What's In a Keystroke? **Working Toward an Automated Process Evaluation Using Paradata**

Catherine Barry, M.A., University of California Berkeley and Rachael Walsh, Ph.D., U.S. Census Bureau

Research Goal

Explore data about the survey collection process, known as paradata, to better understand how interviewers and respondents interact with the survey instrument.

Paradata: Audit Trail Files

Paradata

- Process information such as keystrokes and audit trail files that capture data on survey instrument utilization and time progression
- Useful for detecting problems and errors in the survey instrument

Audit Trail Files

- Blaise software provides an audit trail file, which is a file that records every keystroke made by the interviewer while on an interview.
- Keystroke records provide function key use data used in analysis.
- Each keystroke has a time and date stamp, which is used to produce interview length.

Data Source

Survey of Income and Program Participation Event History Calendar (SIPP-EHC) 2011

A panel series that collects information about income, labor force information, program participation and eligibility data, and general demographics.

- Re-engineered SIPP
- Evolving instrument and training
- Computer Assisted Personal Interview (CAPI)
- Sample:
- Over-sampled low income stratum
- 4,051 household
- 6,981 individuals



Outcomes—Function Key Use

- Don't Know answers;
- Respondents give an 'I don't know' response to a survey question
- Refuse to answer; Respondents refuse to answer a survey question
- Help Screens
- clarify survey questions
- Interviewer Notes
- Enables interviewers to provide comments Skip respondents
- Enables interviewers to change respondents within a household

Function Key Use in the SIPP-EHC 2011



Outcome Measures

— Provides interviewers with additional information to

compare means.

Survey and Interviewer characteristics

- Order of Interview (Order):

- Self: Responses are provide by an individual about her/himself respondent is not available or is a child



The views expressed on statistical, methodological, technical, or operational issues are those of the author and not necessarily those of the U.S. Census Bureau. All comparative statements in this report have undergone statistical testing, and, unless otherwise noted, all comparisons are statistically significant at the 95% significance level.

Analysis and Evaluation



Results

Function Key Associations

- Help screens are more likely to be used in completed interviews than in partially completed interviews
- Proxy respondents have fewer 'don't know' responses than self respondents
- Greater interviewer experience gained through conducting more interviews is associated with fewer 'don't know' responses
- Longer interviews are associated with greater use of Notes function
- Less interviewer experience is associated with greater use of the Notes function

T-Tests: Average Number of Times Function Keys Used by Help Screen Use



indicates statistically significant diiference at the .65 keed

Recommendations

- Multivariate analysis to address confounding variables that may explain relationships uncovered in the results
- Identify survey items with notes to evaluate interviewer comments/concerns
- Identify survey items with greatest numbers of "don't know" and "refuse" non-responses
- Test data quality of survey items with help screens that are not utilized
- Modify or remove SIPP-EHC Help Screens based on evaluation of above recommended items
- Evaluate SIPP-EHC 2012