California Census Research Data Center (CCRDC)
The Future for Social Science Research

• Social science data increasingly comes with security concerns and a limited release for public use.

• Secure research data centers provide the possibility for access to uncensored social science data.

• Data centers are social science using a ‘large scale’ resource - making secure data available to all.
What is a Census Research Data Center?

- A secure physical location where researchers may obtain access to non-public versions of Census Bureau microdata sets and other administrative records
- RDCs incorporate physical security measures, strict operating procedures, and legal safeguards to prevent disclosure
CCRDC Provides

• Research opportunities to California social science researchers.
• Opportunities that would otherwise not have/would not exist without CCRDC.
• Access to, and confidence with, highly confidential survey and administrative data. Working with confidential data resources (security, access, infrastructure) and archiving of work with these data.
CCRDC Organization

- Partnership with US Census Bureau
- Two laboratory sites: UCLA (Southern California) and Berkeley (Northern California).
- Oversight Board
- Project Review Boards
- Faculty Coordinators / P.I.s
- Research Director
- RDC Administrators (US Census Bureau employees)
Census Microdata Available

- http://www.ces.census.gov/index.php/ces/1.00/researchdata
- Examples of which:
  - American Community Survey
  - American Housing Survey
  - Annual Survey of Manufactures
  - Census of Manufactures (plus other industries: Mining, Construction, FIRE, Retail Trade, Services, Transport, Wholesale)
  - Compustat-SSEL Bridge
  - CPS Crosswalk (PIK)
  - Current Population Survey March Supplement
  - Decennial Census Long Form Sample
  - Foreign Trade Data (Import/Export)
Census Microdata Available

- Examples of which:
  - LEHD Business Register Bridge
  - Medicaid Eligibility Database System – California
  - National Center for Health Statistics Data Extract (http://www.cdc.gov/nchs/r&d/rdc.htm)
  - National Longitudinal Mortality Study
  - Ownership Change Database
  - SIPP Public-Use Crosswalk
  - Survey of Income and Program Participation (Panel/Longitudinal)
  - Survey of Manufacturing Technology
  - Survey of Pollution Abatement Costs and Expenditures
  - Unemployment Insurance-Base Wage File -- California
Research

• 8-10 Projects active at any one time.
Research

• Research with State Agencies: California Health and Human Services (CHHS), Department of Health Services (DHS), Employment Development Department (EDD) and Labor Market Information Division (LMID).

• Matching data and creating new research resources.

• Research of direct value to, and in cooperation with, the State.
Research

- Examples of projects (in last 5 years):
  - California Immigrant Families and Welfare Reform
  - A Proposal to Examine the Reporting of Welfare Benefits in the SIPP Using Matched Administrative Records from California
  - Teacher Quality and Wage Compression
  - Analysis of Small Firms Serving a Client Base of Minority Households
  - Productivity Differences Across Firms and Countries
  - Measuring the Productivity Impact of Wireless Technology
  - Corporate Restructuring and Earnings Inequality
  - The Impact of Minimum Quality Standards on the Child Care Market
NCHS data

- Only just become available (MOU signed this year!).
- Proposals will be accepted by CES beginning February 15, 2007.
- NCHS will handle all proposal review and disclosure avoidance review for their data.
- Charge a small fee for creating the researcher's data extract.
- Standard Census RDC access charges, if any, will apply.
- Proposals to use their data will not have to meet Census Bureau standards, just NCHS standards.
- Proposals to use confidential data from both the Census Bureau and the NCHS must use the existing Census Bureau application procedure, and will also be reviewed by NCHS.)
NCHS data (Examples)

- National Health and Nutrition Examination Survey I, II, and III
- National Ambulatory Medical Care Survey
- National Hospital Discharge Survey
- National Employer Health Insurance Survey
- Longitudinal Study on Aging
- National Survey of Family Growth
- State and Local Area Integrated Telephone Survey
- Vital Statistics

Linked data (examples):
- National Health Interview Survey with Mortality data 1986-2000
Proposals

• http://www.ces.census.gov/index.php/ces/1.00/researchprogram
• Proposal MUST be of benefit to the Bureau. Legal requirement.
• Pre-proposal (about 2 paragraphs max.).
• Proposal in four parts: Abstract/Main Proposal/Benefit statement/CV.
• Proposal to be referred by Census and RDC referees: scientific merit and benefit to Bureau.
• Further IRS approval required on economic data (Title 26 data).
Proposals

- Proposal guidelines (approximation): Abstract (1 page); Proposal (10-12 pages); Benefit (3 pages).
- Benefit to fit within certain (Title 13, Ch 5) criteria.

Examples:
- Understanding and/or improving the quality of data
- Improving imputations for non-response;
- Developing links across time or entities for data gathered in censuses and surveys authorized by Title 13, Chapter 5;
- Identifying shortcomings of current data, collection programs and/or documenting new data collection needs;
- Constructing, verifying, or improving the sampling frame for a census or survey authorized under Title 13, Chapter;
Proposals

• Starting from the idea: use Administrator for contact within Census.
• Concentrate on science and Bureau work. Stay away from politics and policy!
• Risk of disclosure: results have to be disclosed by Census.
• Submit at anytime via the website.
• Return hopefully within 6 – 8 weeks (depending). One of four responses:
  – Approve.
  – Clarification Required. (minor requirements/ no resubmission).
  – Revise and Resubmit. fails on one or more of the major criteria.
  – Reject.