Data Service into which government data would flow. It would be open to researchers for statistical purposes. States would be required to provide data that they now may withhold.

The commissioners recommended that federal agencies designate a senior official to oversee the use of data in guiding policymaking and that the White House Office of Management and Budget facilitate the sharing of data between agencies.

Then-Speaker Paul D. Ryan, the Wisconsin Republican, sponsored a bill that became law (P.L. 115-435) to act on some of the recommendations. It aims mainly to inventory the government’s datasets and to identify obstacles to information-sharing with the goal of making much more data available between agencies.

That could provide useful insights. The Social Security Administration, for example, might glean a lot from data held by the Agriculture and Housing and Urban Development departments about the housing and nutrition of the elderly. HUD officials, meanwhile, might care to know about the health of people living in government-subsidized housing.

Congress passed the bill in December and Trump signed it in January.

It was a start, but it leaves some of the hardest questions about shifting away from surveys to administrative data for later.

And it left major obstacles to information-sharing in place. For example, the Internal Revenue Service cannot, by law, share certain data with other agencies, except for the purpose of improving tax administration.

“If the government were a business, it’s like saying the marketing department can’t share information with shipping and sales,” says Meyer. “It means there is a lot of inefficiency and duplication and an inability to improve government programs.”

Public Health Threats

In the meantime, government surveyors are responding by changing their methods. They are relying more on the mail and in-person visits. But that’s expensive and, they fear, may only be further biasing results. It’s hard to know if those they convince to participate in a survey after a second or third attempt further diversify the sample, or further bias it.

The Current Population Survey, despite its travails, is still considered one of the best at garnering a response. The Census Bureau contacts a sample of households by mail, sends a surveyor to the house and later calls to gather data. Households are queried monthly for four months, then removed from the survey sample for eight, then surveyed for four more months. The method is designed to provide continuity in the sample, while not overburdening households.

A 2016 report by John L. Czajka and Amy Beyler of Mathematica, a contractor that works with government surveyors, found that the drop in response rates is universal, affecting surveys crucial to current policy debates,

such as the National Immunization Surveys, the source of government data on how many Americans are opting out of vaccines against infectious diseases like measles, or the National Survey on Drug Use and Health, the source of data on how many Americans are using opioids.

Both have seen their response rates drop fast. The most recent drug use survey, conducted in 2018 by the Substance Abuse and Mental Health Services Administration, reported a response rate of 73.3 percent, a drop of nearly 17 percentage points from 20 years ago. The immunization surveys, conducted

by the Centers for Disease Control and Prevention, is doubly afflicted because it involves first contacting parents, who are asked to allow surveyors to contact their child’s doctor, who is then called and asked to provide immunization records. The parents’ response rate dropped in 2017 to 76.1 percent from 93.5 percent in 1995. But barely half of doctors provided the records.

Filling Some Gaps

The problem facing surveyors is reminiscent of the one afflicting political pollsters. The pollsters, it’s fair to say, face even more pressing questions about their results.

Most pollsters predicted Hillary Clinton would win 2016’s presidential election, for example, though the results, in the end, fell within the margin of error of most of the major polls.

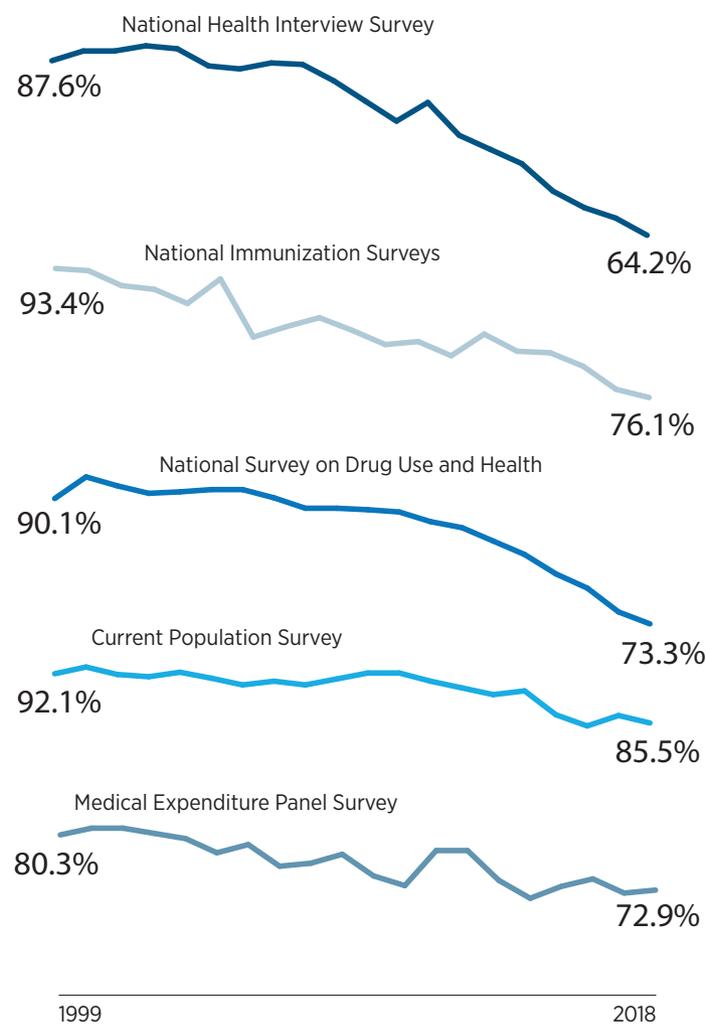
There have been other, less high-profile, but clearer mistakes.

In 2014, for instance, Republican Ed Gillespie overcame a double-digit deficit in the polls to finish within a percentage point of Democratic incumbent Sen. Mark Warner in Virginia. In 2015, only one of 13 publicly released polls showed Republican Matt Bevin leading Democrat Jack Conway in the Kentucky governor’s race, and just barely. The rest showed Conway winning by margins ranging from 2 to 11 points. In the end, Bevin defeated Conway by 9 percentage points.

Pollsters, without the resources the government has to follow up with respondents, now have response rates under 10 percent. As a result, they must adjust their results substantially, weighting more heavily the answers of hard-to-reach demographic groups, like young people and minorities.

Response rates drop

Fewer Americans are answering surveyors’ questions, raising concerns about the data’s reliability.



Source: Mathematica
Jason Mann/CQ Roll Call

The government's surveyors do something similar. They "impute" answers for people who don't fill out their surveys in whole or in part.

"We have very good statistical tools to adjust data," says Washington State University's Dillman. "The question is, do we know what to put weight on? We don't know what variables to adjust from. That's worrisome."

Indeed, there are limits to what surveyors can do. They've found, for example, that Americans — especially at the lower and upper ends of the income scale — do a poor job of estimating how much they're making.

The surveyors will impute an amount based on the earnings of similarly situated people, but only if a respondent reveals that they received some funds from a particular source, whether that be an Earned Income Tax Credit, dividends on stocks or rental property income. If the respondent doesn't acknowledge the income source, the surveyor can't impute a figure. Surveyors believe the government's estimate of median household income is probably a fairly solid number, but not so its figures on the poorest or richest Americans, or all the implications that stem from those, about the wealth gap between the races, or income disparities more generally.

No Good Solutions

Whether or not government surveyors are facing a crisis now, no one expects response rates to recover. It's notable that a polling organization as respectable as the Pew Research Center, facing a response rate of just 6 percent, announced earlier this year that it was mostly abandoning its use of the phone. Instead, it is using a panel of 13,000 Americans it has recruited to be representative of the country. They are paid \$5 to \$20 to fill out the organization's polls.

Courtney Kennedy, Pew's director of survey research, says it will probably be possible for phone pollsters to do "reasonably high-quality surveys for at least a few more years," but, she adds, "We don't see a long-term horizon."

That reality is at the root of the Census Bureau's move to consider administrative data.

In a release last November, Jennifer M. Ortman, the assistant division chief for survey methods and measures in the bureau's American Community Survey Office, explained that the bureau was looking, in particular, at replacing questions about housing and income with data from administrative sources. She said that the census was looking at administrative records to better identify

people by race, age, and ethnicity, and to better calculate numbers of young children.

But a September 2018 presentation by Andrew Keller, a mathematical statistician in the Census Bureau's Decennial Statistical Studies Division, at a workshop on improving the bureau's data gathering, revealed the challenge in that.

Keller looked at data from the 2010 census in which surveyors had not been able to get information on race from households and were forced to either ask neighbors or to impute a race. Keller replaced those figures with administrative data and found substantial differences.

When Keller used administrative records in place of imputed ones, the percentage of whites grew 13.5 percent, blacks 9 percent and Asians 18.4 percent. The decline came out of those imputed to be of some other race, or of mixed race, a reflection of the greater flexibility the census provides to respondents in defining their race.

The bureau declined to make Ortman or Keller available for interviews, saying that Meyer would speak for it. But it appears it will only go ahead with a very limited use of administrative data for the 2020 census, primarily using government information on va-



DRUG DATA: Former House Speaker Paul D. Ryan discusses the opioid crisis during a 2018 news conference.

Mark Wilson/Getty Images

cant housing units to spare census canvassers from having to visit them in person, a huge driver in the census' growing cost.

In placing the 2020 census on its "high risk list," the Government Accountability Office, the investigative arm that helps Congress oversee federal agencies, noted in March that the 2020 census is expected to cost \$15.6 billion, with the average cost for counting a housing unit expected to be higher than the \$92 it cost in 2010. By way of comparison, it cost \$16 in 1970, with both figures measured in 2020 dollars.

The bureau won't do more with administrative data this time, though, because it has found problems with it that the bureau views, as yet, to be insurmountable.

Some data, for instance, are owned by the states, which need not turn them over. And states' skepticism of information-sharing is only getting worse, given the president's polarizing approach to governing and his effort, blocked by the Supreme Court, to ask about citizenship status in the 2020 census.

States are concerned their data "are going to be weaponized against unauthorized immigrants," says Warren A. Brown, research director of the Cornell Federal Statistical Research Data Center. "For this to work, the states and the federal government have to trust each other and come to agreement to put their data together. But states are very suspicious of the federal government depending on who is in the White House and who is in the governor's office. You can't have a data system that waxes and wanes based on the whims of the political currents."

Data Determines Dollars

Even if the Census Bureau were to have access to all the available administrative records, it wouldn't be able to move yet. The administrative data are collected in a way that doesn't square with the questions traditionally asked by surveyors. For example, the IRS has income data on a calendar year basis. The Census Bureau asks for income over the prior 12 months.

And there are concerns about what would be lost if the government drops some survey questions. Replacing survey data with administrative data would deny policymakers information that is useful, says Michael Davern, executive vice president of research at NORC, a University of Chicago-based re-

"If a bunch of people don't respond, it's a huge problem."

— Don A. Dillman, deputy director for research at Washington State University's Social and Economic Sciences Research Center

search organization that conducts surveys for government agencies.

Davern, who studies Medicaid, the state-federal health care program for the poor, says that government surveys routinely undercount the Medicaid population by 30 to 40 percent, compared to the numbers collected administratively.

But the size of the Medicaid undercount is important information for policymakers. Some people on the program may be ashamed to admit they are on public benefits, but some may simply not know that they are.

"People may become enrolled in ways that may not be transparent to them," Davern says. "They may fill out paperwork at a hospital when they are there, and the hospital enrolls them and they may not know it."

This is a problem for policymakers who want to encourage people to seek preventative care, as well as those who want to root out waste in the system. People who don't know they have insurance are reluctant to seek treatment when they are sick, or to get immunized or checked for diseases. States may be paying fees to managed care companies to provide care to people on Medicaid. That money is wasted if people don't use it.

Before the government shifts away from surveys, it should consider conducting smaller, more focused ones, says Robert A. Moffitt, an economist at Johns Hopkins University.

An April article in the *Southern Economic Journal* by Moffitt and a colleague, Kyung Min Kang, looked at data gathered about food stamp recipients by the Agriculture Department's FoodAPS National Household Food

Acquisition and Purchase Survey. In 2012 and 2013, the department's surveyors approached Americans at their homes and interviewed them there. The sample was a relatively small, but nationally representative group of 4,826, and the Johns Hopkins researchers found the survey provided more accurate data than the broader Census Bureau surveys.

The Agriculture Department's approach was expensive but, says Moffitt, it offers benefits over moving to administrative data. "The advantage of surveys is you get to know about family size, where they live, if they are married or not, if they have a grandmother there, how they are using their benefits. You don't get that from administrative data."

And, of course, government surveys not only drive public policy decisions, but also the distribution of federal grants to states and localities. Any effort to rethink them, therefore, could create winners and losers. It's no wonder the Census Bureau is approaching the need to change its practices with caution.

Still, at a September 2018 forum convened by the bureau to discuss ideas for improving its American Community Survey, which collects demographic data on an ongoing basis, Constance Citro, a senior scholar at the Committee on National Statistics, an arm of the National Academies of Sciences, Engineering and Medicine, called out the bureau for moving too slowly. She said it needed to demonstrate some "wins" in improving its processes by committing now to replacing some of its survey questions with administrative data.

One of the reasons the Census Bureau continues to do better in getting Americans to respond than commercial surveys is that its key products, the decennial census and the American Community Survey, have the force of law behind them. They are mandatory. Americans who don't respond to them are, technically speaking, breaking the law.

During the Republican reign in the House, just a few years ago, there was an unsuccessful push to eliminate the threat of fines for those who don't comply, on the grounds that the government's questions are too intrusive.

The Census Bureau believes its response rate would drop by 20 percentage points if its surveys were voluntary. The threat from the GOP has subsided with Democrats now in control of the House, but Citro argued that the bureau must factor in the politics. Congress was still watching, she said. ■