Use of Administrative Data for Research and Evaluation

Deanna Schexnayder

Ray Marshall Center for the Study of Human Resources
The University of Texas at Austin

Presented at

HHS ACF SAMI-CCPRC Conference
August 1, 2007
How to Use Administrative Data for Research

• Identify research questions to be answered
• Assess research method needed
• Assess availability and quality of administrative data
• Collect other data (e.g., survey, policy) not contained in administrative records
• Determine best way to organize data (e.g., relational database or flat files) and to link across files
• Determine sampling strategy
• Create data set(s) needed to answer research questions
Types of Research That Can Be Conducted

- Experimental and quasi-experimental program evaluation
- Demographic characteristics of program participants and families
- Tracking program participation and outcomes over time
- Determination of factors associated with positive program outcomes
Advantages of Administrative Data

- Detail and accuracy of program information
- Large sample size
- Ability to conduct analysis on small states or sub-state areas in large states
- Data on the same individuals over a long period of time
- Ability to expand information by linking across programs
Limitations of Administrative Data

- The number and types of variables available for research may be limited
- Quality and completeness of ‘non-required’ variables may be suspect
- Not well suited for population-based research questions unless combined with survey data
- Computation of statistics across states must be done with caution
Ways to Improve Power of Administrative Data

• Link across programs (e.g., TANF, Medicaid, Food Stamps, child care, education, child support, in-state UI earnings, out-of-state earnings from FPLS)

• Link files over time

• Add survey variables to add detail on topics of interest

• Add economic/geographic variables that are available from published sources
Today’s Examples

• Additional uses of Administrative Data
  – Survey variables added to linked administrative data files (example 1)
  – Analysis of small geographic areas (example 2)

• Interpreting Analysis of Administrative Data Across Jurisdictions (example 2)
  – Differences in populations
  – Differences in policies
  – Changes over time
Example 1:
Texas TANF Leavers Study

- Conducted by Ray Marshall Center and Center for Social Work Research, University of Texas

- Funded by HHS (ASPE) and Texas Department of Human Services

  www.utexas.edu/research/cshr/pubs/TexasFamilies2002pub.htm

- Forthcoming book: *Life After Welfare - Reform and the Persistence of Poverty (December 2007)*
Research Questions

- To what extent are TANF leavers employed and/or receiving other economic supports? (e.g., child support, subsidized child care)

- What types of employment barriers do these families face and how do they manage?

- Which factors are associated with:
  - Employment
  - Returning to TANF?
## Research Methods and Data

<table>
<thead>
<tr>
<th>Population</th>
<th>Time Period of Cohort</th>
<th>Research Approaches and Data Sources Used</th>
<th>Geographic Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>TANF Leavers</td>
<td>April 1998-June 1999 (Cohort 1)</td>
<td><strong>Administrative data</strong> (18 months follow-up)</td>
<td>Statewide universe</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Econometric analysis</strong> (administrative data only)</td>
<td>Statewide universe</td>
</tr>
<tr>
<td><strong>N= 143,491 families</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TANF Leavers</td>
<td>July – Sept. 2000 (Cohort 2)</td>
<td><strong>Administrative data</strong> (6 months follow-up)</td>
<td>Statewide universe</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Telephone/mail survey within 6 months after exit</strong></td>
<td>Statewide universe</td>
</tr>
<tr>
<td><strong>N=23,113 families</strong></td>
<td></td>
<td><strong>Econometric analysis</strong></td>
<td>Statewide sample</td>
</tr>
<tr>
<td></td>
<td></td>
<td>a. administrative data only</td>
<td>Statewide universe</td>
</tr>
<tr>
<td></td>
<td></td>
<td>b. administrative and survey data</td>
<td>Statewide sample</td>
</tr>
</tbody>
</table>
Administrative Data
Used as Framework

• Administrative Data
  – Statewide universe of program data over 5 years
  – Linked data from 9 programs
    TANF, Food Stamps, Medicaid, Choices, quarterly earnings, child support, subsidized child care, child abuse and neglect, foster care

• Survey Data
  – Used to explore a topic in depth:
    details of employment patterns
    child care arrangements not measured by subsidy system
  – Combined with administrative data in regressions
Child Care Strategies

- Less than 25 percent of employed families received child care subsidies

- Survey data used to describe unsubsidized care:
  - Care by TANF leaver
  - Care by relatives
  - Informal arrangements in other homes
  - Self-care by child
## Influence of Child Care on Probability of Employment

For families leaving TANF July-September 2000

<table>
<thead>
<tr>
<th>Event Description</th>
<th>Model 1: Admin. Data State</th>
<th>Model 2: Admin. Data Sample</th>
<th>Model 3 Admin + Survey Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any subsidy after leaving TANF</td>
<td>.136</td>
<td>.229</td>
<td>.240</td>
</tr>
<tr>
<td>Care for youngest child myself</td>
<td>N/A</td>
<td>N/A</td>
<td>-.123</td>
</tr>
<tr>
<td>Youngest child in non-relative’s home</td>
<td>N/A</td>
<td>N/A</td>
<td>.117</td>
</tr>
</tbody>
</table>
Influence of Child Care on Probability of TANF Re-Entry

For families leaving TANF July-September 2000

<table>
<thead>
<tr>
<th></th>
<th>Model 1: Admin. Data State</th>
<th>Model 2: Admin. Data Sample</th>
<th>Model 3: Admin + Survey Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any subsidy after leaving TANF</td>
<td>-.027</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
<tr>
<td>Any subsidy during prior TANF spell</td>
<td>.036</td>
<td>.078</td>
<td>.083</td>
</tr>
<tr>
<td>Youngest child in non-relative’s home</td>
<td>N/A</td>
<td>N/A</td>
<td>.138</td>
</tr>
</tbody>
</table>
Benefit of Linking Administrative and Survey Data

This study conclusion could not have been determined from either administrative or survey data alone:

“Non-relative care in family homes was associated with both increased employment and increased TANF recidivism, suggesting that such care (or the jobs that may cause families to select this type of care) is not stable over time”
Example 2: TX Child Care Devolution Study

• Conducted by Ray Marshall Center and Center for Social Work Research, University of Texas

• Funded by HHS Child Care Bureau and OPRE

• Publications can be found at:
  http://www.utexas.edu/research/cshr/current/devchildcare.htm
Research Questions

• Which combinations of child care subsidy policies did local areas adopt after devolution from the state to the local level?

• Which local policy choices were associated with longer child care subsidy duration?

• Which local policy choices were associated with longer employment duration for families receiving subsidies?
TX Local Workforce Development Areas

1. Panhandle
2. South Plains
3. North Texas
4. North Central
5. Tarrant County
6. Dallas
7. North East
8. East Texas
9. West Central
10. Upper Rio Grande
11. Permian Basin
12. Concho Valley
13. Heart of Texas
14. Capital Area
15. Rural Capital
16. Brazos Valley
17. Deep East Texas
18. South East Texas
19. Golden Crescent
20. Alamo
21. South Texas
22. Coastal Bend
23. Lower Rio Grande Valley
24. Cameron County
25. Texoma
26. Central Texas
27. Middle Rio Grande
28. Gulf Coast
<table>
<thead>
<tr>
<th></th>
<th>Largest (Gulf Coast)</th>
<th>Smallest (Concho Valley)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child Population</td>
<td>1,401,948</td>
<td>38,549</td>
</tr>
<tr>
<td>Funding for Subsidies (FYs 2002 &amp; 2003)</td>
<td>$184 million</td>
<td>$6.1 million</td>
</tr>
<tr>
<td>Total children receiving subsidies (FYs 2002 &amp; 2003)</td>
<td>49,676</td>
<td>2,649</td>
</tr>
</tbody>
</table>

Diversity of Texas Workforce Areas Comparable to Diversity Across States

N=28
Share of Poor Children Served

Share of poor children served (1999)

### Policy Context Across States (2003)

<table>
<thead>
<tr>
<th></th>
<th>IL</th>
<th>MD</th>
<th>OR</th>
<th>TX</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Income eligibility ceiling</strong> (% of SMI)</td>
<td>39%</td>
<td>40%</td>
<td>60%</td>
<td>Varies by local area; 50 - 85%</td>
</tr>
<tr>
<td><strong>Monthly co-pay</strong> (family of 3; 2 children)</td>
<td>$8</td>
<td>$4</td>
<td>$43</td>
<td>Varies by local area</td>
</tr>
<tr>
<td>Minimum</td>
<td>$320</td>
<td>$116</td>
<td>$399</td>
<td>9% of income</td>
</tr>
<tr>
<td>Maximum</td>
<td></td>
<td></td>
<td></td>
<td>15% of income</td>
</tr>
</tbody>
</table>
Texas Child Care Policy Context

Statewide throughout study period:

- TANF Choices recipients have priority for subsidies and are exempt from co-payments
- Child care subsidies are not guaranteed for non-TANF recipients, with waiting lists in some areas

Statewide before devolution in September 1999:

- Income eligibility limits - 150% of FPL or 85% SMI
- Co-payment - 9% of income for 1 child and 11% for 2 or more children
- Reimbursement rates were set at state level but varied based on local market rates
Policy Choices After Devolution

Local Workforce Board Policies: Sept 1999- Aug 2003

N > 28 because some boards changed policies more than once during period

- 5 boards kept baseline policies
- 12 increased payments
- 7 increased payments and eligibility limits
- 8 increased co-pays
- 5 increased eligibility limits
- 4 increased eligibility limit and co-pays
- 3 increased eligibility limits and reduced co-pays
The income eligibility guidelines varied significantly among the 28 local board areas.
Income Eligibility Ceilings Across the State

Income Eligibility Ceilings

The diagram shows the number of LWDBs with income eligibility guidelines in each range from FY2000 to FY2003. Each bar is color-coded to represent different percentage ranges:
- 50% SMI (50% or less of the state median income)
- 51-65% SMI
- 66-80% SMI
- 81-85% SMI
Required Parental Co-Payments

Co-payment policy variation (2000-2001)

Available Data

• Statewide longitudinal data for 6-year period (1997-2003)
  – Child care subsidy participation, demographics and payment
  – Local funding allocations
  – Employment records

• Local subsidy policies, 1999-2003

• Contextual economic and community variables
Research Methods

• Descriptive statistics
  – Families receiving subsidies
  – Patterns of subsidy use
  – Characteristics of local workforce areas

• Cluster analysis to determine variation in local policy choices following devolution

• Cox proportional hazards regression models with time-varying covariates
  – Probability of exit from subsidy
  – Probability of exit from employment
Structuring and Interpreting Regressions

• Need to account for:
  – differences in local workforce board areas
  – differences in population receiving subsidies
  – changes in policies over time and area

• Requires use of complex statistical models

• Requires careful and thoughtful interpretation of results
Suggestions for New Users

1. Start small
   Limit number of files, complexity of questions and geographic scope

2. Use first efforts to learn the nuances of variable meanings, context and limitations

3. Gradually move to more complex topics and files

4. If interested in sophisticated techniques presented today, partner with organization that has experience using administrative data for research
All project descriptions and cited publications can be found at: http://www.utexas.edu/research/cshr/

Deanna Schexnayder
512-471-2193
dschex@uts.cc.utexas.edu