The ATUS and SIPP-EHC: Recent Developments

ROBERT F. BELLi
UNIVERSITY OF NEBRASKA
Acknowledgment

- This material is based upon work supported by the National Science Foundation under Grant No. SES - 1132015. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the presenter and do not necessarily reflect the views of the National Science Foundation.
Background: ATUS

- American Time Use Survey
  - Nationally representative annual household survey
    - Oversampling
      - Hispanic and African American households
      - Households with children
    - CATI
    - Measures with a time diary approach how respondents spend their time
    - Excellent source for measuring unpaid labor and demographic differences in time use
Background: SIPP

- Survey of Income and Program Participation
  - Nationally representative annual panel household survey
    - Oversampling
      - Low income households
    - Primarily CAPI
    - Measures income, employment, and program participation (social security, food stamps, TANF, medicare)
  - One of the premier sources of data for planning, evaluating, and improving government programs
Background: ATUS and EHC-SIPP

Interviewing Method

• Both ATUS and Reengineered Event History Calendar SIPP (EHC-SIPP)
  • Flexible temporally-based interviewing
    • Designed to optimize the quality of retrospective reports
    • Retrieval processes, as cues, are those known to exist in the structure of autobiographical knowledge
Background: ATUS Instrument

So let's begin. Yesterday, Tuesday, at 4:00 AM, what were you doing?

- Use the slash key (/) for recording separate/simultaneous activities.
- Do not use precodes for secondary activities.

1. Sleeping
2. Grooming (shave)
3. Watching TV
4. Working at main job
5. Working at other job
6. Preparing meals or snacks
7. Eating and drinking
8. Cleaning kitchen
9. Laundry
10. Grocery shopping
11. Attending religious service
12. Paying household bills
13. Caring for animals and pets
14. Don't know/Can't remember
15. Refusal/None of your business

Example activity log:

<table>
<thead>
<tr>
<th>Start</th>
<th>Activity</th>
<th>TIME</th>
<th>Who</th>
<th>Where</th>
</tr>
</thead>
<tbody>
<tr>
<td>6:00 AM</td>
<td>Sleeping</td>
<td>2 h, 0 min</td>
<td></td>
<td>Respondent's home</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>Preparing meals and snacks</td>
<td>1 h, 1 min</td>
<td></td>
<td>Respondent's home</td>
</tr>
<tr>
<td>1:00 PM</td>
<td>Playing with kids</td>
<td>1 h, 0 min</td>
<td></td>
<td>Respondent's home</td>
</tr>
<tr>
<td>3:00 PM</td>
<td>Grooming</td>
<td>1 h, 2 min</td>
<td></td>
<td>Respondent's home</td>
</tr>
<tr>
<td>3:30 PM</td>
<td>Eating and drinking</td>
<td>30 min</td>
<td></td>
<td>Respondent's home</td>
</tr>
<tr>
<td>5:00 PM</td>
<td>Watching TV</td>
<td>1 h, 0 min</td>
<td></td>
<td>Respondent's home</td>
</tr>
<tr>
<td>7:30 PM</td>
<td>Sleeping</td>
<td>2 h, 0 min</td>
<td></td>
<td>Respondent's home</td>
</tr>
</tbody>
</table>
Background: Parallel and Sequential Navigation

EHC-SIPP Calendar
Fake data, not from the survey

<table>
<thead>
<tr>
<th>Topic</th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landmarks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marital History</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>More Jobs (if any)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Job</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Food Stamps</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TANF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gen. Assist.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WIC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicare</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical Assist...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Military</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Coverage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Coverage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sequential retrieval
if chronological retrieval of events within the same theme

Parallel retrieval
if retrieval of contemporaneous events across themes

Lee et al. (in preparation)
## Direct Comparisons Between Calendar And Conventional Questionnaires: Calendar Alone Vs. Conventional Alone

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Mode/ Medium</th>
<th>Ref Period</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engel et al., 2001</td>
<td>Test-retest</td>
<td>F-to-f/ P&amp;P</td>
<td>Life</td>
<td>Cal: reports of more jobs and work time</td>
</tr>
<tr>
<td>Belli et al 2001</td>
<td>Experiment</td>
<td>Tel/ P&amp;P</td>
<td>2 years</td>
<td>Cal: more accurate reports of residential and labor histories; Conv: more accurate reports of entitlements</td>
</tr>
<tr>
<td>Yoshihama et al 2005</td>
<td>Quasi-Experiment</td>
<td>F-to-f/ P&amp;P</td>
<td>Since age 12</td>
<td>Cal: more reports of intimate partner violence; more reasonable pattern of results</td>
</tr>
<tr>
<td>Belli et al 2007; Belli et al 2012</td>
<td>Experiment</td>
<td>CATI</td>
<td>Life</td>
<td>Cal: more accurate reports of cohabitation, labor, disability, and health histories; Conv: more accurate reports of marital history</td>
</tr>
</tbody>
</table>
## Direct Comparisons Between Calendar And Conventional Questionnaires:
### Conventional + Calendar Vs. Conventional Alone

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Mode/Medium</th>
<th>Ref Period</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>van der Vaart 2004</td>
<td>Experiment</td>
<td>F-to-f/P&amp;P</td>
<td>8 years</td>
<td>Cal+con: more accurate reports of number, starting date, and types of educational courses</td>
</tr>
<tr>
<td>Van der Vaart &amp; Glasner 2007</td>
<td>Experiment</td>
<td>Con CATI Cal P&amp;P/SAQ</td>
<td>7 years</td>
<td>Cal+con: more accurate reports of when pairs of glasses were purchased and their price</td>
</tr>
<tr>
<td>Glasner 2011</td>
<td>Experiment</td>
<td>Web/SAQ</td>
<td>Life</td>
<td>Cal+con: more reports of unemployment and family leaves, especially for more remote events</td>
</tr>
</tbody>
</table>
Background: Verbal Behavior Coding Studies

- Verbal behavior code audio streams from CATI calendar and standardized interviews
- Infer the memory, cognitive, and communicative processes of both interviewers and respondents from their verbal behaviors
- Determine differences in processes between calendar and standardized interviews
- Associate these processes with data quality in calendar interviews

Background: Verbal Behaviors

• Calendar Interviews More Prevalent than Conventional
  – Interviewer parallel and sequential retrieval probes
  – Respondent spontaneous parallel and sequential retrieval strategies

• Validity analyses
  – Parallel and sequential retrieval probes and strategies associated with higher quality retrospective reports
    • Especially when there is greater retrieval complexity

Belli, Bilgen, & Al Baghal (2013), Public Opinion Quarterly
Paradata: Audit Trails

• Do not have audio recordings from ATUS or SIPP-EHC

• Do have sanitized audit trails from ATUS (2010) and SIPP-EHC field test (2011, 2012)
  – CATI or CAPI automated recordings of interviewer interaction with the instrument
    • entered values (keystrokes, mouse clicks) and time stamps

• Audit trails do permit insight into interviewing dynamics
Audit Trails and Data Quality: ATUS

Controls
Respondent: gender, age, education, race, employment status, family income, marital status, presence of children, home owner, urbanicity; number of activities (exposure)
Interviewer: interviewer experience (within study); interviewer cooperation rate

Estimation
Multilevel negative binomial structural equation models (respondents nested in interviewers); Stata 14 gsem
EHC-SIPP Audit Trail Challenges: Parallel Navigation

• Some navigations are automatic by the instrument
  – Automatic jump back from the last topic completed (e.g., “No Coverage”) to a topic previously not completed, if any (e.g., “Landmarks”)
  – Audit trails do not indicate whether intended or automatic

• Acquiring unsanitized audit trails would appear not to help

• Acquiring response data will not help

Seloske, Lee, et al. (in preparation)
EHC-SIPP Audit Trail Challenges: Timeline Gaps

- Identification requires knowing the beginning and end months of each spell
  - 2012 audit trails contain many instances where beginning and end months were recorded as zero as a result of edited spells
- 2013 audit trails do not suffer from same problem
- Acquiring unsanitized audit trails may help by identifying value changes
- Acquiring response data will not help

Seloske, Lee, et al. (in preparation)
EHC-SIPP Measures of Data Quality: Limitations

- Timeline gaps and answer changes
  - 2012 data will always be deficient

- Answer changes
  - May signal
    - lack of satisficing (proxy for good data quality)
    - Cognitive or interviewing difficulty (proxy for poor data quality)

- Number of final spells
  - Need response data

Seloske, Lee, et al. (in preparation)
EHC-SIPP Need for administrative records: Continuous variables

• Correspondence of responses with records on:
  – Amount of annual earnings at the job-level and at the person-level
  – Amount of program recipient benefits
  – Duration of (un)employment status, program receipt, Medicare, and Medicaid
• Correspondence of responses with records on:
  – Number of spells on jobs, programs, Medicare, and Medicaid
  – Number of household members covered by Medicare and Medicaid
  – Number of transitions in (un)employment status, program receipt, Medicare, and Medicaid
EHC-SIPP: Need for administrative records: Binary/Categorical Variables

• Correspondence of responses with records on:
  – Errors of omission and commission in employment, program receipt, Medicare, and Medicaid
  – Forward and backward telescoping in employment, program receipt, Medicare, and Medicaid

Lee, et al. (in preparation)
Why?

Associate aspects of interviewing dynamics with data quality.
Web ATUS

• Enriched Time Diary Instrument
  – Developed by CES team at UNL
    • Features such as auto-complete, pre-codes, timeline, multi-mode data entry
    • Intelligent features such as highlighting likely events at real-time based on statistical analysis
    • Tracking mechanisms to record paradata with high resolution

Soh, Eck, et al. (in preparation)
What activity were you doing? (Read if necessary: An activity is anything you did during the day. Right now, you are talking to me on the telephone. Talking on the telephone is one type of activity).

Click on the options listed below or type in the activity name in the Activity field.

- Building/maintenance/repair
- Caring for children
- Cooking/cleaning
- Doing hobby
- Doing laundry
- Eating and drinking (on, at home)
- Eating/drinking (here)
- Educational activities
- Emergencies
- General household activities
- Going to parties/meetings
- Interacting with children
- Interacting with children
- Laundry
- Medical care
- Personal care
- Personal time and leisure/relaxing
- Providing medical care
- Reading
- Religious activities
- Don't know/Can't remember
- Sleeping
- Sports and exercises
- Traveling
- Waiting
- Watching TV/movies
- Working
- Refused

Soh, Eck, et al. (in preparation)
Next Steps toward Web ATUS

• Behavior Coding
  – Coding the type of questions asked by interviewers
    • Activity, Who, Where, Time, and Duration
  – Coding the type of responses
    • Match, Match more than one, Relevant Mismatch, No Match, Refuse to answer/No answer

Soh, Eck, et al. (in preparation)
Next Steps (2)

• Data mining of coded behaviors and actual transcript texts
  – To identify sequences of behaviors that are associated with good “quality” and “poor” interviews
    • Quality depending on interviewer’s self report and also quantitative measures of response data matches and number of activities recorded
  – To identify anomalies in data entry
    • e.g., repeated back-and-forth of turns between respondents and interviewers without data entry; repeated data entry without interviewer questions;
  – To correlate texts with such sequences and anomalies

Soh, Eck, et al. (in preparation)
Next Steps (3)

• Developing a casebase as a “brain” repository for an agent-powered self-administered ATUS
  – The agent will perform case-base reasoning at real-time to engage respondents at a timely manner with the texts that are likely to lead to a good response
  – Each case consists of a “problem” and a “solution” with a “utility”
    • Case must capture poor quality sequences (e.g., item nonresponse)
    • Problem: Sequences of coded behaviors (including the turns of questions and responses) together with the correlated texts used prior to item nonresponse
    • Solution: Sequences of coded behaviors that lead to data entry
    • Utility: whether the solution leads to a good quality or a poor quality interview

Soh, Eck, et al. (in preparation)
Intelligent Agent Instruments

• Agent and respondent inputs
  – Mouse clicks, keystrokes, voice?

• Both time diaries and calendars
  – Time diary as a “simplified” version of a calendar data collection instrument

• Both based on developing cases from actual interviews
  – Associating interviewing sequences with data quality
Many Thanks To….

• The NCRN research team at UNL
  – Hariharan Arunachalam
  – Ana Lucia Cordova Cazar
  – Caitlin Deal
  – Adam Eck
  – Antje Kirchner
  – Jinyoung Lee
  – Benjamin Seloske
  – Leen-Kiat Soh