

LABOR CENTER REPORTER

Number 191
September 1986

LIES, DAMN LIES, AND STATISTICS: WHY YOU SHOULD CARE ABOUT HOW THE U.S. GOVERNMENT COLLECTS STATISTICS.

by Amelia Preece

The federal government collects data on a wide range of topics including weather, unemployment, inflation, economic growth, airline safety, incomes, consumption patterns of households, poverty, and industrial output. But Ronald Reagan's Administration has made large cuts in the budget for data collection, analysis, and dissemination. These cuts mean less data are collected. Some information is no longer available. The information which is available is published with less detail than before. For example, since fewer tables showing data for sub-groups of the population are now published, it has become impossible to compare family consumption patterns by income, race, or occupation. This lack of tables dramatically reduces the number of people who can use the data. The quality of data collection has deteriorated, sample sizes are smaller, and procedures are less up-to-date. The data are now made available a longer time after collection, so the information is less up-to-date. Fewer copies of data are published and when these copies are gone the information is only available in research libraries. And prices of the publications containing data and analysis have risen.

The changes in data collection under Reagan have not been haphazard. Many examples serve to illustrate how changes in data collection systematically help the Reagan Administration cut social services and hurt workers. In a recent federal study of the number of homeless people, the statistics were collected using a method that insured an under-count of the homeless. The homeless in temporary housing were excluded and large estimates of the number homeless were not accepted. This information was then used to support the federal policy of not assisting the homeless. The consumption data collected in 1980 were published in a form which made analysis by social scientists impossible. There was no break down by income for different races or regions, and there were very few income categories. However the consumption data were published in a way which made them a very useful marketing tool for private industry, even including information on where purchases were made. Unemployment data have been targeted for changes. There are plans to reduce the number of people interviewed to determine the unemployment rate. This will eliminate unemployment rates by state.

A similar tactic is being used to deflect attention from problems facing ethnic minorities. The Census Bureau now lumps all Asian Americans in one data set instead of providing separate data for different sub-groups. This is important because the problems facing Vietnamese Americans, for example, are quite different from those facing most Chinese Americans. The Census Bureau is also planning to reduce its coverage of Hispanics in the next census, in spite of the fact that Hispanics are the fastest growing minority in the U.S. This lack of data will seriously hamper studies of these special groups or of special problems they face.

The Administration has manipulated some statistics by refusing to update methods of calculation. Poverty for example, is seriously underestimated because of the method of calculation. In the 1960's, when the poverty standard was developed, poor families spent a third of their income on food. The poverty standard was therefore defined as three times the minimum cost of a nutritionally adequate diet. Now poor families spend only about one fifth of their income on food, but the poverty standard has not been increased enough to reflect contemporary consumption patterns.

BERKELEY, CA 94720
(415) 642-0323

UNIVERSITY OF CALIFORNIA, BERKELEY
CENTER FOR LABOR RESEARCH AND EDUCATION
INSTITUTE OF INDUSTRIAL RELATIONS



OCT 29 1986
UNIVERSITY OF CALIFORNIA
BERKELEY

The Administration has also neglected to develop new categories of data for new industries or create new surveys to analyze new problems. Without such innovation, it is difficult to measure the impact of such economic changes as the growth of the service sector and the integration of women into the work force. While the number of women in the labor force has been growing rapidly, the number of job categories where women are predominant in the Dictionary of Occupational Titles has remained relatively stagnant. This probably means that women doing newly created jobs are being lumped into existing "women's job" categories in the statistics.

Data collection is weakest for newer, expanding sections of the economy. For example, no separate category in federal statistics exists for firms producing computer software and no money has been allocated to create a new industry category. As data collection fails to keep up with economic changes we become less able to predict future changes in the economy, including changes in the labor market.

The Administration has increased the difficulty of using government collected data. Many of the statistics originally tabulated by the government are now left for individual researchers to tabulate for themselves using raw data from computer tapes. This means that only people with access to large computers and the skills necessary to use computers can examine the data. It also dramatically increases the time and cost of analyzing the data. Researchers may be spending weeks finding what a glance at a table would have previously shown. The changes in the data availability means that the data are no longer available for lay people. One needs at least a masters degree to be able to change it into a usable form.

Finding funding has become a critical step in analyzing data. Those with money can increasingly determine what research is done and by whom it is done and, ultimately, what researchers find out. The government has traditionally funded all research which uses its data. Researchers now need large amounts of money simply to tabulate data. And funding for research has become extremely political. Social scientists who have supported labor have found it increasingly difficult to obtain funding for anything besides research on joint labor management efforts.

The Reagan Administration claims that the government should not be in the business of statistics gathering, but rather that it should be done by private corporations. This position disregards the fact that census data collection is required by the Constitution and the fact that data is automatically collected by the IRS and other federal agencies. It also disregards the fact that only the government can require the provision of data and insure necessary privacy. Privacy is important both to protect the sources of data and to prevent data from falling into the hands of inside traders before it is made public. Data are valuable for those in the stock and commodity markets if information can be acquired by individuals before it is available to all. No one besides the government would be able to get the data from the sources the government uses. Private data would also probably make access more difficult and expensive. It is logical for the government to collect statistics because information is a "public good." Once the data are available it is impossible to prevent new users from copying them. Accurate statistics are important to collect since everyone is made better off by the availability of good information.

Congressional hearings have been held on the growing inadequacy of government data. At one of these hearings, Senator Paul S. Sarbanes (D-Md) found a "widespread and growing concern that our capacity to provide the statistical information on which sound judgment depends, in both the private and public sectors is increasingly at risk... Concerns about the accuracy and adequacy of our statistics which have been expressed with increasing frequency over the past year ... were underscored (by the explanations given by Administration officials)."

As the amount of data collected falls, the choice of data to be collected becomes more political. The administration in power is able collect data in those areas and in such a way as to support its own political agenda. The data needed to support opposing positions is becoming unavailable. Data can be a valuable political football. We must take steps to insure that this does not happen.

- *Amelia Preece*

This article does not necessarily represent the opinion of the Center for Labor Research and Education, the Institute of Industrial Relations, or the University of California. The author is solely responsible for its contents. Labor organizations and their press associates are encouraged to reproduce any LCR articles for further distribution.