

PRIVACY AND CONFIDENTIALITY ISSUES IN FEDERAL-STATE COOPERATIVE PROGRAMS OF THE BUREAU OF LABOR STATISTICS

By

Stephen H. Cohen¹

Bureau of Labor Statistics; Office of Survey Methods Research;
2 Massachusetts Ave, NE; Washington, D.C. 20212

Key Words: Restricted Access, Research Data Center, Licensing Agreements

Abstract

The Bureau of Labor Statistics (BLS) is the principle Federal statistical agency responsible for collecting and disseminating statistics on labor economics. BLS contracts with the States, the District of Columbia and some U.S. territories to collect and process establishment employment statistics and occupational safety and health statistics. These types of programs are called Federal-State cooperative programs. The BLS and the States enter into an agreement that defines the role of each partner. This paper will discuss the legal, methodological, and organizational issues on privacy and confidentiality associated with Federal-State cooperative programs. Particular emphasis will be made on issues associated with sharing of the data between BLS and the States and researcher access to the data.

Introduction

The U. S. Bureau of Labor Statistics (BLS) is the principle Federal Agency in disseminating statistics in the area of labor economics. BLS collects and disseminates statistics on employment and unemployment, price change, compensation and working conditions and productivity. Data for eight employment programs and two workplace occupational injury and illness surveys are collected through Federal-State cooperative programs—Covered Employment and Earnings (ES-202), Longitudinal Data Base (LDB), The Occupational Employment Statistics Program (OES), Annual Refiling Survey (ARS), Multiple Worksite Report (MRR), Current Employment Statistics Program (CES), Mass Layoff Survey (MLS), Local Area Unemployment Statistics

Program (LAUS), Occupational Safety and Health Survey (OSHS) and the Census of Fatal Occupational Injuries (CFOI). Federal-State Cooperative programs are arrangements where surveys are jointly developed and managed as a joint partnership with the States.

BLS has had cooperative arrangements with certain States since 1902 to use employment statistics collected by the States for National-State statistical programs. Several of these programs have been funded jointly by the BLS and the Employment and Training Administration. Since 1984 the BLS has assumed fully responsibility for managing the CES, OES and Local Area Unemployment Program (LAUS).

These Federal-State cooperative efforts are implemented by BLS via cooperative agreements that define the responsibilities of each party. These cooperative agreements require that the State employees adhere to the same privacy and confidentiality requirements that are required of BLS employees.

BLS reimburses the States for their services through the cooperative agreement. The OSHS requires a 50-50 cost sharing between BLS and the States that participate in this program. BLS collects data directly in non-participating States. Data for these States are needed to generate National estimates.

These cooperative efforts give both partners the ability to shape the programs. For example in the CES, States have substantial influence over BLS due to the Workforce Investment Act. This act requires the Department of Labor to consult with the States in administering labor force programs. BLS has set up a Workforce policy council for advice on State needs. The States elect members to the council from each of ten geographic areas. For example, the policy

¹ Any opinion expressed in this paper are those of the author and do not constitute policy of the Bureau of labor Statistics.

council recommended that BLS develop model based estimates for geographic areas where there isn't adequate sample to support design-based estimates.

In the OSHS program, the States can set the sample allocation design parameters so that the survey can produce estimates for NAICS categories that are of importance to the States. Due to budget constraints that limit the size of the total sample, targeting samples to NAICS codes is iterative.

BLS Confidentiality Requirements

All data collected by BLS directly or through the States must adhere to the requirements found in Commissioner's order 3-93. Specifically: data collected or maintained by, or under the auspices of, BLS under a pledge of confidentiality shall be treated in a manner that will assure that individually identifiable data will be used only for statistical purposes and will be accessible only to authorized persons. Data refers to all elements of information from a statistical program. Individually identifiable data refers to all elements of information (including but not limited to names and addresses) that might identify participants in a statistical program by either direct or indirect means. Statistical purposes refers to the description, estimation or analysis of the characteristics of groups without regard to the identities of individuals or organizations that comprise such groups, and the development, implementation or maintenance of methods, procedures or information resources that support such purposes. This definition does not include any use of individually identifiable data for administrative, regulatory, enforcement or other similar purposes.

BLS denies requests for microdata or special tabulations where individually identify data could be compromised. BLS has successfully defended these denials under the Commissioner's order in court at least three times over the last 25 years.

This order is backed up by Federal Statutes and Office of Management and Budget Statistical Directives that are supported by legislation. In particular, BLS data is protected by the Confidentiality Information Protection and Statistical Efficiency Act of 2002. This act ensures that willful disclosing of information in any manner to a person or agency not entitled to

receive it, shall be guilty of a class E felony and imprisoned for not more than 5 years, or fined not more than \$250,000, or both.

Data Sharing between BLS and the States

The cooperative agreements sets the legal stage that gives cooperating State agencies access to data collected under the agreements. The State agents must follow the limitations in the Commissioner's order 3-93 just like BLS employees. Authorized State employees can handle and discuss survey microdata between each other. However, they are restricted by Commissioner's order 3-93 on what they can release to the public just as BLS employees.

Under the cooperative agreements it is BLS' responsibility to define the methodology and set the parameters on data to be released to the public to ensure adherence to the Commissioner's order 3-93. BLS gives the States the flexibility to publish detail that meets the State's requirements that quite often is different than BLS' needs for providing data to assist Federal policy makers in developing Nation economic policy. However, the States must still adhere to BLS confidentiality requirements. Each program has its unique arrangements and processing system for States to publish additional detail. I will highlight three programs as examples.

The CES program estimates employment by NAICS industry code monthly. The National office releases only National employment counts by NAICS code in a press release or on its web site. The States have the flexibility to release more detail subject to confidentiality and reliability constraints. A State like Maryland releases the following series: State, Baltimore Metropolitan area, Baltimore City, Maryland portion of the Washington D.C. Metropolitan Area, Eastern Shore/Southern Maryland and Western Maryland.

The OES program estimates employment counts of occupations along with the associated wage quartile. BLS publishes on its web site the occupational employments estimates at the National level, by State, and by metropolitan area within State. The States each release their

own State data on their web site. In addition to the same products, States release county detail and geographic configurations of their choice. The computer system is designed for the States to produce county and balance of State estimates on their own provided it meets BLS confidentiality requirements. Of course, the finer the level of geography, the less detail is released.

The OSHA estimates incidence of injury and illness rates by NAICS code with detailed estimates of the case characteristics associated with the incidence rates. BLS only publishes national statistics in its press releases or on its web site. However, the processing system produces estimates for all statistics that meet reliability and confidentiality criteria. The States can and do publish State detail they consider relevant for their State.

The ES-202 data sharing arrangements with the States is more complex. The basic employment and earnings data come from State administrative data bases used to administer their unemployment insurance programs. The data is supplemented by the ARS for NAICS coding and MRR which disaggregates company wide county data into individual establishments. The ARS and MRR pledge that the data will be used only for statistical purposes subject to State laws. Most States depend upon and follow the confidentiality requirements BLS uses internally. However, some States publish ES-202 detailed data that do not meet BLS confidentiality standards. Since the administrative data is not collected under the Commissioner's order this is permissible. It is the datafile created by merging the administrative data with the ARS and MRR that is subject to the confidentiality requirements.

Data Sharing with Researchers

BLS does not produce public datasets of restricted data elements for researchers to access for analysis. However, we do offer researchers three mechanisms for restricted access to confidential data: researcher access to confidential data files at the BLS; licensing agreements; and research fellowships and post doctoral programs.

Researcher Access to Confidential Federal-State Cooperative Program Data Files at the Bureau of Labor Statistics

The BLS has opportunities available on a limited basis for researchers from colleges and universities, government, and eligible nonprofit organizations to obtain access to confidential BLS data files for exclusively statistical purposes. These data files, which are derived from BLS surveys and administrative databases for which no public-use version is available, include:

Longitudinal Database of Establishments Covered by State Unemployment Insurance Programs

Occupational Employment Statistics Survey

Current Employment Statistics Survey

Job Openings and Labor Turnover Survey

National Compensation Survey

Survey of Occupational Injuries and Illnesses

These confidential BLS data are available for research that is exclusively statistical, with appropriate controls to protect the data from unauthorized disclosure. BLS confidential data files are available for use only at the BLS National Office in Washington, D.C., on statistical research projects approved by the BLS. Researchers granted access to the confidential data will be responsible for adhering to the confidentiality policies of the BLS.

Eligible researchers include college and university faculty and researchers in government and eligible nonprofit organizations. Eligible researchers can qualify for an opportunity to use confidential BLS data files for statistical research under terms of the Intergovernmental Personnel Act (IPA). To qualify for an IPA assignment, a researcher must be a citizen of the United States or meet one of several exceptions and be employed by an institution of higher education; an eligible nonprofit organization; a State, local, or Indian tribal government; an association of State or local public officials; an organization representing member State or local governments; or a Federally funded research and development center.

The employee must be in a career position with the organization and have held the position for at least 90 days prior to the IPA assignment.

He/she must not have served under the IPA authority for four continuous years without at least a 12-month return to duty with the organization from which you originally were assigned.

Researchers coming to the BLS to perform statistical research will not be compensated by the BLS for the work or any associated expenses, such as traveling or living expenses. In order for a researcher to participate in this program, the eligible institution with which he or she is affiliated must enter into a written agreement with the BLS. Published outputs from the research are in the public domain and researchers cannot copyright reports or publications based on the research.

The BLS Research Centers are located in the Office of Employment and Unemployment Statistics, the Office of Compensation and Working Conditions, and the Office of Prices and Living Conditions. BLS staff from those offices will be available to answer specific questions about the data files, but no other research assistance is available.

To protect the confidentiality of the data, researchers will not be permitted to use their own computer equipment. In addition, BLS PCs will not provide access to e-mail or the Internet. BLS staff will review all printouts, disks, and research reports and other project outputs to ensure that data confidentiality is protected.

Because resources within each of the research data centers is limited by space, equipment availability and staff time constraint, BLS has an application procedure for researchers to ensure each proposal is evaluated by objective measures.

The BLS considers applications for research proposals four times a year. Applications must be submitted by January 15 for projects beginning in the spring, April 15 for projects beginning in the summer, July 15 for projects beginning in the fall and October 15 for projects beginning in the winter. For a research project to be approved by the BLS, it must have technical merit, must be of significant interest to the BLS, and must further the mission of the BLS to provide objective economic and social research.

Researchers must abide by Federal laws and BLS policies that protect the confidentiality of the individuals and organizations who provide information used in BLS statistical programs.

Research proposals should be between 5 and 10 pages and should contain detailed information about the research project, including a literature review and an indication of how the proposed research contributes to the literature, the hypotheses to be tested, the data set and variables to be used in the analysis, the empirical methods to be used, and the specific data outputs that will result from the project. The proposal should be as specific as possible about the use of confidential information not available on public-use files.

Research proposals should also include:

- A description of how the research results will be presented, such as in a journal article or book.

- A description of whether the research results will be presented as descriptive statistics or frequencies, or via multivariate analysis such as regression coefficients.

- Resumes that include descriptions of the professional background and relevant research record of each individual who will access confidential microdata at the BLS.

- A timeline detailing access dates at the BLS, the frequency and duration of visits, and an estimate of the total length of access to the confidential data. IPA assignments may last up to two years.

- The name and address of the high-ranking official in the research or government organization who will formally sign the agreement with the BLS to permit access to the confidential data. For college and university researchers, this official must be at the level of a dean or higher; a department chair cannot be the signatory.

It is anticipated that prior to submission of a proposal, the researcher(s) will have contacted the BLS to determine the availability of data and the feasibility of the project.

Licensing Agreements

A licensing agreement is a formal agreement that permits confidential microdata to reside on a researcher's personal computer in their home institution. These agreements are formal legal documents between BLS and the host organization that specify the conditions under which the specific data set licensed may be used and the penalties for violation. The principal researcher must demonstrate that the data is required for research; i.e., public use data, if it exists, is not adequate. The goals of the research that require non-public data must be stated in the application. The licensor must approve the goals of the research before the application process can proceed.

The agreement specifies which people in the licensee's institution must sign the form. For an academic department it is typically a Dean and not the department chairman. The principal researcher (PR) must supply a list of names of people who will be authorized to use the data. The list must be updated as personnel working on the data changes. Those people must be informed of their responsibility not to share the data with people outside the group. The PR must indicate the group's experience, if any, with handling other licensed data sets. The agreement also includes a statement concerning which law(s) protects the data (e.g., Privacy Act of 1974). A data security program must be developed and implemented. The licensee's institution must allow inspections of the area where the data are used and stored. The inspections can be unannounced. Penalties for violations of aspects of the agreement are listed on the form (e.g., denial of use of other data from the licensor, fines, prison terms, etc.). There is a requirement that no attempt will be made to determine the identity of respondents. In general, the licensee is not allowed to link the licensed data to other microdata files.

Articles, reports, and statistical summaries generated from the data must be reviewed by the agency before they are published or otherwise communicated. The results must adhere to the agency's disclosure limitation practices (e.g., all non-zero cells in a publicly released table must represent some minimum number of respondents).

BLS makes CFOI available through licensing agreements.

Fellowships and Post Doctoral Programs in Principal Statistical Agencies

Research fellowships and post-doctoral programs provide the unique opportunities for researchers to address some of the complex methodological problems and analytic issues relevant to agency's programs. Fellows and post-doctoral candidates will conduct research in residence at agency, use agency data and facilities, and interact with agency staff. They will adhere to the same confidentiality agreements as regular employees. Proposals are reviewed within the BLS and by an external Program Review Board consisting of representatives of ASA, BLS, the American Economic Association, and the American Association of Public Opinion. Research proposals will be evaluated on the applicability of the research to BLS programs, the value of the proposed research to science, and the quality of the applicant's research record. Research fellows are contract employees and subject to the same confidentiality pledge signed by agency personnel.

Examples of recent research activities are indirect estimation, imputation, statistical computing, quantitative methods and data, confidentiality, data linkage and cognitive processes.

The federal government has many varied post-doctoral programs. In statistics they range from long established programs in the National Institutes of Health to recently developed programs at the BLS and the BOC. The goal of these programs at the BLS is to develop an interest in recent Ph.D. graduates on survey methodology research. It is hoped that these researchers might then become attracted to a career with the statistical agencies in survey methodology or goes on to teach sample design work in academic institutions.

Post-doctoral researchers at BLS would become temporary employees; thus, subject to the confidentiality pledges that all employees must sign. Post-doctoral candidates are restricted to recent college graduates, typically having received their Ph.D. degrees within the last 2 to 3 years. Their research proposal is written in

conjunction with a senior methodologist within the agencies.

Conclusions

BLS has a long tradition going back to 1902 for acquiring data from the States. The Commissioner's confidentiality order applies to this data just as it does to data collected directly by BLS employees. The States are given discretion to publish data that deviates from the National products BLS must produce provided they adhere to BLS confidentiality requirements.

Researchers are allowed restricted access to the data under three scenarios: doing their research at BLS facilities or obtaining a license to use the data at their institution if they agree to abide by BLS security requirements.

References

Doyle et al (2001); *Confidentiality, Disclosure, and Data Access, Theory and Practical Applications for Statistical Agencies*, North-Holland

Cohen, Stephen (2001,02), "Access to Confidential Statistical Agency Data," *Proceedings of the Survey Methods Research Section*

BLS Handbook of Methods, U.S. Department of Labor, Bureau of Labor Statistics, April 1997, Bulletin 2490