Interviewer Training, Monitoring, and Performance in the Survey of Income and Program Participation

Jason Fields, U.S. Census Bureau
Matthew Marlay, U.S. Census Bureau
Holly Fee, U.S. Census Bureau
ESRA - Interviewers’ Deviations in Surveys 3
Tuesday 14th July, 16:00 - 17:30 Room: O-202
Outline

- SIPP Background
- Training and Certification
- Monitoring
  - Contact History
  - Audit Trails (keystroke files)
  - Computer Audio Recorded Interviewing
  - Performance Management
## Survey Design

<table>
<thead>
<tr>
<th></th>
<th>NEW SIPP</th>
<th>SIPP CLASSIC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Instrument</strong></td>
<td>Blaise/C#</td>
<td>DOS-based</td>
</tr>
<tr>
<td><strong>Interview Type</strong></td>
<td>Personal visit/telephone</td>
<td>Personal visit/telephone</td>
</tr>
<tr>
<td><strong>Interview Frequency</strong></td>
<td>Annual</td>
<td>3x/year</td>
</tr>
<tr>
<td><strong>Reference Period</strong></td>
<td>Previous year</td>
<td>Previous 4 months</td>
</tr>
<tr>
<td><strong>Workload Release</strong></td>
<td>Single release of full sample for the 4-month interview period</td>
<td>Monthly releases, each containing one-quarter of the sample</td>
</tr>
<tr>
<td><strong>Panel Length</strong></td>
<td>4 years (planned)</td>
<td>2.5-5 years</td>
</tr>
<tr>
<td><strong>Sample Size</strong></td>
<td>52,000 households (W1)</td>
<td>11-45,000 households (W1)</td>
</tr>
<tr>
<td><strong>Universe</strong></td>
<td>Civilian, non-institutional</td>
<td>Civilian, non-institutional</td>
</tr>
<tr>
<td><strong>Content</strong></td>
<td>Comprehensive</td>
<td>Comprehensive</td>
</tr>
<tr>
<td><strong>File Structure</strong></td>
<td>Person-month data for full calendar year</td>
<td>Person-month data for staggered four-month reference periods</td>
</tr>
</tbody>
</table>
2014 SIPP: Content Overview

- Coverage and Unit Questions
  - Roster
    - Name - Sex - Birthdate/Age
  - Demographics
    - Hispanic origin
    - Race
    - Citizenship
    - Language
    - Marital status
    - Parent/child relationships
    - Educational attainment
    - Armed forces status
    - Prior year corresident people
    - Program/income screeners

- Event History Calendar
  - Residency
  - Marital history
  - Educational enrollment
  - Jobs/Time not working
  - Program receipt
  - Health insurance

- Post-EHC Questions
  - Health insurance wrap-up
  - Dependent care
  - Non-job income
  - Program income
  - Asset ownership
  - Household expenses
  - Health care utilization
  - Medical expenditures
  - Disability
  - Fertility history
  - Biological parents’ nativity and mortality
  - Child care
  - Child well-being
  - Adult well-being

- Closing Screens
  - Respondent Identification Policy
  - Contact information
  - Moving intentions
# Interviewers

<table>
<thead>
<tr>
<th></th>
<th>Wave 1 Staffing</th>
<th>Wave 2 Staffing</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interviewing period</strong></td>
<td>February 1 – June 9, 2014</td>
<td>February 1 – May 31, 2015</td>
</tr>
<tr>
<td><strong>Hiring period</strong></td>
<td>Fall/winter 2013 (<em>significantly delayed by federal furlough in October 2013</em>)</td>
<td>Fall/winter 2014</td>
</tr>
<tr>
<td><strong>Training period</strong></td>
<td>December 2013 – April 2014</td>
<td>December 2014 – March 2015</td>
</tr>
<tr>
<td><strong>Field representatives (FRs)</strong></td>
<td>1,198</td>
<td>1,140</td>
</tr>
<tr>
<td><strong>New hire field representatives</strong></td>
<td>423</td>
<td>310</td>
</tr>
<tr>
<td><strong>Sample Size</strong></td>
<td>Approx. 53,000 households</td>
<td>Approx. 30,000 households</td>
</tr>
<tr>
<td><strong>Average workload</strong></td>
<td>About 40-45 cases per interviewer</td>
<td>About 25-30 cases per interviewer</td>
</tr>
<tr>
<td><strong>Interviewing mode</strong></td>
<td>Interviews all started in-person with some telephone completion</td>
<td>Interviews mostly in-person but with some telephone on request</td>
</tr>
<tr>
<td><strong>Interviewed households</strong></td>
<td>Approx. 30,000 households</td>
<td>Approx. 23,000 households</td>
</tr>
<tr>
<td><strong>Response rate</strong></td>
<td>70.2%</td>
<td>74.2%</td>
</tr>
</tbody>
</table>
# SIPP 2014 Interviewer Training

- Decentralized training after centralized ‘Train-the-Trainer’ at Census HQ

**Two-day generic Census training**
- New hires only
- Communicating with respondents
- Covers cross-survey skills
- Administrative training

**Four-day classroom training**
- All SIPP Interviewers (FRs)
- Decentralized verbatim training
- Paired-practices
- Content specific to SIPP
- Daily quizzes
- Computer based training sequences

**Pre- and post-classroom self-study modules**

**Ends with certification test**
- Required before fieldwork can be started
Certification Test Content

72 questions, divided into 8 sections:
1. Field Procedures (11)
2. Event History Calendar (12) *
3. Programs (6)
4. Movers (15) **
5. Content (10)
6. Noninterviews (6)
7. Medicare vs. Medicaid (7)

*(Wave 2 expanded content)
**(Wave 2+ content only)
Figure 1. Number of Test Takers by Certification Test Score, 2014 SIPP Wave 2 (n=1,362)
Figure 2. Average Certification Test Score for each Subsection, 2014 SIPP Wave 2 (n=1,362)

<table>
<thead>
<tr>
<th>Test Subsection</th>
<th>Test Score (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>82.07</td>
</tr>
<tr>
<td>Field Procedures</td>
<td>85.46</td>
</tr>
<tr>
<td>EHC</td>
<td>73.86</td>
</tr>
<tr>
<td>Programs</td>
<td>85.56</td>
</tr>
<tr>
<td>Movers</td>
<td>77.71</td>
</tr>
<tr>
<td>Survey Content</td>
<td>78.40</td>
</tr>
<tr>
<td>Noninterviews</td>
<td>84.10</td>
</tr>
<tr>
<td>Medicare/Medicaid</td>
<td>95.85</td>
</tr>
<tr>
<td>Blaise</td>
<td>88.87</td>
</tr>
</tbody>
</table>
Figure 3. Average Certification Test Score for each Subsection by Month Administered, 2014 SIPP Wave 2
Figure 4. Mean Certification Test Score in Each Wave by Regional Office for Individuals Who Completed Both Waves (n=853)
Figure 5. Average Interview Duration across Caseload (First 40 Cases) by Certification Test Score, 2014 SIPP Wave 1

- < 70% (n=853)
- 70 - 80% (n=4,079)
- 80 - 90% (n=11,891)
- ≥ 90% (n=9,255)

Legend:
- Linear (< 70% (n=853))
- Linear (70 - 80% (n=4,079))
- Linear (80 - 90% (n=11,891))
- Linear (≥ 90% (n=9,255))
Contact History Instrument

- Keeps a history of every contact attempt for every case
- Collects information about the kind of response received (if contact is made)
  - Reluctant respondent, etc.
- Also collects FR’s observation about housing unit/neighborhood conditions
## CONCERN / BEHAVIOR / RELUCTANCE

- Select the categories that describe respondent concerns, behaviors, or reluctance during this contact attempt.
- Enter all that apply, separate with commas.

<table>
<thead>
<tr>
<th>Concern/Behavior/Reluctance</th>
<th>Other Concerns</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Not interested / Does not want to be bothered</td>
<td>12. Hostile or threatens FR</td>
</tr>
<tr>
<td>2. Too busy</td>
<td>13. Other household members tell respondent not to participate</td>
</tr>
<tr>
<td>3. Interview takes too much time</td>
<td>14. Talk only to specific household member</td>
</tr>
<tr>
<td>4. Breaks appointments (puts off FR indefinitely)</td>
<td>15. Family issues</td>
</tr>
<tr>
<td>5. Scheduling difficulties</td>
<td>16. Respondent requests same FR as last time</td>
</tr>
<tr>
<td>6. Survey is voluntary</td>
<td>17. Gave that information last time</td>
</tr>
<tr>
<td>7. Privacy concerns</td>
<td>18. Asked too many personal questions last time</td>
</tr>
<tr>
<td>8. Anti-government concerns</td>
<td>19. Too many interviews</td>
</tr>
<tr>
<td>9. Does not understand survey / Asks questions about the survey</td>
<td>20. Last interview took too long</td>
</tr>
<tr>
<td>10. Survey content does not apply (retired, healthy, no crimes to report)</td>
<td>21. Intends to quit survey</td>
</tr>
<tr>
<td>11. Hang-up / slams door on FR</td>
<td>22. No concerns</td>
</tr>
<tr>
<td></td>
<td>23. Other - specify</td>
</tr>
</tbody>
</table>
• CONTACT STRATEGIES ATTEMPTED

• Select the categories that describe the strategies used on this contact attempt.
• Enter all that apply, separate with commas.

☐ 1. Advance letter given
☐ 2. Scheduled appointment
☐ 3. Left note / appointment card
☐ 4. Left promotional packet / informational brochure
☐ 5. Called household
☐ 6. Left message on answering machine
☐ 7. FR will request No One Home Letter
☐ 8. FR will request Refusal Letter
☐ 9. FR will request Better Understanding Letter
☐ 10. Called contact persons
☐ 11. Stake-out
☐ 12. Checked with neighbors
☐ 13. Contacted other family members
☐ 14. Contacted property manager
☐ 15. Visited county assessor / post office / permit office
☐ 16. On-line tracking database
☐ 17. Sought help from SFR / RO
☐ 18. Reassignment
☐ 19. Offered incentive
☐ 20. CED double placement
☐ 21. Used MAF or ALMI
☐ 22. No Strategies
☐ 23. Other - specify
Audit Trails

- Audit trail files are a record of all of the keystrokes entered by a field representative (FR) during an interview.

- Audit trail files can be used to create paradata on such things as:
  - Section timers,
  - Don’t know/refused counts,
  - Help screen calls,
  - Checks encountered,
  - Item-level notes left, and
  - FR navigation throughout the instrument.
Blaise audit trail example – sanitized data

"2/dd/20yy 10:02:22 PM", "{KEY:} 15[ENTR]"
"2/dd/20yy 10:03:02 PM", "Action: Store Field Data", "Field: BDemographics.BAge[1].DOB_BDAY"
"2/dd/20yy 10:03:02 PM", "Enter Field: BDemographics.BAge[1].DOB_BYEAR", "Status: Normal", "Value:"
"2/dd/20yy 10:03:04 PM", "{KEY:} 15[ENTR]"
"2/dd/20yy 10:03:05 PM", "(KEY:)
"2/dd/20yy 10:03:10 PM", "{KEY:}
"2/dd/20yy 10:03:15 PM", "{KEY:}
"2/dd/20yy 10:03:21 PM", "Mouse: 906,589", "Message: LeftDown", "HitTest: Client"
"2/dd/20yy 10:03:21 PM", "Mouse: 906,589", "Message: LeftUp", "HitTest: Client"
"2/dd/20yy 10:03:21 PM", "Mouse: 906,589", "Message: LeftDown", "HitTest: Client"
"2/dd/20yy 10:03:21 PM", "Mouse: 906,589", "Message: LeftUp", "HitTest: Client"
"2/dd/20yy 10:03:21 PM", "ErrorDlg action: Goto", "Text: @F That would make you 92 years old. Is that correct? @F @/ @/ @Zs @Z @L If this is correct, suppress and continue. @L @/ @Zs @Z @L If this is not correct, go back to DOB_BMONTH, DOB_BDAY, or DOB_BYEAR and correct. @L
"Involved: BDemographics.BAge[1].DOB_BDAY; 15; BDemographics.BAge[1].DOB_BMONTH; April; BDemographics.BAge[1].DOB_BYEAR; 1918", "Field: BDemographics.BAge[1].DOB_BDAY"
"2/dd/20yy 10:03:24 PM", "{KEY:}
"2/dd/20yy 10:03:25 PM", "{KEY:} 1951[ENTR]"
"2/dd/20yy 10:03:31 PM", "{KEY:} s"
"2/dd/20yy 10:03:31 PM", "ErrorDlg action: Suppress", "Text: @F That would make you 59 years old. Is that correct? @F @/ @/ @Zs @Z @L If this is correct, suppress and continue. @L @/ @Zs @Z @L If this is not correct, go back to DOB_BMONTH, DOB_BDAY, or DOB_BYEAR and correct. @L
"Involved: BDemographics.BAge[1].DOB_BDAY; 18; BDemographics.BAge[1].DOB_BMONTH; Sept; BDemographics.BAge[1].DOB_BYEAR; 1951", "Field: BDemographics.BAge[1].DOB_BDAY"
# Audit Trails

## Statistics (Completed cases)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Total</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Don’t Know (CTRL+D)</td>
<td>13.33</td>
<td>15.61</td>
</tr>
<tr>
<td>Refuse (CTRL+R)</td>
<td>4.46</td>
<td>15.60</td>
</tr>
<tr>
<td>Help Call Screens (F1)</td>
<td>0.37</td>
<td>0.92</td>
</tr>
<tr>
<td>Field Case Notes (F7)</td>
<td>0.76</td>
<td>2.90</td>
</tr>
<tr>
<td>Survey Time (in minutes)</td>
<td>102.41</td>
<td>51.89</td>
</tr>
</tbody>
</table>

18 of 23 cases completed
Computer-Assisted Recorded Interviewing (CARI)

- FRs must obtain consent from each respondent to record the interview
- Records interactions between Field Representatives (FRs) and respondents
- The goal of CARI is to ensure the accuracy and quality of data collected
  - Improve the FR’s performance
  - Identify difficult or problematic questions
Figure 5. Mean CARI Consent Rate (Persons) by Interviewer Certification Test Score

- <70%: 49.74%
- 70 - 80%: 49.78%
- 80 - 90%: 56.06%
- ≥ 90%: 58.64%
Figure 6. Mean CARI Consent Rate (Persons) by Regional Office and Certification Test Score

What Accounts for the Variance in Average Household CARI Consent Rate?

<table>
<thead>
<tr>
<th>Source: 2014 SIPP, Wave 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Regional Office</strong></td>
</tr>
<tr>
<td>New York (n=162)</td>
</tr>
<tr>
<td>Philadelphia (n=251)</td>
</tr>
<tr>
<td>Chicago (n=251)</td>
</tr>
<tr>
<td>Atlanta (n=345)</td>
</tr>
<tr>
<td>Denver (n=239)</td>
</tr>
<tr>
<td>Los Angeles (n=156)</td>
</tr>
</tbody>
</table>
CARI

- Helps ensure a focus on data quality and encourages professionalism
- Listen to recorded cases and code them for:
  - Authenticity (including consent to record)
  - Question administration
  - Behavioral conduct
- Coded Quality Assurance score will directly influence performance rating
- Completely in the control of the interviewer
- May increase non-response and will increase interviewing length
Conclusion

- SIPP (and the Census Bureau more generally) has access to more paradata than we have ever had in the past
- Effective use of this paradata for FR monitoring and performance can help us improve data quality
- In period of declining response rates, focusing on methods to:
  - ensure quality,
  - improve training
  - Improve survey instruments, and
  - leverage administrative data