Introduction to, and re-engineering of, the Survey of Income and Program Participation

Jason Fields
SIPP Survey Director
U.S. Census Bureau

Population, Family and Reproductive Health Noon Seminar
Johns Hopkins Bloomberg School of Public Health
February 22, 2017
Outline

- Background and focus of SIPP
- A brief look back to 2006
- Uses of SIPP (a couple)
- Getting SIPP data
The two primary goals of SIPP should be to provide improved information on the distribution of income and other economic resources for people and families and on eligibility for and participation in government assistance programs.”

- The Future of the Survey of Income and Program Participation, NAS, 1993

“... [The SIPP] provides an unprecedented opportunity to ascertain the nature of income flows and program participation, both for relatively short periods of time and over extended periods of time, for individuals and families as they experience changes in household composition, income, and labor force participation.”

- Improving National Statistics on Children, Youth and Families, 1984
The SIPP Mission

The mission of the Survey of Income and Program Participation (SIPP) is to provide a nationally representative sample for:

- evaluating annual and sub-annual dynamics of income,
- movements into and out of government transfer programs,
- family and social context of individuals and households, and
- interactions between these items.
The SIPP

- Originally designed to compensate for the limitations of the Current Population Survey (CPS)
  - CPS ASEC (March Supplement) uses a very long recall period
    - Not good at measuring irregular/odd sources of income
    - High levels of under-reporting of program participation
    - Doesn’t capture changes in family structure
    - Note: if this makes you panic about the accuracy of our official poverty/insurance estimates from CPS, no-one will blame you
  
- SIPP was designed to have a (much) shorter recall period

- SIPP is meant to provide better estimates of income and public program participation

- Offers the most detailed income and comprehensive program participation variables of the major nationally representative surveys
Classic SIPP Design: National Panel Survey

- First panel began in 1984
  - 4-month recall period (1984 – 2008 Panel design)
  - 3 interviews per year

- Adults (age 15+) interviewed in Wave 1
  - Data collected for all people
  - Proxy interviews for children under 15

- Follows all Wave 1 interviewed adults in subsequent waves
  - Interview all household members at each address with original Wave 1 adult
Classic SIPP Design: National Panel Survey

- About 2½ years

CAPI (1996 – present)
- 4 years

2008 Panel:
- Extended to 16 waves - about 5 years
- Wrapped up interviews in December 2013
- All files now available for data users

- Panel bridges recession
- Provides data over five-year period, spanning crash and recovery
- Monthly, full-sample data from August 2008-May 2013
SIPP Sample Selection

- Nationally representative of the US non-institutional population
- Households from areas with high poverty concentrations are oversampled
- Primary data collection unit is the household, not the individual
- Two-stage sample design:
  - 1) Selection of primary sampling units (PSUs)
  - 2) Selection of address units within PSUs
SIPP Sample Selection & Data editing

- Frame is a list of US counties and independent cities
- Select addresses from frames from the most recent Decennial Census
- Addresses in frames are clustered, then sampled
- When one person is missing, or items are missing, public use data are imputed
- Imputation is largely sequential hot-deck procedure: missing data are matched with a donor who has similar characteristics
- Administrative records are being used in imputation in the 2014 SIPP Panel
## SIPP Panel Sizes and Collection Periods

<table>
<thead>
<tr>
<th>Panel</th>
<th>Number of Waves</th>
<th>Eligible Households</th>
<th>Date of First Interview</th>
<th>Date of Last Interview</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>4</td>
<td>42,456</td>
<td>Feb. 2014</td>
<td>May 2017</td>
<td>(1)</td>
</tr>
<tr>
<td>1987</td>
<td>7</td>
<td>12,527</td>
<td>Feb. 1987</td>
<td>May 1989</td>
<td></td>
</tr>
</tbody>
</table>

(1) The 2014 Panel is the first EHC panel with annual interviewing.
(2) The 2008 Panel start was delayed due to budget and extended into 2013 to overlap with the 2014 Panel.
(3) This is the first CAPI SIPP panel, and first of the non-overlapping panels.
What’s included in Classic SIPP?

- The core monthly files include:
  
  - **Demographics**: race & ethnicity, age, sex, household/family structure & relationships, state identifiers, education, armed forces status, marital status, student status
  
  - **Income**: Person/family/household earned income, total income, property income, “other income” unit-specific poverty thresholds, receipt of severance pay
  
  - **Employment**: Employed, unemployment, NILF, Data on up to 2 jobs/month, industry, occupation, class of worker (public/private), firm size, union membership, tenure, hourly/salaried, employer-based health insurance
What’s included in Classic SIPP?

- The core monthly files include:
  - **Program participation:** Unemployment Insurance, AFDC/TANF, social security, SSI, SSD, workers’ comp, child support, food stamps, public housing, energy assistance, public (and private) health insurance, free/reduced lunch, veterans payments, pell grants receipt, other federal grant program receipt
  - Most program participation variables include both receipt (0,1) and the amount of benefit

- **THIS IS NOT A COMPREHENSIVE LIST!**

What’s included in Classic SIPP?

- **Topical Modules:** Extra questions added to the core once per year/panel in particular waves. These include point-in-time/annualized variables on things such as:
  - Fertility history
  - Migration history
  - Material hardship measures (such as food security)
  - Assets and liabilities
  - Medical expenses/utilization of health care
  - Work schedule

- **THIS IS NOT A COMPREHENSIVE LIST!**

2014 SIPP Content Areas

**Front Sections**
- Roster
- Demographics
- Relationships
- Armed Forces
- Citizenship / Nativity / Immigration

**EHC**
- Residency
- Marital history
- Educational enrollment
- Jobs/Time not working
- Program receipt
- Health insurance

**Post-EHC Questions**
- Health insurance
- 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
File Structure

- Core file wave data are organized into person-month observations
  - Each person will have a record for each month they appear in the wave—so up to 4 observations, per person, per wave

- Household/coverage/person/and other variables are duplicated in each person’s monthly record

- So each household member record contains the household income variables (lot’s of duplication)
SIPP 2014 File Structure for Public Use

- Person-month file structure - 12 month reference period (January – December)

- Household structure is defined by interview month household composition - Relationships captured monthly for reference period

- Fully edited and imputed file with ‘status’ flags - Reported, NIU, hot deck, cold deck, logical, model based, etc.

- Restricted access files available for RDC projects
March 2006
Goals for SIPP Re-engineering

- Include a new household survey data collection
- Modernize the data collection instrument
- Reduce respondent burden
- Integrate survey data and administrative records data
- Require fewer resources than the current SIPP program
- Improve processing efficiency
- Be releasable to the public in a timely manner
### Current SIPP - Calendar year 1

<table>
<thead>
<tr>
<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>Ref. I-1</td>
<td>Reference period – I-2</td>
<td>I-2</td>
<td>(Rotation Group 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ref. pd. I-1</td>
<td>I-1</td>
<td>I-1</td>
<td>Reference period – I-2</td>
<td>I-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference pd. – I-1</td>
<td>I-1</td>
<td>(Rotation Group 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference period – I-1</td>
<td>I-1</td>
<td>I-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Rotation Group 4)</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>

### Re-engineering - Calendar year 1

<table>
<thead>
<tr>
<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>Reference period</td>
<td>EHC Interview</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Interview Sequence

### Event History Calendar (EHC)
- Address verification
- Sample address characteristics and coverage
- Roster creation and demographics
- Type 2 Roster/Info
- Additional person level demographics
- Income and program screener
- Residence History
- Marriage and Cohabitation
- School Enrollment
- Programs (SSI, FS, TANF, GA, WIC)
- Health Insurance (Private, Medicare, Medicaid, Military, Other)
- Health insurance follow-up
- Dependent care expenses
- Social insurance - Annual Programs - Other GI - Lump Sum - Child Support
- Assets / Balances
- Health (Medical Expenditures - Utilization, Disability
- Fertility, Child Care, Child Well-Being
- Housing - Material Wellbeing, Food Security, Basic Needs
- Wrap-up and Missing Follow-up visit/call Info
- Interview closeout
- Contact history instrument (CHI) with Neighborhood Observations Module

### Table: Interview Year 2015 and 2016

<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>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
</tr>
</thead>
<tbody>
<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>
<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>
<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>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XYZ Corp.</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>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABC Ltd.</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>
<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>
<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>
<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>
<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>
<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>
<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>
<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>
<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>
<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>
<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>
<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>
<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>
<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>
<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>
<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>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical Assista...</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>
<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>
<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>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## SIPP 2014 vs. Classic SIPP

<table>
<thead>
<tr>
<th></th>
<th>SIPP 2014</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 calendar year</td>
<td>Previous 4 months</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>Simplified</td>
<td>Complex</td>
</tr>
</tbody>
</table>
SIPP Innovations

- Content enhancements to meet existing and new needs
- Integrated use of an Event History Calendar (EHC)
- Administrative data integration – Model-Based Imputation
- Model-based incentive assignment
- Adaptive design and case prioritization
- Monitoring
  - Computer Audio Recorded Interviewing (CARI)
  - Paradata
What is the SIPP Good For?

- Can use SIPP as a stacked sample of repeated cross sections
- Can generate monthly estimates
- Can generate annualized estimates
- Classic SIPP and current SIPP – Pay attention to recall issues and seams
- Estimates must be adjusted for sample design
- Most powerful use is for longitudinal analysis
What Isn’t It Good For?

- Studying the top of the income distribution (maybe...)
- Long longitudinal analyses (over a life course, say)
- If you need annual estimates for every year
  - (working on this one)
- If you don’t want to deal with the complexities in household/family composition that the SIPP uncovers...
- Producing state-level estimates
Use the SIPP When

- You want to deal with more of the complexity of messy questions
- You want the best possible estimate of the income of the poor
- You want to benefit from overall higher reporting rates for public program participation
- You want to conduct longitudinal analyses over relatively short periods (month-to-month; annualized, up to 4 years)
Survey of Income and Program Participation

Labor force characteristics and dynamics
Welfare duration and poverty studies
Migration
Family structure and marital status stability
Child support
Child care
Program use
Work schedule and commuting
Economic well-being
Utilization of health care services and insurance
Assets, wealth and retirement
Medical expenses and disability
Wages and inequality

Academic Research
Survey of Income and Program Participation

Federal Agency Uses

USDA (FNS): Microsimulations of food stamp eligibility and use

HHS (ACF, OASPE): Disability and Welfare Program Use; Child Care Costs; Child Support

SSA: Microsimulations: SSI benefits; Retirement & Pension benefits; Project retirement income; Disability programs benefits

DOL: Microsimulations of workers with and without pensions and health plan coverage

CBO: participants and non-participants and funds associated with major gov programs

CRS: policy research and Microsimulations: spells of unemployment, spells without health insurance, program participation rates
Survey of Income and Program Participation

Statistical Improvement Projects

- CARI use for reinterview and performance improvement
- Development of adaptive and responsive design procedures for longitudinal surveys
- Development of model based (SRMI) imputation informed by administrative records for production processing
- Development of synthetic data products to facilitate dissemination of linked data information – SIPP SSB and Gold Standard
- ERS/CARRA linking SNAP data to SIPP to understand/model formation of SNAP units
- Mortgage and Home Ownership research using SIPP linked to private mortgage and realty data
- Census research on bias from unit non-response in household surveys using SIPP and CPS linked by address to 1040s
- Augmenting household survey data with job-level administrative data. Collaboration between researchers from Census and the Michigan NCRN.

Bob Belli (UNL – NCRN) working on cognitive aspects of survey design using linked SIPP-EHC and admin data.

Joint imputation of missing data in household surveys. Jerry Reiter et al working with SIPP data at Duke NCRN.
## The Uninsured in America, 2005
### Non-Elderly (Estimates by ERIU)

<table>
<thead>
<tr>
<th>Dataset</th>
<th>During the year</th>
<th>All year</th>
<th>Point-in-time</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPS</td>
<td>N/A</td>
<td>44.4 million</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>17.2%</td>
<td></td>
</tr>
<tr>
<td>SIPP</td>
<td>65.9 million</td>
<td>27.6 million</td>
<td>45.2 million</td>
</tr>
<tr>
<td></td>
<td>25.6%</td>
<td>10.7%</td>
<td>17.6%</td>
</tr>
<tr>
<td>MEPS</td>
<td>66.9 million</td>
<td>36.5 million</td>
<td>50.1 million</td>
</tr>
<tr>
<td></td>
<td>25.8%</td>
<td>14.0%</td>
<td>19.2%</td>
</tr>
</tbody>
</table>
Data Access

- Downloaded from FTP - http://thedataweb.rm.census.gov/ftp/sipp_ftp.html
- Access through DataFerrett – http://dataferrett.census.gov/
- Orlin Research SIPP Data Tool – Coming soon
SIPP 2014: Collection and Release Schedule

Wave 1 (Collected February-June 2014)
- Research file releases – February, July, and November 2016
- Available for RDC use independently of Research File – early 2017

2014 Social Security Administration Supplement on Retirement, Pensions and Related Content (Collected in September and October 2014)
- Telephone interview follow-up
- Released by June 2017
- Dependent on Wave 1 edited input

Wave 2 (Collected February-May 2015)
- Full public use release – by end of December 2017

Wave 3 (Collected April-July 2016)
- Full public use release – by end of June 2018

Wave 4 (Collection expected February – May 2017)
- Full public use release – by end of December 2018
THANK YOU!

Jason Fields
Survey Director
Survey of Income and Program Participation
National Survey of Children’s Health
Jason.M.Fields@census.gov

www.census.gov
www.census.gov/sipp