



CHAPIN HALL

CENTER FOR CHILDREN
AT THE UNIVERSITY OF CHICAGO

Linking administrative data and the types of questions these data can address: How to get "your feet wet"

Robert M. Goerge
Chapin Hall Center for Children

Use of Administrative Data Systems for Program Monitoring, Research and Evaluation
August 1, 2007

**SAMI and CCPRC Institute
Washington, D.C.**



CHAPIN HALL
CENTER FOR CHILDREN
AT THE UNIVERSITY OF CHICAGO

**Illinois/Chapin Hall Center for Children
at the University of Chicago
(Project Manager)**

Robert Goerge, Allison Harris, Lucy Bilaver

**The Ray Marshall Center at the
University of Texas at Austin**

Deanna Schexnayder, Daniel Schroeder

**Maryland/The Jacob France Institute at the
University of Baltimore**

Jane Staveley, David Stevens

NCCP at Columbia University

Lee Kreader

The U.S. Census Bureau

Sally Obenski, Ron Prevost, Dean Resnick



What is linked administrative data good for?

- Used to address questions about comprehensive set of services and programs in which individuals and their families participate
- Individual's records across services and programs are linked. Individuals are linked to others through case or household membership
- Longitudinal, individual-level - which allows for the a rich set of analyses that one may be interested in doing
- Contains addresses of residences and service providers



Administrative Data Use for Research

The basic principle is that you go to the best source of data

A cost effective way of enriching the available data for research through linking

As an ongoing research resource: data is always being collected regardless of researcher's attention to it

Importance of linking population-level data sets



Identifiers used for matching

- Depends on what is available in two datasets
- Typically use name, birthday, race/ethnicity, gender, address, SSN
- As you add more datasets, you can use accumulated data to conduct your matches
- Census/SSA example
- Our example using welfare records to get SSN



Background

- Political environment
- Service recipients are dependent and resources are scarce, so programs are at-risk of cancellation
- Conflict
 - Leaders - Avoid “negative” information
 - Researchers - Disseminate research results
- Compromise is needed



Potentials and Pitfalls

- Advantages
 - Ideal for studying program participation
 - Universe of participants, allows for geographic analyses
 - Collecting sensitive information
 - Data is “in play” -- it is constantly updated



Disadvantages

- Choice, event or participation-based
- Only for period of participation
- Difficult to access
- Lack of documentation



Assessing the Quality of the Data

- One usually gets some kind of extract
- Others may or may not be familiar with this data
- Data must be assessed dataset by dataset
- And, probably field by field



Creating longitudinal files

Access up-to-date data to fill-in event-history

Know what data is updated, when and how

Data are purged or archived

There are lags both in data entry and processing of records (Medicaid)

There are corrections made in the data as both service recipients and providers see errors



Recent Welfare Research

- National or multi-state studies
- SIPP, which is “morphing,” and never really gave us info at the state level
- TANF Leaver studies
- MDRC evaluations (MN)
- 3 city study and other investigator-led studies (FF, IFS)



Problems

- This research effort is not sustainable
- Does not take advantage of the massive data collection efforts now underway



Problems with relying on surveys for participation

- Child care subsidy use is short in duration
- Families start and stop use often
- Families may not know what funding source is being used
 - Head Start and State Pre-k combined funding
- Recall of program use is short



Problems with relying on surveys for eligibility

- Unless you ask all of the questions on application, always uncertain
- Recall about income and employment is poor



Under-reporting of participation

- Bitler, Currie, and Scholz (2003) discovered an undercount of participation for WIC, AFDC/TANF, Medicaid, and the FSP in both the CPS and the SIPP
- Cody and Tuttle (2002) reach similar conclusions.
- Meyer and Sullivan (2006) show that the fraction of FSP dollars that are not reported in the CPS increased from 33 percent to 40 percent between 1993 and 2002.
- Marquis and Moore (1990) show that 23 percent of survey respondents who were food stamp recipients according to administrative microdata, failed to report participation in the 1984 SIPP.
- U.S. Census (2004), which examined food stamps administrative records in Maryland linked to the national 2001 Supplementary Survey (American Community Survey), also found evidence of under-reporting of program participation.
- Jane Stavely yesterday reported undercount of MD TANF caseload in 2001 ACS. (Stavely, Larson, Lynch, Resnick)



Must rely on administrative data for wages and program participation

- Only feasible way to get longitudinal, accurate data on child care subsidy receipt
 - Collected in order to pay vendors
 - Can link to other data
 - Provides data on parent, child, and child care providers
- Major challenges in using it for monitoring and analysis



METHODOLOGY

- Using the Census 2001 Supplementary Survey (SS01), we define a base population of child care subsidy eligible families in 2000-01. The SS01 uses the American Community Survey questionnaire and methods to collect detailed social, economic and housing data from yearly samples of over 800,000 addresses nationwide. The SS01 was conducted in 2001 and interviewed a new sample of households in each month of the year. Questions on income asked about income in the past 12 months.
- We prepare state databases of child care subsidy and TANF receipt and quarterly UI wage data.
- The state databases are combined with the SS01 and used to track quarterly child care subsidy eligibility, child care subsidy use, and employment and welfare outcomes for different groups of low-income families through 2003.

The Base Population

All CCS-eligible families, with children under age 13, in Illinois, Maryland and Texas, in 2000-01, from the SS01. (The SS01 sampled households in each month of 2001 and asked about income in the prior 12 months.)



We need data on target population

- TANF. Medicaid, food stamps or other program populations
- General population
 - Birth certificates (?)
 - Problem with mobility
- That leaves us with survey data



Motivation for using Census data

- Not all households who are eligible for social programs are in administrative data
- Need to have a source of general population data to understand take-up and participation
- While small surveys have provided some information on who does and does not take-up, more detailed information is needed to make policy and programmatic changes--preferably at the county or smaller level, where decisions are made.



Motivation for using Census data (cont'd)

- ACS and decennial Census provide large sample sizes that give us statewide and small region estimates
- Beginning in 2005, very large samples that allow estimates of places over 60,000 individuals
- No data collection costs to states or researchers, although there is minimal cost for analyzing the data



Motivation (cont' d)

- Census survey data provides a large sample of the general population of households and, therefore, a large sample of low-income households that may be eligible.
- Sending state-level administrative data to the CB allows data security and confidentiality to be maintained while the power of combining data is realized.



Taking up the benefit

- Of all the eligible families, how many participate in the child care subsidy program?
- Answer TBA



Determination of eligible population

- Using ACS Public Use Data
 - » i. Microdata files from the American Community Survey show the full range of responses made on individual questionnaires. The files contain records for a sample of all housing units, with information on the characteristics of each unit and the people in it.
<http://www.census.gov/acs/www/Products/PUMS/>
 - » ii. All identifying information is removed to ensure confidentiality. The questionnaire included questions on age, sex, tenure, income, education, language spoken at home, journey to work, occupation, condominium status, shelter costs, vehicles available, and other subjects.
<http://www.census.gov/acs/www/Products/PUMS/>



Determination of eligible population - TANF

- Categorically eligible are included in this data.
- In some TANF data systems, the income from various income sources will be reported.
- Also, the observed quarterly earnings from the UI wage report data will be used to identify those who are eligible (working with quarterly earnings lower than the state eligibility income level) after their entry to TANF.



Determining who is using the subsidy

- **801 data**

- » This administrative dataset provides descriptive information about the families and children served through the federal Child Care and Development Fund (CCDF). CCDF dollars are provided to states, territories, and tribes to provide assistance to low-income families receiving or transitioning from temporary public assistance, in obtaining quality child care so they can work, or depending on their state's policy, attend training or receive education.



Determining who is using the subsidy

- **Child care subsidy administrative data**
 - The source data in each state is a database constructed out of existing administrative data on child care subsidy receipt. Each state's subsidy system records monthly subsidy information as well as subsidized families' basic characteristics and child care arrangements. Each state's database contains longitudinal information on child care subsidy receipt on a monthly basis at individual family and child level.



Effect on employment

- Of those who are eligible, are those who receive a child care subsidy more likely to stay employed?
- Answer TBA



Who is employed

- UI wage data
 - » Unemployment Insurance (UI) wage records consist of total quarterly earnings reported by employers to state UI agencies for each employee. The database contains information on quarterly earnings, employee SSN, employer SSN, and employer address. In each of the proposed states, we have access to a longitudinal wage reports data covering the entire state population at least from 1997 to the present. Any employer paying \$1,500 in wages during a calendar quarter to one or more employees is subject to a state UI tax and must report the quarterly amount paid to each employee.



What we hope to add to the existing knowledge base through our project

- We will estimate CCS eligibility and take-up for all low-income families.
- We will examine employment and TANF family outcomes to determine the impact of child care subsidies.
- We will examine whether the decision to use the subsidy has different self-sufficiency ramifications for different groups of low-income families, including the working poor who have no history with the TANF program.
- The more detailed socioeconomic and demographic information available to us in the census data will enable us to develop better models of the relation between take-up and employment and welfare outcomes.



Data sharing agreements

- University research centers had permission to use administrative data
- States had to provide permission to research centers to share data with Census Bureau
- University research centers entered into agreement with Census Bureau to transfer data to CB
- States have “rights” to review results of work



Illinois Preliminary Results

- Began with 40,000 (16,540 households) individuals over 14
- 12,000 had children under the age of 13
- 2,000 families were eligible for CCS during the quarter of ACS interview or 1st quarter after
- In 2005, we jumped to 160,000 individuals in nearly 60,000 households
- Over 8,000 eligible families



Next steps in Illinois

- Similar project with food stamps--to determine who is taking up those benefits and how it affects employment.
- Intend to include other programs - EITC, WIC, child welfare programs, free and reduced school lunch, SCHIP/Medicaid -- in order to better understand how our “safety net” is working for families



bobg@uchicago.edu
www.chapinhall.org

