The Nuts and Bolts of Working with Administrative Data:

Linking Data Across Multiple States and Multiple Data Sources

Julia Henly

University of Chicago

October 1, 2015
Outline

• Part 1: Using administrative data for multi-state comparison studies

• Part 2: Using multiple sources of data in combination with administrative data
IL/NY Child Care Research Partnership

• Partnership with Univ. of Chicago & Urban Institute & State Child Care Administrators
  – Funded by OPRE/ACF, Grants #90YE0133 & #90YE0151-01-00

• To gain knowledge about child care subsidy dynamics and factors associated with subsidy (in)stability and its implications for child care quality and continuity

• 4 Targeted Regions
  – New York – Nassau and Westchester Counties
  – Illinois – Cook County and Southwestern IL

• Non-school aged new entrants in subsidy program

• Multi-component, mixed-method study, including admin child care records over 18 months in four regions
Part 1: Using administrative data from multiple states

- Multi-state comparisons can be interesting and useful but present challenges
  - Not only inadequate control of contexts,
  - Programs vary,
  - Data vary.

“States don’t collect the same stuff”

-- Amy Claessens
How do state data sources vary?

• Level of the record
  - e.g., family vs child

• Type of information
  - Differences in state definitions of variables, in what is collected, and in reliability/validity.

• Time Unit
  - Weekly vs Monthly Records: Example from MN/MD Child Care Research Partnership (Liz Davis)
Spells with the same number of weeks may include a different number of months

<table>
<thead>
<tr>
<th>CASE NUM</th>
<th>Month 1</th>
<th>Month 2</th>
<th>Month 3</th>
<th>Month 4</th>
<th># Weeks</th>
<th># Months</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X X X X X</td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>X X X X X</td>
<td></td>
<td></td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>X X X X X X</td>
<td></td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>X X X X X X</td>
<td>7</td>
<td>2</td>
</tr>
</tbody>
</table>

State-Specific Research Agreements and Protocols

- States have different orientations toward research
  - Process for establishing data sharing agreement and what the agreement is varies by state
  - States have different in-house data capacities

- These issues affect not only research plan in each state but timeliness of data access
Even with these challenges and caveats, you can nevertheless learn a lot from multi-state comparisons using administrative data.
Kaplan-Meier Survival Curve of First Subsidy Spell, 18-month window

- New York
  - Nassau: 12 months
  - Westchester: 10 months

- Illinois
  - Cook County: 9 months
  - Southwestern IL: 6 months
Smoothed Hazard Curve of the First Subsidy Spell, 18-month window

- New York

- Illinois

![Graphs showing smoothed hazard curves for New York and Illinois with two regions highlighted. The graphs show probability over time since subsidy start.](image)
Figure 1. Illinois: Patterns of Subsidy Use Over 18 Months (N=5,893)

- Continuous use: 28%
- Exit subsidy: 72%
- Returned in 18-month window: 33%
- Did not return in 18-month window: 39%
- Changed subsidized providers after gap: 36%
- Returned with same subsidized provider: 64%

Figure 2. New York: Patterns of Subsidy Use Over 18 Months (N=1,819)

- Continuous use: 33%
- Exit subsidy: 67%
- Returned in 18-month window: 27%
- Did not return in 18-month window: 40%
- Changed subsidized providers after gap: 30%
- Returned with same subsidized provider: 70%
Part 2: Linking administrative data with additional data sources

• Admin data have advantage of telling story over time for large sample

• But story only as good as the data

• Several advantages to linking admin data to other sources
  – Surveys, qualitative interviews, participant observation, etc.
IL/NY Child Care Research Partnership:
Relationship between Data Components

**Phase 1**
- Family-level child care payment records
  - Parent Telephone Survey
  - Parent In-Depth Interviews

**Phase 2**
- Provider Database (R&Rs)
  - Provider Interviews

Policy Stakeholder IW’s & Document Review
Concluding thoughts

• Many benefits to conducting multi-state studies using administrative data
  – But not as easy as comparing or linking a bunch of files together
  – Expect process to be slow, with political, organizational, and data challenges.

• Combining administrative data with other sources has many advantages, and can clarify and expand knowledge gained through administrative data analysis alone.
It takes a lot of people to do multi-state, multi-method research

- **Program Officers**: Susan Jekielek; Ivelisse Martinez-Beck; Ann Rivera
- **Investigators**: PI: Julia Henly; Co-PI’s: Amy Claessens, UC-Harris & Heather Sandstrom, Urban Institute
- **Key Research Staff**: Alejandra Ros Pilarz, JaeSeung Kim, Carolyn Barnes, University of Chicago; Julia Gelatt & Olivia Healy, Urban Institute
Thanks!

jhenly@uchicago.edu