

A User's Guide to the Federal Statistical Research Data Centers

Mark Roberts
Professor of Economics and Director PSU FSRDC

March 2022

- Introduction to the Federal Statistical Research Data Centers (FSRDC)
- Data Resources in Economics
- Data Resources in Health and Demography
- Procedures for Using the PSU FSRDC.

Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the U.S. Census Bureau.

Federal Statistical Research Data Centers

- Network of secure data centers that provide researchers with access to restricted data from
 - Census Bureau
 - Agency for Healthcare Research and Quality (AHRQ)
 - National Center for Health Statistics (NCHS)
 - Bureau of Labor Statistics
 - Department of Housing and Urban Development
 - Bureau of Justice Statistics
- 31 RDCs in operation - most major Research-I Universities have access.
- Penn State opened April 2014, Paterno library
- www.census.gov/fsrdc and www.census.gov/ces

- No public-use data sets for business research
- Public-use data sets are widely used for demographic research (ACS, NHANES, PUMS, SIPP) but limitations to protect confidentiality
 - Limited geographic information.
 - Top-coding of responses (income)
 - Exclude sensitive personal characteristics (weight)
 - Difficult or impossible to link individuals or businesses across surveys
 - Data restrictions have tightened over time

What Does the FSRDC Provide?

- Access to the microdata on individuals or businesses in the surveys
 - Detailed geographic identifiers to allow merging of city, county, or state information.
 - More detail on characteristics - place of birth, date of birth, occupation, income, firm or plant size
 - Individual identifiers that allow merging with administrative data sources and across surveys
- Maintains confidentiality of individual responses
 - Formal review process for research proposals which examines need for confidential data and feasibility of the project
 - Security clearance for researchers
 - Secure computing environment (Linux) to access census data servers
 - Disclosure review of outputs

Economic Research Using Census Data

Primary source for micro data on firms and establishments. Covers all sectors of the economy. Only comprehensive longitudinal micro data source.

- Employer Dynamics - entry, growth, and exit of firms
- Employment Flows - job creation and destruction due to firm turnover
- Productivity Growth - role of investment in capital and R&D
- Vertical Integration- linkages across plants/firms in production chain
- Firms in International Trade - importance of multinational firms in U.S. exports and imports, formation of trading networks

Longitudinal Business Database (LBD)

- annual 1976 - 2019
- covers 8.5 million establishments across whole economy (non-farm, paid employees)
- establishments are linked over time - permanent ID number
- establishments can be aggregated to the firm (enterprise) level - using firm ID
- limited information: geography, industry, payroll, employment
- major strength: plant ID number can be linked to other CB surveys and censuses
- <http://www.census.gov/library/working-papers/2002/adrm/ces-wp-02-17.html>

Five year intervals, mix of establishment surveys and administrative data

- Manufactures 1963 - 2017
- Construction 1972 - 2017
- Retail Trade, Wholesale Trade, Services 1977 - 2017
- Mining 1987 - 2017
- Transport-Communications-Utilities 1987 - 2017
- Finance-Insurance-Real Estate, 1992 - 2017

Detailed establishment information

- six-digit NAICS industry, geography (census place)
- value of shipments (primary and secondary products)
- employees, production workers, hours, payroll, benefits
- materials, energy expenditures, capital investment, inventories

Annual Surveys

Establishments or firms in non-census years

Samples that are stratified by size. Rotating sample every 5 years.

Most are current through 2019

- Surveys that are similar to censuses (output, inputs)
 - Manufacturing, Retail Trade, Wholesale Trade Services
- Specialized Surveys
 - R&D and Innovation
 - Capital Expenditures
 - Business Owners / Annual Business Survey
 - Quarterly Financial Report

Longitudinal Employer-Household Dynamics (LEHD)

- Combined worker-firm data
- Built up from state unemployment insurance records, merged with census surveys.
- Job-level quarterly earnings data, information on where workers live and work, and data on firm characteristics.
- All but 3 states contribute data but access varies
- 2000-2018, some states have older data.
- Underlies reports on job-to-job flows of workers and workforce indicators

Longitudinal Firm Trade Transaction Database

- Export Transactions - monthly 1994-2019, value, date, country of origin or destination, quantity, product code. Covers whole economy (34 million records in 2011).
- Import Transactions - similar to exports (59 million in 2011).
- Includes firm id so they can be linked to the LBD and other census surveys

- National Longitudinal Surveys (NLSY) - follow individuals over time, NLSY79, NLSY79 child and young adult, NLSY97
 - Education, training, family background, fertility, health, criminal activities
 - Geographic identifiers at Zip Code/ Census Tract level, School Surveys
- Current Population Survey (CPS) - employment, hours of work, earnings, person characteristics
- Survey of Occupational Injuries and Illnesses. - employer survey based on OSHA records