

Conducting Research in the Penn State Federal Statistical Research Data Center

Emily Greenman, PhD
Administrator, PSUFSRDC
Center for Economic Studies
U.S. Census Bureau

Disclaimer

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

Demographic and Health Data in the FSRDCs

- Census Bureau Demographic Data
 - Decennial Census and American Community Survey (ACS)
 - Survey data on individuals and households
- Employer-Employee Jobs Data (LEHD)
 - Data on employees linked with data on employers
- Health Data
 - National Center for Health Statistics (NCHS)
 - Agency for Healthcare Research & Quality (AHRQ)

Advantages of Restricted Data

- More detailed level of geographic identifiers
- Very little top or bottom-coding
- Census datasets can be linked together
- Census datasets can be linked to external data

Demographic Datasets

- Decennial Census 1970, 1980, 1990, 2000, 2010
- American Community Survey (annual microdata, 1996-2014)
- Current Population Survey (selected supplements)
- Survey of Income and Program Participation
- American Housing Survey
- National Crime Victimization Survey
- Administrative data from SSA
- **National Survey of College Graduates**
- National Longitudinal Mortality Study

Demographic Example

“Residential Mobility Across Local Areas in the United States and the Geographic Distribution of the Healthy Population.”
Geronimus & Bound 2014 (Demography).

- Used 2000 Decennial Census microdata including current and former place of residence measured at the zip code level.
- Investigate whether “unhealthy” local areas are an artifact of less-healthy people moving into them.
- Find that residential mobility does not tend to change the overall health composition of local areas.

Penn State Research

- Monnat: “Hispanic Health Care Access and Utilization in Different Geographic Locations”
 - Restricted-use SIPP files contain county IDs, enabling a more detailed examination of health care access and use across Hispanic destination types.

Health Data in the RDC

- These data are collected by:
 - National Center for Health Statistics (NCHS)
 - Agency for Healthcare Research and Quality (AHRQ)
- Dual mission: to provide broad access to health data and statistics, while protecting the privacy of respondents

What Types of NCHS Data?

National Health Status Surveys

- National Health and Nutrition Examination Survey (NHANES) I, II, and III
- National Health Interview Survey (NHIS)
- Longitudinal Study on Aging I and II (LSOA)
- National Survey of Family Growth
- National Survey of Children's Health
- National Survey of Early Childhood Health
- National Survey of Children with Special Health Care Needs
- National Asthma Survey

National Health Care Surveys

- National Ambulatory Medical Care Survey

- National Hospital Ambulatory Medical Care Survey
- National Survey of Ambulatory Surgery
- National Hospital Discharge Survey
- National Nursing Home Survey (NNHS)
- National Home and Hospice Care Survey
- National Employer Health Insurance Survey
- National Health Provider Inventory
- National Immunization Survey

Vital Statistics

- Mortality and Multiple Mortality
- Birth
- Fetal Death
- National Death Index
- Marriage and Divorce

What Types of NCHS Data?

Linked Data Sets

- Linked mortality data: NHIS, NHANES, LSOA II, NNHS
- Linked Medicare Enrollment and Claims data: NHIS, NHANES, LSOA II
- Linked Social Security Administration Data: NHIS, NHANES, LSOA II, NNHS
- Linked EPA data

What's Restricted?

- Geographic variables (state, county, or metropolitan area)
- Most dates (date of interview, date of death, date of birth)
- Income and employment data (detailed industry/occupation codes)
- Specific diagnoses (ICD-9 codes are generally coarsened)
- Details about facilities (accreditation, payments, number of employees)
- Some information about children and adolescents (e.g. height and weight, depression, behavior problems, and drug use)
- Some information about race, ethnicity, and country of origin
- Contextual data (nearest hospital, % of population with diploma)
- Sample design variables (necessary for estimating variances)

NCHS Examples

- NHANES-Mortality: Individuals in the lowest quartile of bone mineral density had a 1.53 times greater risk of death than those in top quartile (Mussolino and Gillum 2008)
- NHIS: Military service increased the probability of smoking, but in later adulthood, the effect attenuated and did not lead to large negative health effects (Eisenberg and Rowe 2009)

Penn State Research

- Van Hook et al.: “Neighborhood Context, Weight, and Weight-Related Behaviors among Mexican American Children”
 - Restricted-use NHANES files contain county and tract IDs, as well as detailed place of birth, food security, and occupation.
 - Preliminary findings suggest that Mexican-origin children who live in neighborhoods with higher concentrations of foreign-born Mexicans had lower levels of dietary acculturation, but dietary acculturation was higher among children living in neighborhoods with higher percentages of non-Hispanic whites and persons with low educational attainment.

What types of AHRQ Data?

- Medical Expenditure Panel Survey (MEPS) files include:
 - Household Component
 - Provider Component
 - Insurance/Employer Component
 - Nursing Home Component (1996 only)
 - Area Resource File
 - Two-year two panel file
 - MEPS-NHIS linked data
- Only Household Component and portions of Provider Component are publicly available

AHRQ Examples

- MEPS-IC and Economic Census:
- Matched MEPS with Census Bureau economic data to add information on firm characteristics
- Employers who offer health insurance have 25% greater productivity and 32% higher pay, all other variables held constant (McCue and Zawacki 2006)

How to Access the RDC

- Develop proposal
 - Different guidelines for Census data vs. NCHS/AHRQ guidelines
 - For Census data, begin by contacting Emily with a brief project summary
 - For NCHS/AHRQ, contact Emily as soon as you submit a proposal to NCHS/AHRQ.
- Submit proposal for agency review
 - Census (and agency sponsors)
 - NCHS/AHRQ
- Obtain Special Sworn Status (SSS)
- Pay one-time fee for NCHS/AHRQ data

Guidelines for Census proposal

- Proposal must meet Basic Requirements
 - Need for *Non-Public* data
 - Maintains Confidentiality
 - Must emphasize statistical models vs. tabular output
 - Feasibility
 - Describes Census Benefits (LEGAL REQUIREMENT)
 - Scientific Merit
- Work with Census Administrator to Craft Final Proposal

Working in the RDC lab

- All analysis conducted in the RDC lab:
 - Data located on server in Maryland
 - Access data via thin client terminals located in cubicles
- No internet access or personal computers allowed in lab
- Statistical software available: SAS, Stata, R, Matlab, Sudaan, etc.
- Agency reviews output before releasing
 - Penalty for disclosure is \$250,000 and/or 5 yrs in prison (inadvertent or otherwise)

Timeframe – “Patience is a Virtue”

- Census Data
 - Plan on 4 to 12 months for proposal review
 - Title 13 (Census approval only) vs. Title 26 (Census & IRS approval)
- NCHS/AHRQ Data
 - Timeline dependent on agency approval process
 - Census approval automatic and quick after agency approval
- Special Sworn Status
 - 2-4 additional months for your “security clearance”

Contact Information

- People:
 - Emily Greenman, PSUF SRDC Administrator
emily.k.greenman@census.gov
 - Mark Roberts, PSUF SRDC Executive Director
mroberts@psu.edu
- Resources:
 - PSUCRDC website:
<http://www.psurdc.psu.edu/>
 - CES Working Paper Series:
<https://ideas.repec.org/s/cen/wpaper.html>
 - CES FSRDC website:
<http://www.census.gov/fsrdc>

Resources & Links for Health Data

- AHRQ restricted data & application info:
http://meps.ahrq.gov/mepsweb/data_stats/onsite_datacenter.jsp
- NCHS restricted data:
<http://www.cdc.gov/rdc/b1datatype/dt100.htm>
- NCHS application process:
<http://www.cdc.gov/rdc/b3prosal/pp300.htm>
- Center for Economic Studies Annual Research Reports (list all approved RDC studies and publications, including NCHS/AHRQ):
<http://www.census.gov/ces/publications/annualreport.html>

Resources & Links for Health Data

- NCHS Data Linkage information
- http://www.cdc.gov/nchs/data_access/data_linkage_activities.htm