Census Research Data Centers: NYCRDC Baruch

Rosemary Hyson
U.S. Census Bureau
NYCRDC Baruch
May 29, 2007
Outline

Census RDCs and applied research
Data available
Proposal process
Working at NYCRDC Baruch
Further information
Census RDCs and applied research

Census RDCs

- are partnerships between the Census Bureau’s Center for Economic Studies (CES) and research institutions
- give researchers access to confidential Census Bureau microdata for approved research projects
- benefit the Census Bureau by increasing the utility and quality of Census data products
Census RDCs and applied research

Census RDCs offer qualified researchers

- access to Census Bureau microdata
  - covering businesses, individuals and households
  - non-public data
    - no PUMS equivalent
    - containing information excluded/aggregated in public data
- opportunity to enhance Census Bureau programs
Census RDCs and applied research

Center for Economic Studies (CES)
  – Research unit at Census, created in 1982 to provide secure access to confidential microdata

RDCs
  Boston 1994       Michigan 2002
  UC Berkeley 1999  Chicago 2002
  UCLA 1999        New York Cornell 2004

www.ces.census.gov
Census RDCs and applied research

Why is access to microdata restricted?
Public perception of improper use could
  – reduce response rates
  – induce Congress to cut funding or entire program

Title 13 U.S.C. protects confidentiality
  – identifying microdata cannot be released
  – only Census employees and temporary staff can look at individually-identifiable data
  – this access must potentially provide legitimate benefits to Census Bureau programs
Census RDCs and applied research

“Without the availability of micro data on establishments it is unlikely that the work of Dunne et al (1989) or Davis and Haltiwanger (1992) would have had their current intellectual influence.”

-- Daniel Hamermesh, 2006
Census RDCs and applied research

“What we need is more empirical work. . . . This greater availability of data [through the Center for Economic Studies] and the encouragement given to all researchers working on the institutional structure of production by the award to me of the Nobel Prize, should result in a reduction in that elegant but sterile theorizing so commonly found in the economics literature on industrial organization and should lead to studies which increase our understanding of how the real economic system works.”

-- Ronald Coase, 1992
“Economic” research
(aka research on businesses)

Job creation & destruction
Firm dynamics
Market structure and productivity
Pollution abatement
Firm location and spillover effects
Trade and productivity
“Economic” research
(aka research on businesses)

Small businesses
Sectoral reallocation
Capacity utilization
Energy use
Technology adoption
Mergers, spin offs and asset sales
Employer provided health insurance
“Economic” research

Recent CES Discussion Papers

What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns (CES-WP-07-13)
  William Kerr, Edward Glaeser, Glenn Ellison
Estimating the Distribution of Plant-Level Manufacturing Energy Efficiency with Stochastic Frontier Regression (CES-WP-07-07)
  Gale Boyd
Efficiency Implications of Corporate Diversification: Evidence from Micro Data (CES-WP-06-26)
  Ekaterina Emm, Jayant Kale
The Production Decisions of Large Competitors: Detecting Cost Advantages and Strategic Behavior in Restaurants (CES-WP-06-19)
  Clarissa Yeap
Older Workers' Access to Employer-Sponsored Retiree Health Insurance, 2000-2004 (CES-WP-07-12)
  Elaine Zimmerman, Alice Zawacki, Christine Eibner
Firms in International Trade (CES-WP-07-14)
  Andrew Bernard, J Bradford Jensen, Stephen Redding, Peter Schott
Demographic research

Poverty and inequality
Migration
Immigrant and neighborhood segregation
School quality and housing value
Demographic research

Social networks and labor markets
Crime rates
Female labor supply
Welfare reform
Smoking behavior
Linked Economic and Demographic research

LEHD projects
Human capital spillovers
Segregation in the workplace: race/ethnic, sex, and skill
Economic data available at the RDCs

Annual Capital Expenditures Survey
Annual Survey of Manufactures
Business Expenditure Survey
Commodity Flow Survey
Economic Census
  Construction
  Finance, Insurance, & Real Estate
  Manufactures
  Mining
Retail Trade
Services
Transportation
Wholesale Trade
Economic Census of Puerto Rico
Exporter database
Foreign Trade Data – Imports/Exports
LEHD Employer Characteristics, QWI
Longitudinal Business Database
Manufacturing Energy Consumption Survey
Medical Expenditure Panel Survey (IC)
National Employer Survey
Quarterly Financial Reports
Research and Development Survey
Survey of Business Owners
Survey of Manufacturing Technology
Survey of Plant Capacity Utilization
Survey of Pollution Abatement Costs and Expenditures

See the complete list with descriptions at www.ces.census.gov
Linking economic data at the RDCs

Longitudinal Business Database (LBD)

- all establishments with employees
- history for each, 1976 to 2005
  - age/tenure
  - industry
  - geography
  - employment, payroll
  - establishment births, deaths
- can link LBD history to all other Census Economic data to create panel data
- Longitudinal Research Database was a precursor; LRD was the “panel” constructed for manufacturing sector
Linking economic data at the RDCs

Standard Statistical Establishment List (SSEL)
- link external data to Census economic data
- aggregate establishments up to a firm

SSEL-Compustat® bridge (2001)
- crosswalk to match Compustat IDs to SSEL & other Census Economic data

SSEL-LEHD bridge
- links LEHD employer data to SSEL & other Census Economic data

Other linkage using geography, industry, etc.
Demographic data available at the RDCs

American Community Survey
American Housing Survey
1990, 2000 Decennial Long Form
Current Population Survey March Supplements
National Longitudinal Mortality Study
National Longitudinal Survey (original cohorts)
Survey of Income and Program Participation

...linkage is possible at detailed levels of geography
Data soon to be available

Historical Decennial Data (1970, 1980)
Integrated Longitudinal Business Database
Medical Expenditure Panel Survey-Household Component
Survey of Business Owners (2002)

Interested in other Census Bureau data?
Talk to the RDC Administrator
National Center for Health Statistics data

NCHS RDC data now available through the Census RDCs

NCHS-data-only proposals for projects at a Census RDC
  – do not need to be approved by Census, only NCHS
  – do need to be entered in Census Bureau system
  – if approved, researchers must also obtain SSS

Proposals combining NCHS and Census RDC data must also meet Census Bureau criteria
NCHS Data

National Health & Nutrition Examination Survey I, II, III
National Ambulatory Medical Care Survey
National Hospital Ambulatory Medical Care Survey
National Survey of Ambulatory Surgery
National Hospital Discharge Survey
National Nursing Home Survey
National Home and Hospice Care Survey
National Employer Health Insurance Survey
National Health Provider Inventory
National Health Interview Survey 1967-2005
National Immunization Survey
Longitudinal Study on Aging
National Survey of Family Growth
Vital Statistics (Birth, Mortality, Marriages and Divorces, Fetal, Death National Death Index)

State and Local Area Integrated Telephone Survey
- Health
- Child Well-Being and Welfare, 1997
- National Survey of Early Childhood Health
- National Survey of Children with Special Health Care Needs
- National Survey of Children's Health
- National Asthma Survey
- National Survey of Children with Special Health Care Needs
NCHS data (cont’d)

National Health Interview Survey with Mortality Data 1986-2000
National Health Interview Survey with Medicare Enrollment and Claims Data 1991-2000
National Health Interview Survey with Social Security Administration Retirement, Survivors, and Disability Insurance Data 1962-2003
National Health Interview Survey with Social Security Administration Supplemental Security Income Data 1974-2003
National Health and Nutrition Examination Survey I Epidemiologic Follow-up Study with Mortality Data 1971-2000
National Health and Nutrition Examination Survey I with Social Security Administration Retirement, Survivors, and Disability Insurance Data 1962-2003
National Health and Nutrition Examination Survey II with Mortality data 1976-2000
National Health and Nutrition Examination Survey II with Medicare utilization and expenditure Data 1991-2000
National Health and Nutrition Examination Survey III with Mortality Data 1988-2000
Proposal process

Proposal criteria
- Benefits to the Census Bureau
- Scientific merit
- Need for non-public data
- Disclosure risk
- Feasibility

Review Process
- Internal reviewers (Census employees)
- External reviewers (subject-matter experts)
Proposal process

Talk to the NYCRDC Administrator

Proposals include:
1. Preliminary information entered on CES web site
2. Project description and abstract
3. Statement of benefits to Census
4. Researcher accounts on CES web site
Benefits to the Bureau

Benefits Statement

- "The predominant purpose of this study is to increase the utility of Title 13, Chapter 5 data of the Census Bureau by meeting the criteria listed below and as described."

Benefits criterion listed on CES web site—examples:

- Understanding and/or improving data quality
- Enhancing the data
- Identifying shortcomings and/or documenting new data collection needs
- Preparing estimates of population and characteristics of population

At project end, technical memo and working paper (part of contract)
Funding for access to NYCRDC

Currently, use is free if researcher is part of member institution
NSF-funded projects can use NYCRDC free until May 2009
Non-members can use NYCRDC for a fee
  – Build into grant proposals
Working at NYCRDC Baruch: RDC Computing

RDC data are accessed via “thin client” Linux/UNIX
SAS, Stata, Gauss
Documentation for data is on servers and available from RDC Administrator
Secure space has cubicles and drawers to lock up project materials
Research assistance is available for projects from member institutions (some limitations)
Working at NYCRDC Baruch: Disclosure and confidentiality

No output allowed out of the RDC unless it has been cleared for disclosure
  – Release of intermediate output discouraged
  – Projects must emphasize statistical models vs. tabular output

Penalty for advertent or inadvertent disclosure ($250,000 and/or 5 yrs in prison)

Only researchers with approved projects and Census Special Sworn Status can enter
NYCRDC Baruch

Centrally located at Baruch College

NYCRDC Baruch
135 E 22nd St, Room 801
New York, NY 10010

Computers are available outside the secure RDC area to access the internet and publicly available research-related materials.
NYCRDC Baruch

Our funders

National Science Foundation
Office of the Manhattan Borough President

Member institutions:

Baruch College
City University of New York
Columbia University
Cornell University
Federal Reserve Bank of New York
Fordham University
National Bureau of Economic Research
New York University

Pace University
Princeton University
Russell Sage Foundation
Rutgers University
Stony Brook University/SUNY
University at Albany/SUNY
Yale University
More Info

www.ciser.cornell.edu/NYCRDC/home.shtml

www.ces.census.gov

NYCRDC Baruch RDC Administrator:
Rosemary Hyson
rosemary.t.hyson@census.gov
646-660-6788

NYCRDC Baruch Executive Director:
Sanders Korenman
Sanders_Korenman@baruch.cuny.edu
Resources for learning about RDC data and projects

CES web site
- data descriptions
- discussion papers
- Annual Report
- CES seminar series
- NCHS

NYCRDC web site
- Helpful Links

Census web site
- survey/program pages
- Census program contacts

RDC and CES staff
RDC Researchers
Cornell/CISER’s Virtual RDC
Information sources to help identify benefits to Census

- Potential RDC Methodological Topics
- Census Bureau program contacts
- CES Annual Report
- CES Discussion Papers
- Research Opportunities at the Census Bureau
- Summary of Census Bureau Research Problems Identified by Senior Staff
- RDC Administrator