Research Opportunities at the Triangle Census Research Data Center

Gale Boyd – Director
Department of Economics, Duke University

Bert Grider – Administrator
Center for Economic Studies, U.S. Census Bureau
Mission of the RDC Program

There is a natural “coincidence of interests” between the Census Bureau and its data users:

- Researchers want to conduct important research and...
- Census values the feedback to its data programs.

The RDC program is based on MOU with Census and the local host institution.

The host institution may create additional access arrangements.
Stated Objectives of RDC Program

• Improve Census programs (*required by law*)
• Maintain confidentiality (*required by law*)
• Provide high quality research
• Expand community of scholars
Research in the RDC Network

• Microdata on Firms (Economic)
  • sector studies: productivity, learning by doing
  • dynamics: jobs, entry and exit
  • topics: pollution, healthcare, capital expenditure

• Microdata on Households (Demographic)
  • topics: residential segregation, immigration, welfare reform, intergenerational mobility

• Bridging Across Data
  • firms and households, external sources, geo-spatial
Economic Data

- Annual Survey of Manufactures
- Longitudinal Business Database
- Economic Census
  - Census of Construction
  - Census of Finance and Insurance
  - Census of Manufactures
  - Census of Mining
  - Census of Real Estate
  - Census of Retail
  - Census of Services
  - Census of Transportation
  - Census of Wholesale
- Survey of Business Owners
- Commodity Flow Survey
- Auxiliary Establishment Survey
- Manufacturing Energy Consumption Survey
- Medical Expenditure Panel Survey, Insurance Component
- National Employer Survey
- Pollution Abatement Costs and Expenditures
- Quarterly Financial Reports
- Research and Development Survey
- Survey of Manufacturing Technology
- Worker Establishment Characteristics Database
Demographic Data

- March CPS Earnings Supplements
- Survey of Income and Program Participation
- American Housing Survey
- American Community Survey
Linked Household / Business data

Longitudinal Employer-Household Dynamics (LEHD) data

- Links households to place of employment
- Based on unemployment insurance administrative records
- Covers most states
- Quarterly starting in 1990

- "tracks" a person based on their place of employment
  - Establishment (i.e. the place of work) is exact for single plant companies
  - Establishment is assigned for all others (using geography and industry to improve matches)
Other Agency Data

The RDCs provide infrastructure to access data from
- National Center for Health Statistics
- Agency for Healthcare Research and Quality

Other agencies are in talks with Census
- National Institute for Child Health and Human Development
- Bureau of Labor Statistics
- Bureau of Economic Analysis
- Energy Information Administration

Protocol for other agency survey data
- Agency retains control of all approvals and clearances
- Merged data possible with both agency approvals
Role of the RDCs in NCHS/AHRQ Research

NCHS & AHRQ operate their own RDCs in Hyattsville, MD and Rockville, MD, respectively

- Project approvals are required
- Researchers either submit jobs remotely or work in Maryland
- NCHS or AHRQ conduct disclosure analysis

The Census RDC network provides local access points for confidential NCHS or AHRQ files for approved research

- A data extract is transferred securely to the Census Bureau
- RDC users obtain SSS and are given a computer account on the Census server.
- Analysis is conducted in a Census RDC
- Agency approval still required, and the agency still conducts disclosure analysis
NCHS Data

National Health and Nutrition Examination Survey (NHANES) I, II, and III
National Health Interview Survey (NHIS)
National Home and Hospice Care Survey (NHHCS)
National Hospital Ambulatory Medical Care Survey (NHAMCS)
National Hospital Discharge Survey (NHDS)
National Nursing Home Survey (NNHS)
Lots of others…
AHRQ Data

Medical Expenditure Panel Survey

-Household Component
-Linked Household Component-Insurance Component
-Nursing Home Component
-Medical Provider Component
-Two-Year, Two-Panel File
The Census RDC Network

- The Center for Economic Studies (CES) oversees the Census RDC network

- Census RDCs are located around the country near major research universities
  - Washington DC (Census HQ)
  - Boston (NBER)
  - UC Berkeley and UCLA
  - University of Michigan (ICPSR)
  - Chicago Consortium (The Chicago Fed)
  - New York Consortium (Baruch & Cornell)
  - University of Minnesota (open soon)

- Census RDCs are remotely connected to Census servers
Costs and Funding

• Standard rates (based on Census HQ RDC fee per project)
  - Full-time access for a purchased seat = $50,000
  - Typical project using the lab one day a week = $10,000/yr
  - Additional fees may be charged if the project is resource-intensive

• The TCRDC has now shifted to the consortium model
  - Following the lead of the Chicago & New York RDCs
  - Fees waived for UNC system researchers and affiliates
  - Duke and UNC system agreement formally started Fall ‘08
Why is Microdata Access Restricted?

• Title 13 U.S.C. protects confidentiality
  -Identifying microdata cannot be released
  -Only Census employees & temporary staff have access

• Census collects data containing tax information
  -Title 26 (IRS) protects tax data (usually business data)
  -Additional reviews for projects are required

• Confidentiality protection requires:
  -SECURE RESEARCH ENVIRONMENT
  -STRICT ACCESS AND OUTPUT CONTROLS
Research Output and Disclosure

• No output released until approved by the CES Disclosure Officer
  - Release of intermediate output discouraged
  - Output should emphasize statistical models instead of tabular output
  - Results must stay within project scope
  - Disclosure Review Board

• Penalties up to $250,000 and/or 5 years in prison per disclosure for Title 13 data
How to Apply

• www.ces.census.gov

• Preliminary Proposal must meet basic requirements
  - Need for *Non-Public* data
  - Maintains Confidentiality
  - Feasibility
  - Scientific Merit

  ***Describe Census Benefits***
  (LEGAL REQUIREMENT)

• Work with the RDC Administrator to craft a final proposal
  for submission to Census
Why Apply?

- Competition for funding
  - Many research institutions now have reasonable access to an RDC
  - Projects that can demonstrate access to “the most detailed data available” should have an advantage

- Obtaining both lab fees (grant applications) and Census approval (project proposal) often creates a double hurdle
  - For researchers in the UNC system, Census and grant applications can now occur in parallel
  - Census-approved projects may be looked at more favorably by funding entities

- NSF and other funding agencies are increasingly aware of the RDCs

- TCRDC RAs are available to assist new projects at no cost
  - Priority will be given to start-ups and remote researchers
Timeline

• Census Data
- plan on 6 to 12 months
- Title 13 data (Census approval only) vs. Title 26 data (Census & IRS approval)

• NCHS/AHRQ Data
- timeline dependent on agency approval process
  (typically much quicker than approval for a Census project)
- Census approval NOT required
Contact Information: Triangle Census RDC

Gale Boyd - Director,
Phone:  (919) 660-6892
E-mail:  gale.boyd@duke.edu

Bert Grider – Administrator
Phone: (919) 660-6893/919-597-5123
E-mail: grider@email.unc.edu

TCRDC website
http://econ.duke.edu/tcrdc/

CES website
www.ces.census.gov