

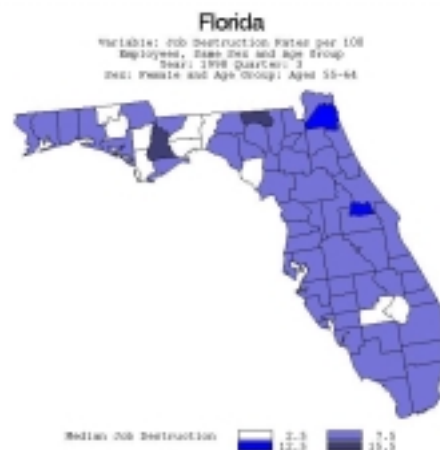
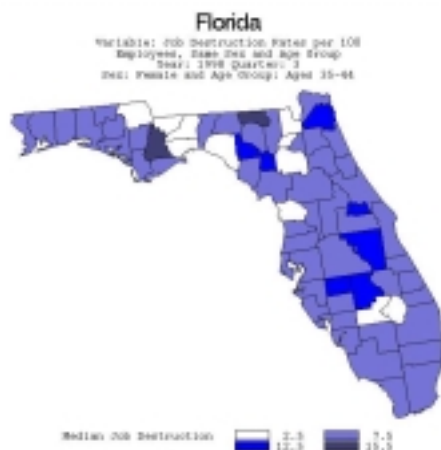


Longitudinal Employer-Household Dynamics Program

FOB #3 Room 2138 Washington DC 20233 Phone: 301-457-8304 Fax 301-457-8430

U.S. CENSUS BUREAU
Helping You Make Informed Decisions

LEHD Newsletter November 7, 2001





Longitudinal Employer-Household Dynamics Program

FOB #3 Room 2138 Washington DC 20233 Phone: 301-457-8304 Fax 301-457-8430

U.S. CENSUS BUREAU
Helping You Make Informed Decisions

Why this newsletter?

We've been busy doing data work, but we wanted to keep you informed about all of the programmatic research and development activities conducted by the staff of the Longitudinal Employer-Household Dynamics Program. We've organized the newsletter along the lines of our on-going projects. We hope that you find the research summaries helpful. Here's the list of projects

1. State Employment Dynamics Estimates
2. BLS-Census Prototype Expansion of the EDE
3. LEHD Infrastructure
4. Human Capital, Worker Flows and the Definition of a Business
5. Demographic Survey Improvement
6. American Community Survey Place of Work Coding
7. SSN Validation
8. Population Estimates Improvements

For additional information about items in this newsletter contact Ron Prevost, Director of the LEHD Program (Ronald.C.Prevost@census.gov, 301-457-8304)

I. State Employment Dynamics Estimates (EDE)

A. Version 2 enhancements

1. Place of Work Enhancements

Historical and updated data from our original partner states (CA, FL, IL, MD, TX) are being integrated into new versions of our individual characteristics, employer characteristics, and employment history files. These new files will be used to update and enhance the version 2 employment dynamics estimates. Updates will provide data through 2000:4 (for states that have already submitted those data) with improved quality arising from extensive SSN editing, extensive, updated Census Numident information, improvements to the confidentiality protection, and the state-requested addition of full quarter new hires.

Estimated availability December 2001.

2. Place of residence-based version 2 estimates.

Enhancements to the individual characteristics file now permit LEHD to estimate worker flow and earnings data by county of residence and workforce investment board of residence. These estimates will supplement the employer location-based employment dynamics estimates (versions 2 and 3).

Estimated availability January 2002.

3. Decennial Census quality control for the Employment Dynamics Estimates.

We have begun a research project to compare our version 2 and version 3 estimates of the age and sex composition of the work force by county and by workforce investment board using the 1990 and 2000 Decennial Censuses of Population and Housing. Comparison of estimates of April 1 (beginning of quarter 2) employment will be conducted for all partner states.

Estimated availability: first quarter 2002.

B. Version 3 research and development

1. Improved place of work information.

One of the core enhancements in version 3 of the Employment Dynamics Estimates is to improve the accuracy of the place of work by using the ES-202 establishment as the business-entity that defines the worker and job flow and earnings estimates. In order to do this, LEHD must develop an algorithm for imputing ES-202 reporting units for the UI wage record data that form the core of our employment history files. The methodology is being extensively tested using the UI wage record and ES-202 data from the state of Minnesota, the only state with reporting unit information coded onto the wage record data. The methodology will then be presented for discussion with our state partners. Following these discussions, LEHD will implement the methodology for current partners (CA, FL, IL, TX, and MD) as well as new partners (MN, NC, NJ, and PA).

Estimated availability: Methodology report January 2002.

2. Improved confidentiality protection.

The second core enhancement in version 3 is refinement of the disclosure-proofing methodology to permit release of some industry within county estimates. The research is completed on the methodology. All LEHD EDE estimates are benchmarked to an existing BLS ES-202 estimate of employment in the first month of the quarter. Other confidentiality protections include the use of various estimation techniques and the insertion of noise at the business unit level. The Census Bureau's Disclosure Review Board has fully vetted the LEHD methods. The BLS's Office of Statistical Research (equivalent to the Census Disclosure Review Board) is studying the method now. The refined methods have been fully implemented for version 2 estimates. The refined methods, with any modifications requested by the BLS, will be used for all version 3 estimates.

Availability: version 2 immediate, version 3 first quarter 2002.

3. Reconciliation of historical SIC-based estimates and on-going NAICS-based estimates.

The LEHD program and the BLS ES-202 program are cooperating to discuss the feasibility of historical NAICS-based EDE estimates. The BLS is reconsidering its original decision not to release historical ES-202 employment and payroll tabulations on a NAICS basis. Historical NAICS-based data are essential to LEHD's current confidentiality protection system. Research is underway to develop both NAICS and SIC-based historical employment dynamics estimates (all variables).

Expected availability: version 2 first quarter 2002, version 3 second quarter 2002.

4. Graphical Delivery System

LEHD has begun experimenting with presenting EDE's in graph or map format – the first pass is on the front cover. Input will be requested of all the states at the January workshop (and before!)

Expected availability of prototype: January, 2002

C. SSN editing

1. Editing SSNs to improve the accuracy of employment histories based on UI wage records.

LEHD performed extensive research on the use of statistical matching methods to detect and correct errors in the coding of SSNs on the UI wage records that are used to create our employment history files, which in turn form the core of the EDE. The test state was California. The results were very encouraging, with major corrections of work histories that contained single-quarter holes using work histories from the same UI account that existed for only that quarter. Results were

reported to California in July, 2001. The methodology is being applied to FL and TX now.

Availability: Report immediate, SSN-edited files 6 months from receipt of original UI records (if specified in MOU).

2. Effectiveness of SSN editing for worker and job flow estimates.

Research comparing the employment dynamics estimates before and after SSN editing is underway. CA, TX, and FL version 2 estimates are being created on a “before and after” basis for comparison.

Availability: January 2002.

D. Successor-predecessor editing

1. UI-account level successor-predecessor editing.

LEHD has an extensive ongoing research project to integrate the information from worker flows with the ES-202 successor/predecessor information on the UI account (called an SEIN by LEHD). The method determines probable successors, predecessors, breakouts, and consolidations based upon the flow of workers from one (or several) SEIN(s) to one (or several) different SEIN(s). The methods dramatically improve the accuracy of the employment dynamics estimates for SEINs with more than 5 employees. The research is being conducted on Florida and will be ported to all partner states in version 3.

Availability: report January 2002.

2. UI-account/Reporting unit successor/predecessor research.

The development of version 3 estimates using the reporting unit as the basic business entity necessitated a research program to supplement the ES-202 successor/predecessor information on reporting units (called SEINUNIT by LEHD) within SEIN. We conducted this research in collaboration with BLS experts on the methods used to create the LDB. The primary research state was Illinois. The research shows that most candidate successor/predecessor relationships (including breakouts and consolidations) can be inferred from quarter-to-quarter changes in the first month employment of the SEINUNITs within a given SEIN. The SEINUNIT edit is then used to create a flag for succession, breakouts and consolidations that will be used in the version 3 estimates to eliminate the effects of these business recombinations on the job and worker flows.

Availability: report January 2002.

E. Cornell University support site.

The Cornell University LEHD support site is implementing a full simulation of the EDE micro data as well as a full installation of the SAS-based employment dynamics estimates. The confidentiality protection methods used for this support site were developed

over the past year. They permit state partners and other authorized users to conduct research on the flow data using a variety of statistical tools that may not be available to the individual states as well as software developed by LEHD staff. In addition, the support site permits work on a full-scale confidentiality-protected simulated version of the EDE micro data from the individual characteristics, employer characteristics, and employment history files. Some results based on the confidentiality-protected simulated data can be used directly. Other analyses can be reproduced on the original confidential data, then released to the authorized user after appropriate confidentiality protections have been performed.

Availability: report immediate, system: prototype December 2001, production first quarter 2002.

F. Related programmatic and sponsored research

1. Low wage workers.

A Russell Sage/Rockefeller Foundation/HHS sponsored research program is studying the labor force dynamics of low wage workers using the LEHD individual characteristics, employer characteristics, and employment history files from cooperating partner states (CA, FL, IL, MD, TX, MN).

Availability: first report January 2002; final report August 2003.

2. Demand for older workers.

NIA sponsored research on the demand for older workers will focus on the kinds of firms that employ older workers and on the differences between the employment dynamics of older and younger workers. The research will be conducted using the EDE data from cooperating states (CA, FK, IL, MD, TX, with other states asked to join as their data become available).

Availability: report fourth quarter 2002.

3. Displaced workers.

One of the most striking facts to emerge from the literature on displaced workers is the earnings declines experienced by workers several years before displacement occurs. Little attention, however, has been paid to other changes in compensation and employment in firms prior to the actual displacement event. This research examines changes in the composition of job and worker flows before displacement, and compares the "quality" distribution of workers leaving distressed firms to that of all movers in general. The research is being conducted using cooperating partner states EDE files (IL, MD).

Availability: report immediate.

II. BLS-Census Prototype Expansion of the EDE

A. Inclusion of Minnesota, New Jersey, North Carolina, and Pennsylvania.

The program has been authorized to expand from five to nine state partners. MOUs (Data Use Agreements) have been completed for MN, NC, and PA. These states will receive historical version 2 estimates in conformance with the terms of their MOUs (six months from the arrival of data) and will receive version 3 estimates as soon as their data can be incorporated into the version 3 system files.

Availability: second quarter 2002.

III. LEHD Infrastructure

A. Individual characteristics, employer characteristics, and work histories.

The LEHD Program, as a part of its Title 13 Census mandate and under NSF, NIA and Sloan Foundation support, creates, maintains and enhances data products that permit the integration of Census demographic products (surveys like the SIPP, CPS, and ACS), administrative data (Federal tax information, state UI wage records, ES-202 records) and Census economic data (Business Register, Economic Censuses, and Economic Surveys). These data products are then used directly to improve the Census Bureau's demographic surveys; economic censuses and surveys.

Availability: ongoing.

B. SIPP, CPS, SSA integration.

Continuing Title 13 research studies the quality of the administrative data integration into the SIPP and CPS. The core issue is the quality of the SSN provided by the respondent (SSN validation) and the accuracy of the integration of the SSA-provided federal tax information (W-2 and summary earnings records).

Availability: report first quarter 2002.

C. Business Register, Economic Censuses and Survey integration.

Continuing Title 13 research studies the quality of the integration of the business units defined by the Business Register, Economic Censuses, and Economic Surveys with the Census Bureau's demographic products. This research involves the testing of different methods of exact (identification number) integration and statistical matching (using name and address information).

Availability: report first quarter 2002.

IV. Human Capital, Worker Flows and the Definition of a Business

A. Measuring the human capital inside a business.

LEHD's Title 13 mandate includes developing a method to measure the inputs that a business uses in order to better understand differences in the productivity differences among businesses. Human capital is measured by the "portable" part of an individual's earnings—the part that another labor market competitor would pay for an individual's services. LEHD uses information on the education and labor force experience of individuals in the individual characteristics files, earnings and work histories from the employment histories, and employer characteristics from the employer characteristic files of participating states (CA, FL, IL, MD, MN, TX) to estimate the human capital within a business. LEHD uses the "non-portable" part (associated with working in a particular business or industry) to characterize differences in the compensation policies of the businesses, which are also related to measurable productivity differences.

Availability: report December 2001.

B. Technological investments and human capital flows.

LEHD's Title 13 mandate includes using the measures of human capital developed above to study differences among entrant, continuing, and exiting businesses. To date, this research has been conducted using integrated IL files that include individual, employment history, employer characteristics, Economic Census data (1992, 1997), and various Economic Surveys. The research shows that persistent differences in business productivity are directly related to persistent differences in the human capital employed by the business. Exiting businesses have less human capital than continuing or entering businesses.

Availability: report currently available.

C. New Health and Pension Benefit Information.

The accurate measurement of employer-provided health and pension plan coverage is critically important for policy makers who want more information about both the quality of jobs and their future financial responsibilities towards their poorest citizens. Yet two major Census Bureau surveys -- the Current Population Survey (CPS) and the Survey of Income and Program Participation (SIPP) -- produce substantially different estimates of the uninsured. Similar problems surround the collection of private pension information. Current information could be vastly improved by improving estimates of coverage and by bettering the firm level information currently available. LEHD is working to exploit an underutilized resource -- the 5500 file -- to achieve this improvement. The 5500 file, which contains data from annual returns filed annually with the Internal Revenue Service by all employers covered by ERISA, provides information on employee benefit plans, including data on different types of pension and health insurance plans. The LEHD program is undertaking an investigation of the quality of the 5500 file with the intent of answering the following questions.

1. Examining what kinds of firms offer different types of health insurance/pension plans -- particularly by size class, industry and legal form of ownership.
2. Examining how firms change coverage types

3. Examining how workers at firms offer health insurance/pensions differ from workers whose firms do not – and how this has evolved over time.
4. Examining whether firms that offer health insurance/pensions have lower turnover rates than firms that do not – and how this has evolved over time.
5. Examining differences in turnover rates for different types of workers at firms with health insurance/pension availability. For example, what are the effects of health insurance on the job mobility of a low-wage worker, relative to a high wage worker?

Availability: no report.

V. Demographic Survey Improvements

A. SIPP earnings and work history improvements.

The Survey of Income and Program Participation is the source of official US statistics on participation in a wide variety of income support and poverty assistance programs. LEHD's Title 13 mandate focuses on improving this survey by using administrative records (from federal sources) to enhance the survey information on work history and earnings. This research program has already provided major enhancements the SIPP data from the 1990s (including the 1996 SIPP, which is the source of current statistics and public use data).

Availability: report second quarter 2002.

B. CPS earnings improvements.

The Current Population Survey is the source of official US unemployment and poverty statistics. LEHD's Title 13 mandate focuses on improving the earnings measures in the March Annual Demographic Supplements by using integrated administrative data (from federal sources).

Availability: report second quarter 2002.

C. SIPP/SSA/CBO public use data project.

LEHD's Title 13 mandate includes researching the feasibility of creating a public use file that combines some SIPP variables with federal information from the Social Security Administration on employment histories and earnings from the SSA master earnings and benefit files. This research is being done in collaboration with the SIPP branch at Census, SSA, and the CBO. Difficult issues in confidentiality protection remain to be solved.

Availability: no report.

VI. American Community Survey (ACS) Place of Work Coding

A. Business Register, ES-202, ABI integration.

An essential requirement of LEHD's Title 13 research improving the ACS requires the creation of a reliable, extended master address file that permits latitude and longitude coding of both the residence and employer addresses. The integration of the Census Business Register, the partner ES-202 addresses, and the commercial ABI addresses is accomplished through the use of a common address standardizer. Resulting geo-coding is used for the ACS place of work research and for other geo-coding requirements such as the place of residence in the EDE version 2 enhancements.

B. Place of work coding in the American Community Survey

External reviews of the ACS, which is proposed as the replacement for the Decennial Census long form, suggest that Census should invest in improving the place of work coding. Place of work coding is used for employment statistics and for transportation and other urban planning. LEHD's Title 13 research compares the quality of place of work coding based upon the respondent's reported place of work and the place of work obtained from the integration of administrative records (federal and state sources).

Availability: no report.

VII. SSN Validation

A. Improvement of the validation of SSNs on administrative records.

This Title 13 research is being conducted jointly with the Administrative Records Research Staff at Census. The goal is to develop a method for using the Census Personal Characteristics File (based on the SSA Numident) to validate SSNs on administrative records. This replaces the work of clerks validating the SSNs from microfiche copies.

Availability: December 2001.

B. Alternatives to using SSN for administrative record integration.

This Title 13 research is also being conducted jointly with the Administrative Records Research Staff at Census. The goal is to extend the SSN validation methods to methods that permit the integration of data from the Census Personal Characteristics File onto administrative records and surveys that do not have SSNs.

Availability: no report.

VIII. Population estimates improvement

A. Immigration.

Census is using a combination of administrative records and modeling techniques to enable them to correctly count the number of immigrants in the US, and understand why their estimates were erroneous in the 1990's. LEHD program data are matched to Census data to model the sources of the gap. We expect there to be some relationship between jobs and population – which is likely to vary by region and ethnic origin. Once this historical relationship is modeled, LEHD is investigating the potential to merge ACS and LEHD data to provide better estimates of the immigrant population in the intercensal years. LEHD is particularly focused on modeling both the static and dynamic link between jobs and population.

As part of this LEHD is exploring the wage growth and employment patterns for recent immigrants, conducting comparisons across and within different immigrant groups and contrast these results with native workers. Differences in firm size, location, payroll practices, and possibly even health care coverage will give a more detailed picture of the immigrant experience than previously available. In addition, LEHD is investigating the clustering of immigrant groups in certain geographic areas, particularly the evidence that immigrants also cluster together at specific firms. Examining the rise and fall of ethnic clusters at both firms and neighborhoods, and the impact of clusters of workers at one firm on the wages and employment composition of nearby firms is an important component of the project. Finally, the progress made by immigrant offspring will provide insights into the long term effects of immigration on population estimates.

Another component of the research, which will be used to model cross state migration of native born and immigrant workers, re-examines the impact of immigration on native earnings and employment during the 1990s. The focus is particularly on the effect of the interaction between the effects of immigrant competition at the establishment level and at industry, regional, and statewide levels.

Availability: Preliminary report available first quarter 2002; final report fourth quarter 2002.