DEMOGRAPHY

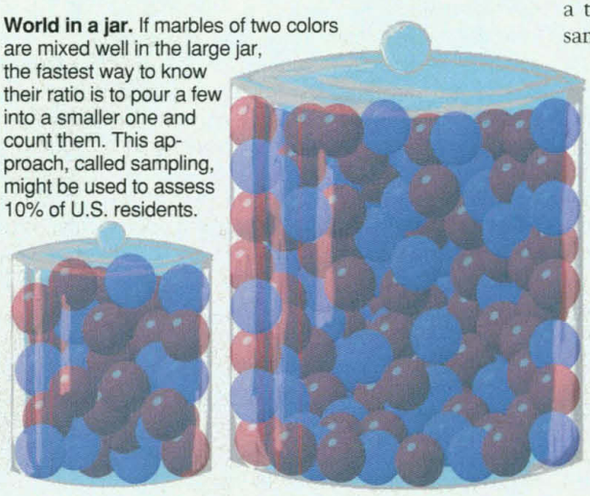
Census 2000: Where Science And Politics Count Equally

Just 2 years from now, the United States will undertake the biggest social science exercise in history. An army of professionals and volunteers will try to paint a statistical portrait of the entire U.S. population, producing a wealth of data that will influence politics, economics, and social research for decades to come. But even at this late stage, the U.S. Bureau of the Census, which will conduct the 2000 survey, doesn't know how it will carry out critical parts of the exercise. It is locked in a high-stakes political battle with Congress over a plan to reduce uncertainties through statistical techniques rather than direct head counts. The Census chiefs are in the uncomfortable position of a general on the eve of a major campaign who doesn't know what weapons he will be able to deploy.

The crux of the dispute is how to account for those Americans who statisticians say are too costly to track down or might elude a head count entirely. The Census Bureau—backed by the American Statistical Association (ASA) and a panel of the National Academy of Sciences—has recommended that the most accurate and cost-effective way is to directly count about 90% of U.S. residents and assess the numbers and demographics of the rest by extrapolating from a sample of them. "Who has to eat the whole pie to know what it tastes like?" asks Census statistician Tommy White. Critics, led by several key Republican members of Congress, contend, however, that 10% of a pie may not in fact taste like the rest, and they have tried to block the bureau from using sampling. This battle has claimed its first victim: Census director Martha Farnsworth Riche, who resigned her post last week. "It's a mess," says Margo Anderson, a University of Wisconsin historian and author of an upcoming book on the census. "There is no easy resolution in sight."

The level of passion aroused by what may seem like a dry statistical argument is not surprising. The Census count will determine how House seats are divvied up among the states—and that final 10% might shift the balance between urban and rural areas—as well as how tens of billions of dollars in federal funds are distributed. It will also become grist for countless studies. Governments use census data to plan the construction of roads, hospitals, and schools. Educators, businesses, and nonprofits plot their course using the census as a compass. Says Congressman Dan Miller (R–FL), who heads the House committee that oversees the bureau, "The census is the basis of power and money for the next decade." With all that at stake, notes Richard Rockwell, a social demographer at the University of Michigan, "we've got to get it right."

That imperative is keenly felt at the Census Bureau, which in a 1997 report to Congress concluded that its last premier event, the 1990 census, was a "large step backward." That head count cost $2.6 billion—more than twice the previous one—and according to subsequent analyses missed a record 4.7 million people, a disproportionate number of them low-income minorities (*Science*, 1 November 1996, p. 713).



World in a jar. If marbles of two colors are mixed well in the large jar, the fastest way to know their ratio is to pour a few into a smaller one and count them. This approach, called sampling, might be used to assess 10% of U.S. residents.

ILLUSTRATION: P. RIEHN SOURCE: TOMMY WHITE/CENSUS BUREAU

Statistical sampling was supposed to be the remedy for the bureau's ills. It's not a new idea: The Census Bureau has without fanfare used statistical techniques since 1940 to interpret its questionnaires, and in 1970 even used statistics to correct for apparently vacant apartments. "There was essentially a black-box method that created people and put them in households where we had no evidence that they existed," says Carnegie Mellon statistician Steve Fienberg. Those corrections, he says, added 1 million people to the total population. But the extensive use of sampling in the 2000 census would be unprecedented—if it is allowed to happen.

By boat, dogsled—and laptop? The 2000 census, which is expected to cost at least $4 billion, would proceed in three phases. In the first one, the bureau would cast as wide a net as possible, compiling a master list of the estimated 118 million U.S. housing units according to postal records and other sources. Next, it would deliver detailed questionnaires to every address by mail or by hand—employees have reached some locales by swamp boat or dogsled. With an aggressive ad campaign, the bureau hopes to get about 77 million sets of forms back, representing two of every three households.

Rather than try to reach all the remaining households—an arduous effort that the bureau estimates would add $675 million to the census's price tag—some 200,000 enumerators would hit the pavement to track down a random subset. The workers would ring enough doorbells to raise the response rate to 90%. That target must be met in each census "tract," an area containing approximately 4000 people. So if an initial response rate in a tract is only 50%, census workers would sample four of every five remaining households. Because the four would be randomly selected, bureau scientists say, the characteristics of the fifth can be statistically inferred. According to Rockwell, if you tried to count everyone you would fail—and worse, you'd have no idea who had been left out. Without sampling, he says, the error bars on the final population count are "unknowable."

Because this first round of sampling isn't designed to catch all the "wily trouts" who continually evade capture in the census net—including those not on the initial address lists—the bureau intends to conduct an intensive sampling effort to correct and supplement the figures. In this exercise, called an Integrated Coverage Measurement (ICM), census workers would fan out in randomly picked regions, eyeballing every house and compiling a second, independent address list. By interviewing inhabitants of 750,000 homes and comparing these results to the initial survey, the bureau can correct the overall population numbers.

The Census Bureau will also use the ICM to eliminate errors in which people get counted twice—for instance, a child claimed by two divorced parents. Using sampling and ICM, the bureau says it will be able to reduce the error rate from under- and overcounting to an unprecedented 0.1% at the national level, and 0.6% at the congressional district level. Without these statistical tools, the bureau says, the uncertainties would run about 1.9% across the board. The techniques are so robust, Rockwell claims, that "you could even do a really good census" by sampling 75% to 80% of the population. Although such a reduced headcount is not being considered for the 2000 census, down the road it could cut costs by several hundred million dollars, says associate census director John Thompson.

But some statisticians are skeptical that these analyses would produce a more accurate count. The bureau's error estimates "are all wrong," contends David Freedman, a statistician at the University of California, Berkeley, who is a vocal opponent of using ICM in the 2000 census. A variation on the ICM, Freedman points out, was used after the 1980 and 1990 censuses to estimate how many people slipped through the cracks. The Commerce Department, which oversees the Census Bureau, opted not to adopt corrections from the ICM calculations. That was a good move, Freedman says, because the corrections turned out to be wrong: In 1990, he says, the bureau found that a computer bug in the ICM software had erroneously inflated the corrections by about 1 million people. "In 1980 and 1990, it was very hard to tell if the adjustment would have moved us closer to the truth," says Freedman, who favors a more costly but complete head count. The ICM is so complex, he contends, that it introduces errors that are larger than the ones it fixes.

Other statisticians disagree. Sampling is a time-tested technique, "and this is not a particularly unusual application," says John Rolph, who led an ASA panel that strongly endorsed the sampling plan. Others acknowledge shortcomings in the methods but contend that they are better than relying on a head count. "The debate is not over whether we have a perfect method, but whether we have a sensible method," says Fienberg.

The arguments for and against sampling will be put to a key test later this year. In April, the bureau will conduct a census of Sacramento, using all the techniques it hopes to use in the 2000 national census. Workers will contact about 90% of the city's 400,000 people, then use sampling to estimate the rest. This "dress rehearsal could have a major role in the debate," says Ed Spar, executive director of the Council of Professional Associations on Federal Statistics.

But some fear that the experiment could be an exercise in futility. The political costs of sampling are too high to justify the potential improvement in accuracy, contends Alan Heslop, a professor of government at Claremont McKenna College in California. "I suspect that a lot of ivory tower people don't understand the real world situation," he says: By correcting for the undercounting of predominantly poor, urban dwellers, sampling "pits states against one another for House seats and federal funds." Indeed, Congressman Miller told *Science* that while he agrees that sampling is "good in theory," he predicts it would cause fewer people to return the forms because they would assume that the statisticians will save them the time and effort. "The possibility of a failed census is very real," he says. Others contend that partisan politics, plain and simple, is thwarting sampling. There is a "deep distrust by some members of Congress" that the Clinton Administration would rig the data in their favor, says Rockwell.

Last summer, House Republicans tried unsuccessfully to tack a sampling ban to a disaster-relief bill. The next salvo, observers say, could come in the Commerce Department's 1999 appropriations bill, which legislators will start working on with this week's release of the Administration's 1999 budget proposal. It promises to be a year of intense debate on the frontier where statistics meets politics.

–David Kestenbaum

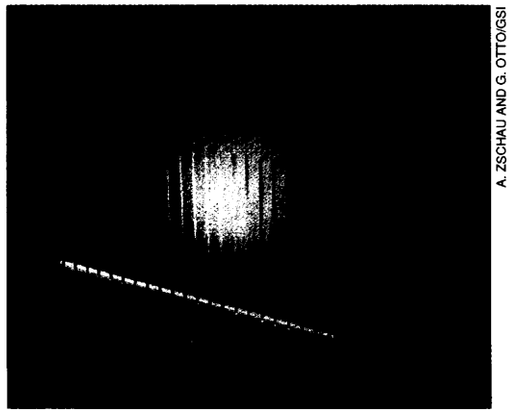
BIOPHYSICS

Blasting Tumors With Particle Beams

The physics of particle beams may seem like an esoteric branch of basic research, but two new research programs in Europe have begun exploring a down-to-earth use for two different types of such beams: treating patients with intractable brain cancer. Last week, the European Union's (EU's) commissioner in charge of research, Edith Cresson, inaugurated a five-nation program of tumor therapy using neutron beams from the high-flux reactor at the EU's Joint Research Center in Petten, the Netherlands. And later this month, two patients will get a first indication of whether a carbon-ion beam therapy developed at Germany's heavy-ion research center (GSI) in Darmstadt has halted growth of their tumors.

Researchers at Petten are the first in Europe to use a technique called boron neutron capture therapy, which builds on current work in Japan and at Brookhaven National Laboratory (BNL) in New York. Patients are injected with a boron-containing compound that selectively accumulates in the tumor tissue. Researchers then expose the patients to neutrons from the reactor, which are absorbed by the boron nuclei. The nuclei then emit damaging short-range ionizing particles that can kill the surrounding tumor cells while, in theory, leaving healthy cells comparatively undamaged.

The Petten team has already treated five patients and hopes to treat 40 more over the coming months. "There's an urgent need for controlled clinical trials to test the benefit of this treatment," says team member Wolfgang Sauerwein. The team plans to spread the radiation over four sessions to lessen damage to normal cells—BNL used just a single dose.



A. ZSCHAU AND G. OTTO/GSI

Depth charge. Varying beam energy alters the depth of energy deposition in tissue.

"We're very interested to see how that goes," says a BNL spokesperson.

In Darmstadt, GSI, in collaboration with the University of Heidelberg, the German Cancer Research Center in Heidelberg, and the Rossendorf Research Center near Dresden, has recently begun using highly focused beams of high-energy carbon ions to attack tumor tissue. Such heavy-ion therapy was first attempted at the Lawrence Berkeley National Laboratory in 1975, but its 1940s-vintage accelerator did not have a sufficiently stable beam or tunable energy, says GSI's head of radiation biophysics, Gunter Kraft. Japanese researchers at the National Institute of Radiation Research have also treated more than 150 patients with carbon ion beams. Although they have had some success, their beams cannot be focused entirely on the tumor, so some damage to normal tissue has occurred, says Kraft.

To overcome these problems, the German team has developed a new beam steering system. It divides the target tumor into layers and, starting with the deepest, scans each tumor layer in the same way an electron beam scans a TV screen. "By changing the beam's energy, it is possible to change the maximum dose at a particular depth very precisely," says Kraft. The ion bombardment produces gamma rays that researchers use to track the beam during treatment. "It's a very promising technique because of the selectivity of the target tissue," says radiobiologist Andre Wambersie at the Catholic University of Louvain in Belgium.

During the pilot phase, which began in December, two patients with tumors at the base of the skull were treated, and these will soon be followed by another 10 patients. A second phase will treat a larger trial group. The treatment proved that the beam can be manipulated with "millimeter precision," says Kraft. But the key test will come when the first two patients will be checked to see how much the treatment has shrunk their tumors.

–Nigel Williams