Paradata Session
Framework for this Session

• Four case studies
  – US National Survey of Family Growth
  – China Mental Health Survey
  – US Panel Study of Income Dynamics
  – Kingdom of Saudi Arabia Mental Health Survey
Constraints and Challenges

• Replace traditional evaluation techniques; No audio-recordings on NSFG and KSA
• Explore an unanticipated interviewer production issue (PSID)
• Focus and prioritize resources (CMHS).
Using Paradata for Interviewer Data Quality Monitoring

Nicole Kirgis
March 2014
Outline

• Background on the use of paradata dashboards for production monitoring on the National Survey of Family Growth (NSFG).

• Extension of dashboard use for monitoring interviewer data quality.

• Next steps.
Background

• Use of paradata from sample management system to monitor production and guide field interventions (responsive design).
• Compilation of daily graphs arranged into categories: Effort, Active Sample, Productivity, Data Set Balance.
• Interviewer-level reports as well.
Importance of Paradata for NSFG

• Understanding the work pattern of quarterly sample.
• Monitoring efficiency compared to past quarters.
• Given 12-week schedule, ability to make quick management decisions to change course.
The NSFG Dashboard

**Effort**
- Interviewer’s Working
- Hours
- % Productivity
- Calls/Day
- Calls/Hour
- % Peak Calls
- Screener/Main Calls

**Active Sample**
- % Occupied
- % Eligible
- % Nonworked
- Noncontacts
- Mean Calls
- % 8+ Calls
- % Locked Bldgs
- % Resistant
- % Hard Appt.
- Propensity

**Productivity**
- Interviews
- Cumulative Interviews
- Hours/Interview
- Calls/Interview

**Data Set Balance**
- Response Rate
- % With Kids
- % Sexually Active
- Group Rates
- CV Group Rates
Paradata and Responsive Design


Monitoring Interviewer-Level Data Quality

• Audit trail data from the actual Blaise interview.
• Created three factors based on Principle Component Analysis from the past cycle of data collection.
• Nine individual performance indicators.
Indicators

- Field time
- Error escapes, suppressions, jumps
- Backups
- Don’t know and refused responses
- Help key use
- Remarks used
Three Factors

• Factor 1: Too Fast
• Factor 2: Many Error Checks
• Factor 3: Many ‘Don’t Know’ and ‘Refused’
Example of Factors

<table>
<thead>
<tr>
<th>Row Labels</th>
<th>Column Labels</th>
<th>W08</th>
<th>W10</th>
<th>W12</th>
<th>W06</th>
<th>W08</th>
<th>W10</th>
<th>W12</th>
</tr>
</thead>
<tbody>
<tr>
<td>factor1</td>
<td>-0.36</td>
<td>-0.52</td>
<td>-0.54</td>
<td>-0.79</td>
<td>-0.83</td>
<td>-0.74</td>
<td>-0.82</td>
<td></td>
</tr>
<tr>
<td>factor2</td>
<td>3.48</td>
<td>3.41</td>
<td>3.25</td>
<td>-0.83</td>
<td>-0.36</td>
<td>0.26</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>factor3</td>
<td>0.53</td>
<td>0.74</td>
<td>1.47</td>
<td>0.08</td>
<td>0.50</td>
<td>1.20</td>
<td>1.44</td>
<td></td>
</tr>
</tbody>
</table>
## Example of Indicators

<table>
<thead>
<tr>
<th>Row Labels</th>
<th>Column Labels</th>
<th>W08</th>
<th>W10</th>
<th>W12</th>
<th>W06</th>
<th>W08</th>
<th>W10</th>
<th>W12</th>
</tr>
</thead>
<tbody>
<tr>
<td>avg_backup_perfield_z</td>
<td></td>
<td>-0.26</td>
<td>-0.35</td>
<td>-0.42</td>
<td>-1.26</td>
<td>-1.41</td>
<td>-1.44</td>
<td>-1.46</td>
</tr>
<tr>
<td>avg_DK_perfield_z</td>
<td></td>
<td>1.11</td>
<td>1.37</td>
<td>1.40</td>
<td>0.55</td>
<td>0.77</td>
<td>1.23</td>
<td>1.88</td>
</tr>
<tr>
<td>avg_err_esc_perfield_z</td>
<td></td>
<td>-0.38</td>
<td>-0.36</td>
<td>-0.42</td>
<td>-1.12</td>
<td>-1.17</td>
<td>-0.79</td>
<td>-0.74</td>
</tr>
<tr>
<td>avg_err_jump_perfield_z</td>
<td></td>
<td>0.93</td>
<td>0.62</td>
<td>0.58</td>
<td>-0.11</td>
<td>0.21</td>
<td>0.70</td>
<td>0.32</td>
</tr>
<tr>
<td>avg_err_supp_perfield_z</td>
<td></td>
<td>4.85</td>
<td>4.82</td>
<td>4.86</td>
<td>-0.49</td>
<td>0.18</td>
<td>0.71</td>
<td>0.52</td>
</tr>
<tr>
<td>avg_fieldtime_pervisit_z</td>
<td></td>
<td>-0.54</td>
<td>-0.51</td>
<td>-0.49</td>
<td>0.02</td>
<td>-0.06</td>
<td>-0.06</td>
<td>-0.14</td>
</tr>
<tr>
<td>avg_qhelp_perfield_z</td>
<td></td>
<td>-0.92</td>
<td>-0.75</td>
<td>-0.77</td>
<td>0.11</td>
<td>0.11</td>
<td>-0.09</td>
<td>-0.19</td>
</tr>
<tr>
<td>avg_remclk_perfield_z</td>
<td></td>
<td>2.89</td>
<td>2.91</td>
<td>2.99</td>
<td>2.01</td>
<td>2.58</td>
<td>3.30</td>
<td>3.30</td>
</tr>
<tr>
<td>avg_RF_perfield_z</td>
<td></td>
<td>-0.49</td>
<td>-0.42</td>
<td>0.61</td>
<td>-0.30</td>
<td>0.10</td>
<td>0.71</td>
<td>0.50</td>
</tr>
</tbody>
</table>
Process for Monitoring

• Small production group meets every two weeks to review.
• Discuss interviewers with factors/indicators that look troublesome.
• Decide on interviewer-level intervention.
• Monitor outcomes—look for improvement.
Types of Intervention

• Practice interview: Trip error checks and re-train on techniques for resolving discrepancies.
• Re-train on proper interviewing techniques.
• Increase number of verification interviews.
• Group re-training.
• Investigation at case level.
Next Steps

• Continue to fine-tune variables of interest.
• Develop more user-friendly ways to view the data.
• Better documentation of problems and interventions.
• Implementation on a large panel study.
Thank you!
nkirgis@umich.edu