

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

Mark Roberts  
Professor of Economics and Director PSU FSRDC

September 2016

- Introduction to the Federal Statistical Research Data Centers (FSRDC)
- Data Resources in Economics
- Data Resources in Health and Demography
- Procedures for Using the PSU/Philly 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
  - Census Bureau
  - Agency for Healthcare Research and Quality (AHRQ)
  - National Center for Health Statistics (NCHS)
  - Bureau of Labor Statistics
- 24 RDCs in operation - most major Research-I Universities have access.
  - Washington, New York, Boston, Atlanta, Chicago areas.
  - Minnesota, Texas A&M, Michigan, Cornell, Duke, Berkely, Stanford, UCLA, USC, Penn State, Yale
- 6 under construction this year including Philadelphia
- [www.census.gov/fsrdc](http://www.census.gov/fsrdc) and [www.census.gov/ces](http://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 RDC 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 turnover and investment in capital and R&D
- Vertical Integration of Firms - linkages across plants 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 - 2011
- 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/ces/pdf/CES-WP-02-17.pdf>



Five year intervals, mix of establishment surveys and administrative data

- Manufactures 1963 - 2012
- Construction 1972 - 2012
- Retail Trade, Wholesale Trade, Services 1977 - 2012
- Mining 1987 - 2012
- Transport-Communications-Utilities 1987 - 2012
- Finance-Insurance-Real Estate, 1992 - 2012

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

Establishments or firms in non-census years

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

Variables are similar to census (output, inputs)

- Manufacturing - establishment based. 1973-2011
- Retail Trade - firm based. 1999-2010
- Wholesale Trade - firm based. 1995-2010
- Services - 1991-2010 except 1999-2001

# 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 50 states contribute data but access varies - Requires approval of every state used. Some states review every project, others give blanket permission. Approval for large, multi-state projects is difficult.
- 2000-2008, some states have older data.

## Other Specialized Economic Surveys

All have unique sampling frames. Can be linked to LBD.

- **Quarterly Financial Reports** - 10,000 firms in manufacturing, wholesale, retail, mining. Debt structure, liquidity, profits.
- **R&D and Innovation Surveys** - firm level, 1972-2009. R&D expenditure, employment, management, innovation, tech transfer, intellectual property.
- **Survey of Business Owners** - firm level, 1972-, Ethnicity, gender, race, vet status. Includes non-employee firms.
- **Capital Expenditure Survey** - firm level. Spending on equipment and structures. Includes non-employee firms.
- **Manufacturing Plant Capacity Utilization** - quarterly, 1974 - 2011

Complete listing at

<http://www.census.gov/ces/dataproducts/economicdata.html>

# Longitudinal Firm Trade Transaction Database

- Export Transactions - monthly 1994-2011, 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 - 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
- National Compensation Survey - 40,000 jobs sampled quarterly. Salary and benefits
- Survey of Occupational Injuries and Illnesses. - employer survey based on OSHA records
- Census of Fatal Occupational Injuries

Can apply to match BLS and Census data