The Good, the Bad, and the Ugly of Playing a Data Custodian

Presented by Chiu-chuang (Lu) Chou

Data and Information Services Center
Center for Demography of Health and Aging

IASSIST 2009, Tampere, Finland
What’s Happening to the Family?

• Increasing Divorce Rate
• Women’s Participation in Labor Market
• Cohabitation and Marriage
• Childbearing and Parenting
• Marital Relationships
• Intergenerational Relationships
• 1983: NICHD issued a Request for Proposal for a large-scale data collection to study the causes and consequences of the changes happening in the U.S. families and households.

• A UW-Madison team of 6 researchers responded with a proposal to design a national survey.

• 1984: UW-Madison’s proposal was accepted.

• 18 months later Larry Bumpass and Jim Sweet submitted their full proposal to conduct the National Survey of Families and Households. A three-year grant of $4.8 million was awarded.
Characteristics of NSFH

• National probability sample
• Pair interviews (husband/wife, cohabiting couples, and parent/child)
• Comprehensive coverage
• Holistic approach
• **Respondent’s current family situation**
  - Household composition
  - Relationship and activities with spouse/partner
  - Dating behavior and marriage orientations of unmarried persons
  - Fertility plans and orientations
  - Parenting behaviors, priorities and expectations for children
  - Parent’s assessment of quality of relationship with children
• Respondent’s current family situation (continued)
  – Parent’s involvement with absent minor children and with step-children
  – Respondent’s contact and quality of relationship with parents, parents-in-law and siblings
  – Inter-household assistance given and received
  – Assessment of the respondent’s “well-being” on several dimensions
  – Division of household labor
  – Current income, assets and debt
• **Retrospective information**
  – Family living arrangement in childhood
  – The experience of leaving the parental home
  – Marital and cohabitation experience
  – Education history
  – Fertility history
  – Employment history
  – Social background
Extent of NSFH Data and Its Documentation Files

- Wave 1 (13,007 cases, 4,355 variables, 1 data file, 27 document files)
- Wave 2 (23,072 cases, 7,503 variables, 15 data files, 54 document files)
- Wave 3 (18,169 cases, 5,111 variables, 12 data files, 27 document files)
ICPSR Related Literature Database: 1,046 publications
- 188 theses
- 641 journal articles
- 188 reports
- 34 book sections
- 11 conference proceedings
- 11 books
Dissemination of Public Use NSFH Data and Documents

- ICPSR (http://www.icpsr.umich.edu)
- Sociomometrics (http://www.socio.com)
- NSFH website (http://www.ssc.wisc.edu/nsfh)
- BADGIR (http://nesstar.ssc.wisc.edu)
• Wave 1 (study 6041)
  – Data, documentation and setup files (SAS and SPSS)

• Waves 1 and 2 (study 6906)
  – Data and documentation files but no setup files

• Wave 3 (study 171)
  – Only study description
  – Users are referred to NSFH website
• American Family Data Archive (AFDA)
  – Wave 1 data, documents, SAS and SPSS statement files and MIDAS online data analysis
  – Wave 2 data, documents, SAS and SPSS statement files and MIDAS online data analysis
  – No Wave 3
NSFH Website

- Data and document files from all three waves are freely available
- Content outlines organized by waves allow easy browsing of the topics
- Waves 1 and 2 data files are in raw ASCII format (layout files only no SAS, SPSS or Stata statement files)
- Wave 3 data files are in SPSS sav format
- FAQ, links to Working Papers and appendices
- Powered by Nesstar Suite
- Waves 1, 2 and 3 were all marked up to the variable level
- Study, document, file and variable are compiled with the DDI standard
- Easy access via Web interface
- Simple tabulation and regression can be run interactively
- Downloading data in various formats
Type of NSFH Users

- Undergraduates
- Graduate Students
- Experienced Researchers
- Mass Media
- General Public

<table>
<thead>
<tr>
<th>Time Spent</th>
<th>Undergraduates</th>
<th>Researchers-nonUW</th>
<th>Graduates</th>
<th>Community</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-9 minutes</td>
<td>1</td>
<td>69</td>
<td>3</td>
<td>8</td>
<td>81</td>
</tr>
<tr>
<td>30 minutes</td>
<td>1</td>
<td>56</td>
<td>2</td>
<td>1</td>
<td>60</td>
</tr>
<tr>
<td>a couple hours</td>
<td>0</td>
<td>51</td>
<td>1</td>
<td>0</td>
<td>52</td>
</tr>
<tr>
<td>substantial</td>
<td>0</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Grand Total</td>
<td>2</td>
<td>183</td>
<td>7</td>
<td>9</td>
<td>201</td>
</tr>
</tbody>
</table>
• Can I use question items for my own survey?

A: Yes. NSFH is a publicly funded project. Its question items can be reused as long as acknowledgement is given to the project.
"It is better to get married than to go through life being single." from NSFH Wave1 SE11, Q59B

<table>
<thead>
<tr>
<th>SEX OF RESPONDENT</th>
<th>BETTER-PERSON BE MARRIED THAN SINGLE</th>
<th>MALE</th>
<th>FEMALE</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>STRONGLY AGREE</td>
<td>21.2</td>
<td>15.8</td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>AGREE</td>
<td>29.5</td>
<td>23.2</td>
<td></td>
<td>25.7</td>
</tr>
<tr>
<td>NEITHER A/D</td>
<td>30.9</td>
<td>35</td>
<td></td>
<td>33.3</td>
</tr>
<tr>
<td>DISAGREE</td>
<td>8.5</td>
<td>12.9</td>
<td></td>
<td>11.1</td>
</tr>
<tr>
<td>STRONG DISAGREE</td>
<td>4.3</td>
<td>8</td>
<td></td>
<td>6.5</td>
</tr>
<tr>
<td>REFUSED</td>
<td>0.2</td>
<td>0.1</td>
<td></td>
<td>0.2</td>
</tr>
<tr>
<td>DON'T KNOW</td>
<td>0.1</td>
<td>0.1</td>
<td></td>
<td>0.1</td>
</tr>
<tr>
<td>NO ANSWER</td>
<td>5.2</td>
<td>4.9</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>N=</td>
<td>5226</td>
<td>7781</td>
<td></td>
<td>13007</td>
</tr>
</tbody>
</table>
“Women are less likely to agree with this statement than men—51 percent of women and 66 percent of men agreed or strongly agreed.” National Surveys of Family Growth, 2002 (cycle 6) Report.
Hello –
I’ve emailed you a few times previously as I’m trying to use the NSFH 3 data for the first time (I’ve used NSFH 1 extensively).

I have looked through most of the codebook files and cannot find any mention of a case weight variable for Wave 3.

Where is that variable located? If there is no such variable, what variable have users employed in weighting their analysis of Wave 3 data?

Could you please check on this important question for me? I am in need of running some basic frequencies very soon and need to apply the weighting variable.
We asked the PI for help. Larry Bumpass said because the respondents in Wave 3 do not represent a national sample, weight variables would not be created. However, he suggested two subgroups, respondents older than 45 and focal child for creating weight variables.
• Waves 1 and 2 only
• Link cases to users’ geo contextual files for residency characteristics information
• Confidentiality Agreement
• User-prepared Contextual Files
  – Use the 11 geographic units in our geo file
  – Geographical characteristics data from other sources like U.S. Census, County and City Data Book
  – Data Dictionary
  – Descriptive Statistics of the Contextual Variables
Future Plans for NSFH

• Create a geocode file for wave 3
• Match NSFH respondents to National Death Index
• Create a three-wave combined longitudinal data file
• “Changing the Shape of the American Family No Longer Undercover, Living Together Is Replacing Marriage” USA Today (April 18, 2000)
• ‘Is Social Stability Subverted If You Answer “I Don't”?; Fears for Children's Well-Being Complicate a Debate Over Marriage’ New York Times (November 4, 2000)
• “Divorce Rate: It's Not as High as You Think” New York Times (April 19, 2005)
• “Love by Numbers” Guardian (London) (July 12, 2008)
Playing the Data Custodian for NSFH

• The Good
• The Bad
• The Ugly
Questions and Comments

nsfhhelp@ssc.wisc.edu