Challenges and Opportunities in Working with Large Scale Datasets

Haksoon Ahn, Sarah Kaye, & Terry Shaw
Ruth H. Young Center for Families and Children
University of Maryland School of Social Work

Society for Social Work and Research Conference
San Francisco, January 14-17, 2010
Workshop Outline

• Provide a brief overview of secondary data
  - Advantages/disadvantages to using secondary data
  - How to access existing data
  - Things to consider when selecting a dataset

• Illustrate analytic strategies using 3 datasets
  - AFCARS: multinomial logit, survival analysis
  - SIPP: Longitudinal Data Analysis
  - NSCAW: Propensity Score Matching (PSM)

• Discuss State/University collaboration
  - How to build and maintain relationships
  - Examples of work in Maryland and California
Using Secondary Data Sources for Research: Challenges and Opportunities
Advantages

• Cost and access
• Large sample sizes
• Population representation
• Timeliness for measuring current issues

Disadvantages

- Fit between research question and dataset
- Imprecise measurement
- Data-driven research
- Data mining
- Cost of learning a new dataset

Hofferth (2005)
Data Warehouses

• National Data Archive on Child Abuse and Neglect

• Inter-University Consortium for Political and Social Research

• Social Science Electronic Library

• Census Bureau
Questions to Consider

• Does the design of the study fit the research question?

• Do I understand methods and tools appropriate for the research design?

• Am I familiar with the question sequences and missing data patterns?

Hofferth (2005)
Questions to Consider

- Are sample sizes of subgroups large enough?
- Does the data set have the measures I need?
- What previous analyses have been done with these data?

Hofferth (2005)
Adoption and Foster Care Analysis and Reporting System
Research questions

What is the effect of ASFA on rates of exit from out-of-home care?

1. What is the effect of ASFA on the likelihood of exit from foster care?

2. What is the effect of ASFA on the timing of exits from foster care?
Data Needs

• Service experience of children in foster care nationwide
  – Length of stay in out-of-home care
  – Discharge/permanency details
  – Demographic and case characteristics

• Consistent data collection before and after implementation of ASFA (1999)
Data Solution: AFCARS

• Adoption and Foster Care Analysis and Reporting System (AFCARS)
  – Child-level administrative data
  – National dataset compiled from states semi-annually
<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population-level dataset</td>
<td>Data quality varies across states</td>
</tr>
<tr>
<td>Consistent data collection/reporting over time</td>
<td>Few required elements about family characteristics</td>
</tr>
<tr>
<td>Collects important demographic and case flow information</td>
<td>Uses dichotomous variables to represent complex constructs</td>
</tr>
</tbody>
</table>
Study Variables

- **IV: ASFA**
  - date of entry into OOH
  - date of implementation in each state

- **DV1: Exit from OOH**
  - Adoption
  - Reunification
  - Emancipation

- **DV2: Length of Stay**
  - Exits: Date of exit – Date of entry
  - Current: Report date – Date of entry
Control variables

• Reason for placement
  – Neglect, abuse, death of parent

• Parental characteristics
  – Incarceration, substance abuse, illness

• Child Characteristics
  – Age, gender, disability

• Family structure
• Race/ethnicity
Empirical Strategy #1: Multinominal Logit

What is the effect of ASFA on the likelihood of exit from foster care?

- **DV:** more than 3 categories
  - Exhaustive and mutually exclusive
  - Sufficient distribution of cases across categories
- Simultaneously predicts likelihood of each category
- Results are predicted probabilities of falling into each category, relative to reference group
Empirical Strategy #2: Survival Analysis

What is the effect of ASFA on length of stay in foster care?

- Some cases will have experienced an event while others have not
- DV: Time to experiencing an event
- Results present the cumulative proportion of the population experiencing the event over time
Results: Life Tables

![Graph showing the proportion of children exiting to adoption pre- and post-ASFA over different lengths of stay.](image-url)
U.S. Census Bureau Panel Datasets

Survey of Income and Program Participation (SIPP)
Survey of Income and Program Participation (SIPP)

• Purpose
To collect source and amount of income, labor force information, program participation (AFDC/TANF, SSI etc) and eligibility, demographic characteristics to evaluate federal, state programs.

• History
Interviewing for the first panel, the 1984 panel, began in October 1983 with a sample size of 26,000 households.

• Current operation
2008 panel began in September 2008 to be interviewed 13 times (13 waves) until December 2012.
2008 panel Wave 2 Core file is available.
• **Respondents**
  All household members 15 years old and over are interviewed by self-response

• **Periodicity**
  Monthly interview
  4 rotation groups, 4 reference months (1 wave)

• **Survey design and sample size**
  National panels with sample size ranging from 14,000 to 36,700 households. Duration of each panel ranges from 2 years to 4 years.
  Core Module – income, government programs, labor force
  Topical Module – child care, disability, health care, taxes
### Table 2. SIPP Panel 2004 Reference Months (horizontal) for Each Interview Month (vertical)

<table>
<thead>
<tr>
<th>Month of Interview</th>
<th>Wave / Rotation</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1st Q</td>
<td>2nd Q</td>
<td>3rd Q</td>
<td>4th Q</td>
<td>1st Q</td>
</tr>
<tr>
<td>Feb 04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun</td>
<td></td>
<td>2/1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>July</td>
<td></td>
<td>2/2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug</td>
<td></td>
<td>2/3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sept</td>
<td></td>
<td>2/4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oct</td>
<td></td>
<td></td>
<td>3/1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nov</td>
<td></td>
<td></td>
<td>3/2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dec</td>
<td></td>
<td></td>
<td>3/3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan 05</td>
<td></td>
<td></td>
<td>3/4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feb</td>
<td></td>
<td></td>
<td></td>
<td>4/1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mar</td>
<td></td>
<td></td>
<td></td>
<td>4/2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apr</td>
<td></td>
<td></td>
<td></td>
<td>4/3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>May</td>
<td></td>
<td></td>
<td></td>
<td>4/4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5/1</td>
<td></td>
</tr>
<tr>
<td>July</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5/2</td>
<td></td>
</tr>
<tr>
<td>Aug</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5/3</td>
<td></td>
</tr>
<tr>
<td>Sept</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5/4</td>
<td></td>
</tr>
<tr>
<td>Oct</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6/1</td>
</tr>
<tr>
<td>Nov</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6/2</td>
</tr>
<tr>
<td>Dec</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6/3</td>
</tr>
<tr>
<td>Jan 06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6/4</td>
</tr>
<tr>
<td>Feb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>July</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sept</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oct</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nov</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dec</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan 07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jul</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sept</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oct</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nov</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dec</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jan 08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Comparison of SIPP with Current Population Survey

<table>
<thead>
<tr>
<th>Feature</th>
<th>SIPP</th>
<th>CPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size and design</td>
<td>1996 Panel: 40,188 households; new panel periodically; each original-sample adult in panel for no. of months in survey; interviews every 4 months</td>
<td>50,000 households; each household in sample for 8 months over 2-year period; rotation group design; monthly interviews (income supplement once per year)</td>
</tr>
<tr>
<td>Sample designed to be representative within states?</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Income data</td>
<td>Data for about 70 cash and in-kind Sources at each 4-month wave, with monthly reporting for most Sources</td>
<td>Data for prior calendar year for about 35 cash and in-kind Sources</td>
</tr>
</tbody>
</table>
# Comparison of SI PP with Current Population Survey

<table>
<thead>
<tr>
<th>Feature</th>
<th>SI PP</th>
<th>CPS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tax data</strong></td>
<td>Information to determine federal, state, and local income taxes; payroll taxes; property taxes</td>
<td>None</td>
</tr>
<tr>
<td><strong>Asset-holdings data</strong></td>
<td>Detailed inventory of real and financial assets and liabilities once each year for panels from 1996 forward and at least once per panel in prior years; more frequent measures for assets relevant for assistance programs</td>
<td>None, except home ownership</td>
</tr>
<tr>
<td><strong>Expenditure data</strong></td>
<td>Information at least once each panel before 1996 and once a year 1996 and beyond on previous month’s out-of-pocket medical care costs, shelter costs (mortgage or rent and utilities), dependent care costs, and child support payments</td>
<td>None</td>
</tr>
<tr>
<td>-----------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>Wave 1</td>
<td>1/92</td>
<td>1/93</td>
</tr>
<tr>
<td>Wave 2</td>
<td>5/92</td>
<td>5/93</td>
</tr>
<tr>
<td>Wave 3</td>
<td>9/92</td>
<td>9/93</td>
</tr>
<tr>
<td>Wave 4</td>
<td>1/93</td>
<td>1/94</td>
</tr>
<tr>
<td>Wave 5</td>
<td>5/93</td>
<td>5/94</td>
</tr>
<tr>
<td>Wave 6</td>
<td>9/93</td>
<td>9/94</td>
</tr>
<tr>
<td>Wave 7</td>
<td>1/94</td>
<td>1/95</td>
</tr>
<tr>
<td>Wave 8</td>
<td>5/94</td>
<td>5/95</td>
</tr>
<tr>
<td>Wave 9</td>
<td>9/94</td>
<td>9/95</td>
</tr>
<tr>
<td>Wave 10</td>
<td>1/95</td>
<td></td>
</tr>
<tr>
<td>Wave 11</td>
<td></td>
<td>7/99</td>
</tr>
<tr>
<td>Wave 12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wave 13</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Uses of SIPP

• Program participation
  One of the most important reasons for conducting SIPP is to gather detailed information on participation in income transfer programs.
  - How have changes in eligibility rules or benefit levels affected recipients?
  - How have changes in the eligibility rules affected the program target population, that is, those eligible to receive benefits?
  - How does income from other household members affect labor force participation and reasons for not working?
  - How do wealth and income patterns differ for various age, gender, and racial groups?
Uses of SI PP

- Longitudinal Survey

- What factors affect change in household and family structure and living arrangements?
- What are the interactions between changes in the structure of households and families and the distribution of income?
- What effects do changes in household composition have on economic status and program eligibility?
- What are the primary determinants of turnover in programs such as Food Stamps?
<table>
<thead>
<tr>
<th>Income Source</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family Earning</td>
<td>Personal wage - salary income - self-employment income</td>
</tr>
<tr>
<td>Asset Income</td>
<td>Person’s interest from savings accounts - money market funds/securities/bonds - stock dividends received - stock dividends reinvested - net rental income - mortgage interest - royalties or other investment income</td>
</tr>
<tr>
<td>Means-tested Cash Transfers</td>
<td>Sum of federal SSI - state SSI - Veteran’s compensation - AFDC/TANF - General Assistance - Indian, Cuban or Refugee Assistance, other welfare</td>
</tr>
<tr>
<td>Food Stamps</td>
<td>Cash amount of food stamps</td>
</tr>
<tr>
<td>Other Income</td>
<td>Social Security - Railroad Retirement - State Unemployment Compensation - Supplemental Employment Benefits - Other Unemployment Compensation - Veteran's Compensation - Worker's Compensation - Employer or union temporary sickness payments - Income from a private sickness, accident or disability insurance policy purchased on your own - Foster child care payments - Child support payments - Alimony payments - Company or union pension - Federal Civil Service or other Federal civilian employee pension - U.S. military retirement pay - State government pension - Local government pension - Income from life insurance policies or annuities - Estates and trusts - Other payments for retirement, disability, or survivors - GI Bill education benefits - Other VA Educational Assistance - Income assistance from a charitable group - Money from relatives or friends - Lump sum payments - Income from roomers or boarders - National Guard or reserve pay - Incidental or casual earnings - Other cash income not included elsewhere - Black Lung Payments, State temporary sickness or disability benefits, National Guard or Reserve forces retirement - Child and dependent care expense credit - EIC - Elderly or disabled credit</td>
</tr>
<tr>
<td>Category</td>
<td>Topical Module</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Health, Disability, &amp; Physical</td>
<td>Adult Well-Being; Children's Well-Being; Functional Limitations and Disability; Health and Disability; Health Status and Utilization of Health</td>
</tr>
<tr>
<td>Well-Being</td>
<td>Care Services; Long-Term Care; Medical Expenses and Work Disability; Work Disability History</td>
</tr>
<tr>
<td>Financial</td>
<td>Annual Income and Retirement Accounts; Assets and Liabilities; Real Estate Property and Vehicles; Recipiency History; Retirement Expectations</td>
</tr>
<tr>
<td></td>
<td>and Pension Plan Coverage; School Enrollment and Financing; Selected Financial Assets; Shelter Costs and Energy Usage; Support for Nonhousehold Members; Taxes</td>
</tr>
<tr>
<td>Child Care &amp; Financial Support</td>
<td>Child Care; Child Support Agreements; Child Support Paid; Support for Nonhousehold Members</td>
</tr>
<tr>
<td>Education &amp; Employment</td>
<td>Education and Training History; Employment History; Job Offers; School Enrollment and Financing; Work-Related Expenses; Work Schedule</td>
</tr>
<tr>
<td>Family &amp; Household Characteristics &amp; Living Conditions</td>
<td>Extended Measures of Well-Being; Family Background; Fertility History; Household Relationships; Marital History</td>
</tr>
<tr>
<td>Personal History</td>
<td>Education and Training History; Employment History; Fertility History; Marital History; Migration History; Recipiency History; Work Disability History</td>
</tr>
<tr>
<td>Welfare Reform</td>
<td>Eligibility for and Recipiency of Public Assistance; Benefits; Job Search and Training Assistance; Job Subsidies; Transportation Assistance; Health Care; Food Assistance; Electronic Transfer of Benefits; Denial of Benefits</td>
</tr>
<tr>
<td>W 1</td>
<td>W 2</td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
</tr>
</tbody>
</table>
| • Recipiency History  
• Employment History |
| • Work Disability  
• Marital History  
• Fertility History  
• Household Relationships  
• Education & Training History  
• Migration History |
| • Child Well-Being  
• Work-related Expenses  
• Child Support Paid  
• Medical Expenses/Health Care Services  
• Assets and Liabilities  
• Real Estate, Dependent Care, Vehicles  
• Mortgage, Stocks, Rental, Other |
| • Annual Income & Retirement Accounts  
• Taxes  
• Child Care  
• Work Schedule |
| • Adult Well-Being  
• Child Support Agreements  
• Functional Limitations/Disabilities-Adult  
• Functional Limitations/Disabilities-Child  
• Support for Non-household members  
• School Enrollment & Financing  
• Employer-Provided Health Benefits |
| • Assets and Liabilities  
• Real Estate, Dependent Care, and Vehicles  
• Mortgage, Stocks, Rental, Other  
• Medical Expenses/Health Care Services  
• Work-related Expenses  
• Child Support Paid |
| • Annual Income & Retirement Accounts  
• Taxes  
• Informal Care Giving  
• Retirement & Pension Plan Coverage |
| • Welfare Reform  
• Child Care  
• Child Well-Being |
Access to SIPP datasets:

How to find publicly available SIPP datasets
Survey of Income and Program Participation

Welcome to the DataFerrett Tutorial for SIPP users!

Extracting Data: step 1
Follow along with the yellow text boxes!

In DataFerrett it is also possible to extract the data into other formats. With the variables left over from your previous exercises go to the "Step 3: Download/Make a Table" tab and click on the "Extract" button.
SIPP Utilities

News Headlines - September 21, 2009

- The March 2009 CPS data is now available to order!

- For the first time in 8 years, we will be adjusting our pricing modestly to reflect our current costs. Effective October 1st, 2009 all prices will increase by 10%. We remain committed to bringing you tremendous value and quality, and look forward to the privilege of serving you in the years to come.

SIPP Utilities

Download Data
Create your requested file
Delete your requested file
Copy/Move your requested file
R23AMT  Amount of foster child care payments (ISS Code 23)

Source:

<table>
<thead>
<tr>
<th>Panels</th>
<th>Waves</th>
<th>File Type</th>
<th>Census Name</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>All</td>
<td>Core</td>
<td>T23AMT</td>
<td>5</td>
</tr>
<tr>
<td>1996</td>
<td>All</td>
<td>Core</td>
<td>T23AMT</td>
<td>5</td>
</tr>
<tr>
<td>1993</td>
<td>All</td>
<td>Core</td>
<td>S23AMT</td>
<td>5</td>
</tr>
<tr>
<td>1992</td>
<td>All</td>
<td>Core</td>
<td>S23AMT</td>
<td>5</td>
</tr>
<tr>
<td>1991</td>
<td>All</td>
<td>Core</td>
<td>S23AMT</td>
<td>5</td>
</tr>
<tr>
<td>1990</td>
<td>All</td>
<td>Core</td>
<td>S23AMT</td>
<td>5</td>
</tr>
<tr>
<td>1989</td>
<td>All</td>
<td>Core</td>
<td>123AMT1 to I23AMT4</td>
<td>5</td>
</tr>
<tr>
<td>1988</td>
<td>All</td>
<td>Core</td>
<td>123AMT1 to I23AMT4</td>
<td>5</td>
</tr>
<tr>
<td>1987</td>
<td>All</td>
<td>Core</td>
<td>123AMT1 to I23AMT4</td>
<td>5</td>
</tr>
<tr>
<td>1986</td>
<td>All</td>
<td>Core</td>
<td>123AMT1 to I23AMT4</td>
<td>5</td>
</tr>
<tr>
<td>1985</td>
<td>All</td>
<td>Core</td>
<td>123AMT1 to I23AMT4</td>
<td>5</td>
</tr>
<tr>
<td>1984</td>
<td>All</td>
<td>Core</td>
<td>123AMT1 to I23AMT4</td>
<td>5</td>
</tr>
</tbody>
</table>

Category: General income

Type: Topcodes

Periodicity: Month

Allocation Flag: R23AMT A

Description:

Core: 1990w1 to 1993w9

Amount of foster child care payments received this month.

Notes:
The associated recipiency variable is R23. See Appendix E for list of income types covered in SIPP. See other variables named with R followed by the ISS code number and "AMT" for the amounts for other income types.
Combining options

'Combining' means including multiple instances of your selected variables in the same output file. (Multiple instances occur when the same variable is measured more than once, e.g. once per reference month, or in both core and topical files)

Combining can be done in two ways. You may place your multiple instances in separate records, which will be stacked (or appended) one after another to create a 'Long' file. Or you may merge them in the same record, creating a shorter file of 'Wide' records.

Combining file types or waves is optional. If you do not check either box, you will get one output file for each selected source file.

<table>
<thead>
<tr>
<th>File types</th>
<th>Combine</th>
<th>Configure combination</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Long (stacked)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Always wide</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Waves</th>
<th>Combine</th>
<th>Configure combination</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Same as Reference months below</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(person-month records)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reference months</th>
<th>Combine</th>
<th>Configure combination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Always combined</td>
<td></td>
<td>(person records)</td>
</tr>
</tbody>
</table>

Number of IDs to output (edit as required): 100

More information
Click 'Extract' to save settings and begin extraction, or return to prior pages to make revisions.

Request file: c:sipp\test.req
specifies 4 variables and 39 sources

Options selected:
- Data format: ASCII with SAS loading code
- Variable P32s included
- Combining and merging options: reference months appended as separate records
- Preview: no
- IDs to be written: 100

Extraction will produce 39 output data files (left list below)

Highlighted output file will contain 3 variables (center list), and up to 4 records per person per reference month per wave, identified by auxiliary variables in right list

Number of IDs to output (edit as required): 100

More information
Technical Issues

• Inconsistencies across panels
  ex) child care arrangement

• High nonresponse

• Undercoverage
<table>
<thead>
<tr>
<th>1993 SI PP panel</th>
<th>1996 SI PP panel</th>
<th>2001 SI PP panel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calendar Year: 1993 -1994</td>
<td>Calendar Year: 1997-1999</td>
<td>Calendar Year: 2002</td>
</tr>
<tr>
<td>2. Child’s brother/sister</td>
<td>2. ...cared for child while at work or at school</td>
<td>2. ...cared for child while at work or at school</td>
</tr>
<tr>
<td>6. Day/group care center</td>
<td>6. Other relative of child</td>
<td>6. Other relative of child</td>
</tr>
<tr>
<td>7. Nursery/preschool</td>
<td>7. Family day care provider caring for 2 or more children outside of ...’s home</td>
<td>7. Family day care provider caring for 2 or more children outside of ...’s home</td>
</tr>
<tr>
<td>8. Organized school-based activity (before/after school)</td>
<td>8. Child care or day care center</td>
<td>8. Child care or day care center</td>
</tr>
</tbody>
</table>

**For child ages 0-5:**
9. Nursery/preschool
10. Federally supported Headstart
11. A non-relative such as a friend/neighbor/sitter/nanny

**For child age 6 and older:**
9. Organized sports, including practices
10. Lessons (music, art, dance, language, computer)
11. Clubs (boys/girls clubs, scouts, and other organizations)
12. Before or after school care program
13. A non-relative such as a friend/neighbor/sitter/nanny
14. Other Arrangements
<table>
<thead>
<tr>
<th></th>
<th>Total Sample</th>
<th></th>
<th>Single Mothers</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Did not report Tax data</td>
<td>Reported Tax data</td>
<td>Did not report Tax data</td>
<td>Reported Tax data</td>
</tr>
<tr>
<td>Age</td>
<td>35</td>
<td>43</td>
<td>32</td>
<td>35</td>
</tr>
<tr>
<td>Schooling level</td>
<td>11th grade</td>
<td>college</td>
<td>high school graduate</td>
<td>college</td>
</tr>
<tr>
<td>Annual number of weeks at work</td>
<td>22 weeks</td>
<td>40 weeks</td>
<td>27 weeks</td>
<td>45 weeks</td>
</tr>
<tr>
<td>Annual family income</td>
<td>$60,348</td>
<td>$66,046</td>
<td>$29,148</td>
<td>$36,537</td>
</tr>
<tr>
<td>Monthly family earnings</td>
<td>$3,648</td>
<td>$3,591</td>
<td>$1,400</td>
<td>$1,871</td>
</tr>
<tr>
<td>Tax expenditures</td>
<td>NA*</td>
<td>$5,763</td>
<td>NA</td>
<td>$2,282</td>
</tr>
<tr>
<td>Net family disposable Income</td>
<td>$47,264</td>
<td>$46,125</td>
<td>$22,849</td>
<td>$25,603</td>
</tr>
<tr>
<td>Observations</td>
<td>190,251</td>
<td>27,464</td>
<td>6,290</td>
<td>1,601</td>
</tr>
</tbody>
</table>
Single Mothers Employment Status

Single mothers

<table>
<thead>
<tr>
<th>Year</th>
<th>Working Mom</th>
<th>Non-working Mom</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993</td>
<td>53.2</td>
<td>46.8</td>
</tr>
<tr>
<td>1994</td>
<td>56.7</td>
<td>43.3</td>
</tr>
<tr>
<td>1997</td>
<td>64.8</td>
<td>35.2</td>
</tr>
<tr>
<td>2002</td>
<td>65.8</td>
<td>34.2</td>
</tr>
</tbody>
</table>

Low-income single mothers

<table>
<thead>
<tr>
<th>Year</th>
<th>Working Mom</th>
<th>Non-working Mom</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993</td>
<td>40.7</td>
<td>59.3</td>
</tr>
<tr>
<td>1994</td>
<td>45.8</td>
<td>54.2</td>
</tr>
<tr>
<td>1997</td>
<td>55.9</td>
<td>44.1</td>
</tr>
<tr>
<td>2002</td>
<td>56.4</td>
<td>43.6</td>
</tr>
</tbody>
</table>
Family Earning, AFDC/TANF, and Family Food Stamp Trends for Low-Income Single Mother Families, 1993-2002

- **Earning**: 10388, 11427, 12711, 13340
- **AFDC/TANF**: 3099, 2714, 1721, 1477
- **Food Stamp**: 2267, 2108, 1682, 614

2007 dollars

- **1996 Welfare Reform**

Childcare Costs Percentage of Total Family Income
1993-2002

- Working mothers
- Non-working mothers

Childcare Costs Percentage of Total Family Income Trends Between Only Mother with Children Families and Other Adults with Children Families, 1993-2002

- Other adults with children
- Only mother with children
National Survey of Child and Adolescent Well-being
Research Question

What is the effect of the care environment in kinship and foster care on adolescent behavior?
Data Needs

• Adolescent behaviors

• Measures of the quality of the home environment

• Sufficient sample sizes of youth in kinship and foster care
Data Solution: NSCAW

- Longitudinal, nationally representative study of children/youth
  - CPS sample (n=5501)
  - LTFC sample (n=727)
- Multi-informant survey methodology
- Developmental-ecological risk-resiliency perspective

NSCAW Child Domains

- Social competence, relationships
- Health, cognitive status
- Adaptive behavior, functional status
- Behavior regulations, emotional and mental health
- Life experiences
- Service experiences

Biemer, Dowd, & Webb (2010)
Parent/Caregiver Domains

• Health status
  – Mental health and physical health
  – Services received

• Caregiver attributes/behaviors (observation)

Biemer, Dowd, & Webb (2010)
Contextual Domains

Caregivers report

- Neighborhood
- Family demographics
- Social support
- Domestic violence
- Criminal involvement

Biemer, Dowd, & Webb (2010)
Caseworker Domains

• Job role
• Work unit
• Caseload
• Work environment
• Demographics

Biemer, Dowd, & Webb (2010)
Operationalization

• Adolescent behavior problems (youth report)
  – Internalizing and externalizing subscale on YSR

• Care environment (youth report)
  – Relationship with caregiver
  – Parental monitoring
  – Activities with caregiver

• Current placement (derived from caseworker report)
### NSCAW Data Collection

<table>
<thead>
<tr>
<th>Wave</th>
<th>Months after Investigation</th>
<th>Informant</th>
<th>Response Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wave 1</td>
<td>2-6</td>
<td>Youth, Caregiver, Caseworker, Teacher</td>
<td>64.2</td>
</tr>
<tr>
<td>(99-01)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wave 2</td>
<td>12</td>
<td>Caregiver, Caseworker</td>
<td>86.7</td>
</tr>
<tr>
<td>(00-02)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wave 3</td>
<td>18</td>
<td>Youth, Caregiver, Caseworker, Teacher</td>
<td>86.6</td>
</tr>
<tr>
<td>(01-02)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wave 4</td>
<td>36</td>
<td>Youth, Caregiver, Caseworker, Teacher</td>
<td>85.3</td>
</tr>
<tr>
<td>(02-04)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wave 5</td>
<td>59-96</td>
<td>Youth, Caregiver, Caseworker, Teacher</td>
<td></td>
</tr>
<tr>
<td>(05-07)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Biemer, Dowd, & Webb (2010)
NSCAW Sampling

• Two stage stratified sample
  – 9 sampling strata ("states")
  – 92 Primary Sampling Units ("counties")
  – 8 sampling domains (cases)

• Oversampled cases
  – Infants
  – Sexually abused children/youth
  – Cases receiving child welfare services

Biemer, Dowd, & Webb (2010)
What is the effect of the care environment in kinship and foster care on adolescent behavior?

First... what if youth who are placed in kinship or foster care are different (for non-random reasons) from youth who remain at home?
Empirical Strategy #1: Propensity Score Matching

1. Predict placement into out-of-home care using a logistic regression using possible
2. Save the resulting probability of placement, or propensity score
3. Match youth on their propensity score
4. Keep only cases with matches and omit cases without close matches
(Other Selection Correction Strategies)

- Control for wave 1 covariates
- Model change in behavior over time
- Heckman’s selection

What is the effect of the care environment in kinship and foster care on adolescent behavior?

Second… does the care environment mediate the relationship between out-of-home placement and adolescent behavior?
Empirical Strategy #2: Series of Regression Models

On the matched sample, regress
a: mediator on IV
b: DV on IV
c: DV on IV and mediator

Administrative Data: Developing a State/University Partnership
Measuring Child Welfare Outcomes

Outcomes, outcomes, everywhere

- Government Performance Results Act of 1993 (GPRA)
- Annual Outcomes Report to Congress mandated by Adoption and Safe Families Act (ASFA) of 1997
- Statewide Data Indicators in Child and Family Services Reviews -- a subset of the Annual Outcomes—from National Child Abuse and Neglect Data System (NCANDS) and Adoption and Foster Care Analysis and Reporting System (AFCARS)
- State specific accountability acts such as the Child Welfare Accountability Act in Maryland which went into effect in 2007.
- More regulations/opportunities on the horizon.
The use of administrative data to improve child welfare programs and policy

**Opportunities**
- Enhanced ability to meet expectations of accountability.
- Ability to examine policy relevant questions.
- Opportunities for replication and linkage with other data systems.

**Challenges**
- Predetermined variables limit the extent of research questions.
- Incomplete or inaccurate recording.
- Potential bias.

Factors determining if an administrative database will be useful for public policy/practice analysis

• The nature and quality of the underlying data in the administrative database. (Is the data reliable and accurate?)

• The extent to which the administrative database is integrated into the decision making process. (Is the data being used for anything?)
Rules of the Game

- The State should be in the drivers seat
  - The state has mandates both internal and external.
  - The state is the expert in their child welfare system.
  - The data belong to the state

- The University has resources and expertise that can facilitate the successful implementation/evaluation of state goals
  - Should be knowledgeable of reporting requirements
  - Should understand state context
  - Should have technical expertise in data
  - Should have STRONG familiarity with state data
Ultimate Child Welfare Goals

- **Family** - *Children deserve a family*
- **Safety*** - *Children need to be safe and protected*
- **Permanency*** - *Children need to be with families.*
- **Well-being*** - *Children need care, guidance, and support.*

* Ultimate Child Welfare Outcomes
Advantages of partnerships

- Enhanced understanding that data are our friends, not our dictators
- Increased participation on state and county workgroups and committees
  - The University can contribute their expertise
  - The State can educate the University on practice.
- Enhanced understanding of administrative data (strengths/weaknesses).
- Enhanced ability to utilize available data to inform policy and practice decisions.
- Sharing resources and expertise.
- The University can assist in answering ad hoc questions
Maryland Context

PLACE MATTERS

Nothing Matters More to a Child Than a Place to Call Home
THE FOCUS

Family
Children deserve a family

Safety
Children need to be safe and protected

Permanency
Children need to be with families.

Well-being
Children need care, guidance, and support.
Place Matters Goals

• Keep children in families first
• Maintain children in their communities
• Reduce reliance on out of home care
• Minimize the length of stay
• Manage with data and redirect resources
Goals of the Collaborative

• Improved accuracy and reliability of Maryland Children’s Electronic Social Services Information Exchange (MD CHESSIE) data
• Increased capacity of DHR/SSA to utilize data in decision making
• Integration of evidence-based practices into DHR/SSA service provision
• Enhanced community-based resources
• Enhanced collaboration across child- and family-serving systems
• A competent child welfare workforce
Analysis of Performance Indicators

Maryland Child Welfare
Performance Indicators
3rd Annual Child Welfare Accountability Report

Terry Y. Shaw, Ph.D., MSW, MPH
Co-Principal Investigator
Natassia Ann, Ph.D., MSW
Co-Principal Investigator
Diane DeFilippis, Ph.D., MSW
Principal Investigator

December 2009

Quality Assurance Processes in
Maryland Child Welfare
3rd Annual Child Welfare Accountability Report

Dianne DeFilippis, Ph.D., MSW
Principal Investigator
Nina Esaki, Ph.D., MSW, MBA
Co-Principal Investigator
Gillian Gregory, MSW
Research Specialist
R. Anna Hayward, Ph.D., MSW
Co-Principal Investigator
Terry Y. Shaw, Ph.D., MSW, MPH
Co-Principal Investigator

December 2009
(1) Using administrative child welfare data to identify sibling groups?

- Practice
  - Frequency of sibling placement
  - Circumstances of sibling placement
- Compliance
  - State legislation
  - Federal CSFR outcomes
- Research
  - Associations between sibling placements and outcome

Administrative Data Methods

• Child Method: Children linked to each other as full, half or step siblings
• Maternal Method: Children linked to a common mother
• Paternal Method: Children linked to a common father
• Removal Address Method: Children removed from a common address
(2) Constructing a Caseload Staffing Formula for Child Welfare In-Home Services Based on the Results of Risk and Safety Assessments

• Caseload size may impact:
  - The quality of child welfare services
  - The outcomes of child welfare services
  - The retention of child welfare staff

(3) Measuring racial disparity in Child Welfare

• The Disproportionality Index compares the proportion of those experiencing an event that are members of one group to that group’s proportion of the underlying population.

• If the group of interest is overrepresented, a comparison group is likely to be underrepresented.

• The Disparity Index compares the likelihood of one group experiencing an event to the likelihood of another group. Disparity Indices can easily be computed in either of 2 ways ----
  • Disproportionality Index for group of interest/Disproportionality Index for comparison group
  • Rate group of interest/rate comparison group

(4) Children Ever in Care

• What does administrative data tell us about lifetime incidence of child welfare system involvement?
• What are the differences between using birth cohorts and synthetic cohorts?

(5) Linking data administrative systems

• The need to meet legal mandates and ethical responsibility.
• The need for real time responsiveness to practice issues.
• The need to answer questions at different levels of analysis.
• Evidence-based policy: the need to contextualize and triangulate.


• In addition we believe, developing the skill set and capacity to link administrative data systems will put the state (and their partners) in a stronger position to successfully access external funds.
Contact Information for the Presenters

Haksoon Ahn, hahn@ssw.umd.edu
Research Assistant Professor
Ruth H. Young Center for Families and Children

Sarah Kaye, skaye@ssw.umd.edu
Research Assistant Professor
Ruth H. Young Center for Families and Children

Terry V. Shaw, tshaw@ssw.umd.edu
Assistant Professor
Ruth H. Young Center for Families and Children