

THE WALL STREET JOURNAL.

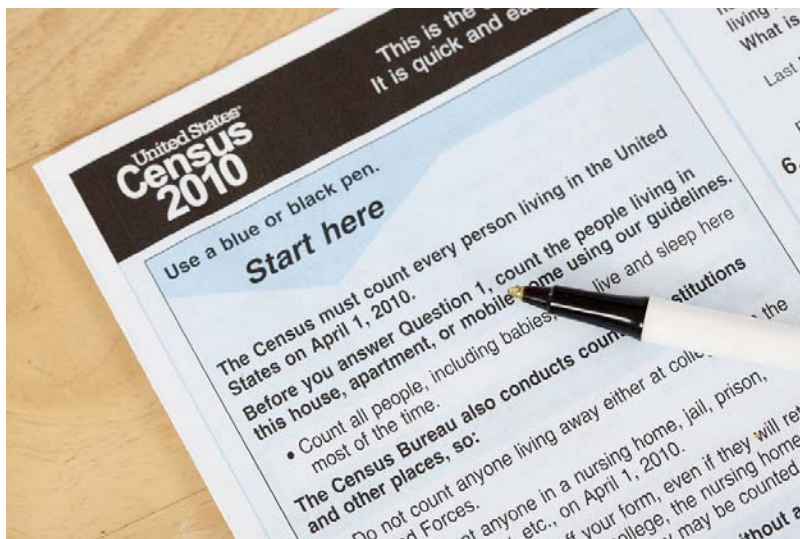
This copy is for your personal, non-commercial use only. To order presentation-ready copies for distribution to your colleagues, clients or customers visit <http://www.djreprints.com>.

<http://www.wsj.com/articles/how-uncle-sam-conducts-surveys-1461334374>

U.S. | THE NUMBERS

How Uncle Sam Conducts Surveys

Complex sampling gives U.S. representative data on big issues at less cost than simple random sampling



The Census Bureau uses a variety of sampling techniques to conduct its many different surveys. PHOTO: GETTY IMAGES

By JO CRAVEN MCGINTY

Updated April 22, 2016 10:15 a.m. ET

We need to talk.

When that message comes from Uncle Sam, it's often delivered in the form of a survey. The federal government amasses an astounding amount of data each year from nationally representative surveys on topics ranging from public

health to consumer spending.

Part of the challenge is to collect the information—much of it gathered in expensive face-to-face interviews—as efficiently as possible. That means a simple random sample, the gold standard of scientific surveys, is often out of the question.

A random sample is a reliable way to ensure the survey responses of a few people accurately represent an entire population. It works because every person has an equal chance of being selected, and the attitudes and practices of the sample will resemble those of the whole group.

But in a country the size of the U.S., it's too costly to send survey teams ping-ponging across the nation to conduct in-person interviews in randomly selected households that may be separated by hundreds of miles.

The solution is to use a complex sample that affords the benefits of a simple random sample but minimizes travel and the attendant costs.

“You want to design the most cost-efficient sample that gives you the most representative data,” said Brady T. West, who teaches survey methodology at the University of Michigan.

Complex samples are selected in stages. First, the country is divided into strata, often by geographical regions, to increase representation. The strata are subdivided into what survey designers call clusters including counties or cities and towns or some other geographical unit. Those units are randomly sampled, and then survey participants are randomly selected from only those areas.

Instead of selecting participants from across the country, they are pulled from, say, 20 or 30 cities and towns—drivable expanses that are easily covered by survey teams.

“It's less efficient statistically than spreading the sample out more, but you save a large amount of money,” said Colm O'Muircheartaigh, a senior fellow at the National Opinion Research Center who is an expert in survey design. “It's cheaper to go to a set of places and do work in each place.”

But there is a trade-off.

“Because elements tend to be alike within clusters, you lose precision by concentrating the sample in a limited number of clusters,” Mr. O'Muircheartaigh

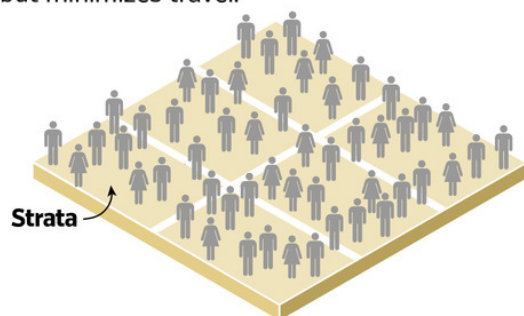
said. “The decision as to how much you cluster is based on the trade-off between cost and precision.”

Sample for Example

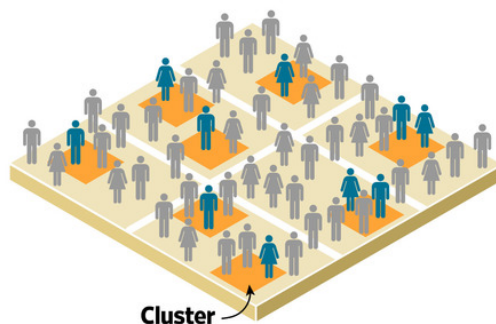
In a country the size of the U.S., it's too costly to send survey teams to conduct face-to-face interviews in randomly selected households that may be separated by hundreds of miles. The solution is to use a complex sample that affords the benefits of a simple random sample but minimizes travel.

Complex samples are selected in stages.

1. The country is divided into 'strata' to increase representation.



2. Strata are divided into geographical units, which might include counties or cities and towns or some other geographic unit. These are called clusters, and they are randomly sampled.



3. Households or individuals are randomly sampled from within the selected clusters. These are the people who ultimately are interviewed in the survey.



Source: survey experts

THE WALL STREET JOURNAL.

In-person surveys, which produce some of the richest and most complete data available, are conducted when collecting information by telephone or mail isn't feasible. That could be because the survey includes a physical examination or the information sought is particularly sensitive or the survey is too lengthy and complex for a telephone call or mailed questionnaire.

The National Health and Nutrition Examination Survey, known as NHANES for short, is an example.

Each year, the Centers for Disease Control and Prevention sends tractor trailers loaded with medical equipment to 15 counties where mobile survey teams conduct around 5,000 interviews and medical and dental exams.

Altogether, there are 30 field interviewers, two physical examination teams of 16 people each, and three field-office teams of four to five people. At any given time, home interviews are being conducted in three counties and physical exams are being conducted in two counties, according to a representative for the National Center for Health Statistics.

The data reveal the prevalence of major diseases, provide the basis for national standards of height, weight and blood pressure, and guide the development of

public health policies and programs.

Child growth charts are one product of this survey. It also has revealed the prevalence of undiagnosed diabetes, provided a means to track obesity and, thanks to the data accumulated from blood tests, led to the elimination of lead from gasoline.

The survey costs \$30 million to \$35 million annually. If it were conducted with a simple random sample, the mobile teams would potentially have to travel to 5,000 locations to assemble comparable data at a greater cost.

Another example of a government survey that uses complex samples is the National Crime Victimization Survey, which interviews 90,000 households each year to assemble nationally representative data on personal and property crime. The primary source of information on unreported crimes, it's used to estimate the likelihood of victimization by rape, robbery, assault, burglary and other crimes for the whole country as well as segments of the population such as the elderly or members of various racial and ethnic groups.

MORE NUMBERS

- Like numbers? Check out past columns (<http://www.wsj.com/news/types/the-numbers>)
- Join the conversation at The Numbers blog (<http://blogs.wsj.com/numbers/>)
- Follow The Numbers on Twitter (<https://twitter.com/WSJnumbers>)

There's also the Consumer Expenditure Survey, which polls 7,000 households in

91 areas of the country to capture every dollar spent by the survey participants during the year, according to Taylor J. Wilson, an economist for the Bureau of Labor Statistics who works on the survey.

The survey provides the market basket of consumer goods and services used by the consumer-price index, which is the most widely followed monthly indicator of inflation and the basis for adjusting things like Social Security payments.

“There is a herculean amount of persistence required to collect this information, and it's an incredible, terrifically rich resource for any researcher who's interested in understanding American household spending patterns,” said Anthony Damico, an independent consultant who analyzes a variety of government surveys.

So, if Uncle Sam asks for some face time, the request might sound simple. But the

method is complex.

Write to Jo Craven McGinty at Jo.McGinty@wsj.com

Copyright 2014 Dow Jones & Company, Inc. All Rights Reserved

This copy is for your personal, non-commercial use only. Distribution and use of this material are governed by our Subscriber Agreement and by copyright law. For non-personal use or to order multiple copies, please contact Dow Jones Reprints at 1-800-843-0008 or visit www.djreprints.com.