Introducing the Survey of Income and Program Participation (SIPP)

Matthew Marlay
U.S. Census Bureau
Agenda for the Webinar

1. Survey background and overview
2. Content overview
3. Processing overview
4. Data overview
Survey Overview
SIPP’s Mission

Provide a nationally representative sample to evaluate:

• Annual and sub-annual dynamics of income
• Movements into and out of government transfer programs
• Family and social context of individuals and households
• Interactions between these items
SIPP’s Advantages

• Comprehensive, detailed data on a variety of topical areas
• Measures dynamics of:
  – Households
  – Families
  – Individuals
• Longitudinal
• Follows movers
Current Status

2008 Panel:
• Wrapped up interviews in December 2013
• All files now available for data users
• Panel bridges Great Recession
  —Provides data over five-year period, spanning crash and recovery
  —Monthly, full-sample data from August 2008-May 2013
Current Status

2014 Panel:
• Wave 1 interviews completed June 2014
  – Dataset now available
• SSA supplement interviewed in fall 2014
  – Dataset now available
• Wave 2 interviews completed June 2015
  – Dataset available mid-2018 (tentative date)
• Wave 3 interviews completed July 2016
• Wave 4 interviews completed June 2017
Current Status

2018 Panel:

• Wave 1 interviews begin February 2018
  – Currently preparing instrument, materials, training, etc.
  – Anticipate sampling ~53,000 households for Wave 1

• Panel intended to last four years (through 2021)

• Design very similar to 2014 design
### 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>June 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.
SIPP Design

• Survey universe is the civilian, non-institutional population of the United States
• Annual interviews
• Adults (age 15+) interviewed in Wave 1
  – Data collected for all people
  – Proxy interviews for children under 15
• We follow all interviewed adults in subsequent waves
SIPP Design: Following Rules

• Everyone in the Wave 1 household is considered an Original Sample Person (OSP)
• In subsequent waves, we attempt to interview every OSP, as well as any new people who have moved in with the OSPs
• For OSPs who leave the original household, we will follow them as long as they remain in the survey universe (civilian, non-institutionalized, in the U.S.)
SIPP Design: Feedback

• In Waves 2+, we feed back data from the previous wave’s interview to shorten the current interview and improve data quality
  —Aka “Dependent data” or “Preload data”
  —Respondents have to affirmatively consent to having their data fed back

• For example, we feed back demographic info, since that is unlikely to change from wave to wave
SIPP Design: Sample

• Sample universe mostly based on the 2010 Decennial Census frame
  – 2014 SIPP first Census Bureau survey to use updated frame
• Drawn from Master Address File (MAF)
  – Un-duplication with other surveys
  – Housing unit eligible for sampling 5 years after their last interview
• Generally follows population distribution

Adapted from Westra and Sundukchi (2017)
SIPP Design: Weighting

• The **weight** is the number of units in the population that the responding unit represents.

• Weights intended to make final dataset representative of survey universe (civilian, non-institutionalized U.S. population)

Adapted from Westra and Sundukchi (2014)
Content Overview
Some Important Definitions

• **Respondent**: Someone in the household for whom SIPP collects information

• **Household respondent**: The first eligible adult household member interviewed

• **Reference person**: The owner or renter of the housing unit (the first, if more than one resident owns/rents the unit)
Some Important Definitions

- **Type 2 person**: A person who lived with a respondent during the reference year but not at time of interview
  - Do not get a person record
  - Can be at any address a respondent lived at, not just interview address
- **Adult**: Someone 15 years old or older
- **Family**: A set of people related by blood, marriage, or adoption
- **Household**: A set of people living together
Some Important Definitions

• **Self vs. proxy**: How we collect information for a given respondent
  – **Self-response**: Respondent provides information for him/herself
  – **Proxy response**: Another adult in the household provides information for respondent

• Everyone under 15 is a proxy by design – an adult provides a child’s interview information
2014 SIPP: Content Summary

- Opening Screens
- Roster
  - Sex
  - Birthdate/Age
- Demographics
  - Hispanic origin
  - Race
  - Citizenship
  - Language
  - Marital status
  - Parent/child relationships
  - Educational attainment
  - Armed forces status
  - Type 2 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
  - 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
2014 SIPP: Content Highlights

Demographics:
- Roster: Interview-month household residents
- Roster of people you lived with during the year but who aren’t in the interview month household (Type 2)
- Age, race, Hispanic origin, sex
- Relationship to householder, including options for same-sex partners and spouses
- Marital status, spouse pointer, year of marriage, times married, ever widowed, ever divorced, fertility screener
2014 SIPP: Content Highlights

Demographics:

- Gender neutral parent identification (Parent 1 and Parent 2)
  - Facilities same-sex parent identification
- Type of parent for parent 1 and 2
- Nativity, citizenship, year moved to U.S., immigration status
- Education: Attainment, vocational/technical schooling, educational certificates, professional certifications and licenses
- Armed forces service: Veteran status, period of service, active duty
## 2014 SIPP: Content Highlights

<table>
<thead>
<tr>
<th>Topic</th>
<th>Reference Year 2013</th>
<th>Interview Year 2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residency</td>
<td>Jan</td>
<td>Feb</td>
</tr>
<tr>
<td>Marital History</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job 6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>More Jobs (if any)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Job</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Food Stamps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TANF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gen. Assist.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WIC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medicare</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical Assistance...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Military</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Coverage</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

United States Census Bureau
Economics and Statistics Administration
U.S. Census Bureau
census.gov
2014 SIPP: Content Highlights

Residency:
- Up to 5 residences during the reference period for each person
- Tenure status of each residence
- Public housing status: Housing subsidy receipt, voucher receipt
- Reason for moving to address
- Left censor: Year and month moved into January address, tenure of prior residence
2014 SIPP: Content Highlights

Marital History:
- Up to 3 marital status spells
- Monthly marital status, with spouse pointer
- Monthly cohabitation status, with partner pointer
- Registered domestic partner item for cohabitations
2014 SIPP: Content Highlights

Educational Enrollment:

- Up to 3 spells of educational enrollment
- Grade attended
- School type (public, private, charter, home)
- Enrollment type (full-time/part-time)
- Grade repetition
- Head Start for children 7 and under
2014 SIPP: Content Highlights

Labor Force:
- Up to 7 discrete job/business timelines, each with 2 spells possible
- Timeline for additional work beyond the first 7 jobs
- Type of pay/pay rate
- Job earnings and business profits
- # of hours usually worked per week
- Changes in earnings or hours worked (up to 3)
- Industry, occupation, and class of worker
- Business/employer name, size, type, address
- Union status
- Incorporation status
- Presence of partners (business)
- Time away without pay
2014 SIPP: Content Highlights

Unemployment:

- Spells of unemployment
- Time away from work
- Time out of the labor force
- Reason not working
- Availability for work
- Reason not available for work
2014 SIPP: Content Highlights

Commuting:
- Mode(s) of travel
- Time to work
- Miles to work
- Costs for a typical week
- Reimbursement of costs

Work Schedule:
- Days worked
- Start and stop times
- Working from home
- Type of schedule
- Reason for schedule
2014 SIPP: Content Highlights

Programs:

- Supplemental Security Income (SSI)
- Temporary Assistance for Needy Families (TANF)
  - Pass-through child support payments
- Supplemental Nutrition Assistance Program (SNAP, aka Food Stamps)
- General Assistance (GA)
- Women, Infants, and Children’s Nutrition Program (WIC)
2014 SIPP: Content Highlights

Programs:

- Up to 3 spells per reference period recorded for each program
- Who in the household is covered
- Ownership of the coverage
- Reasons for starting and stopping
- Amount received each month
- Up to three changes in amounts
2014 SIPP: Content Highlights

Health Insurance (in EHC):

- Private, Medicare, Medicaid, Military, and Other Coverage timelines
- Two timelines for private coverage to allow overlapping spells
- Who in the household is covered
- Ownership of the coverage
- Whether anyone outside household covered
- Type of coverage
- Cost
- Type of private plan
- Type of deductible
- ACA/exchange participation (Waves 2+)
Health Insurance Follow-up (post-EHC):

- ACA/exchange participation (Wave 1)
- Premiums and reimbursement
- Reconciliation of time without coverage
- Reasons no private if employed but not covered
- Reasons no public if not covered at all
- Reasons for not being covered
2014 SIPP: Content Highlights

Programs (Annual and Other):

- Paid care of children or disabled persons so that a person could work, attend training, or look for work in December of reference year
- Total cost of that care in December
- Social Security retirement income and deductions for Medicare (monthly amts.)
- VA benefits (monthly amts.)
- Worker’s compensation (monthly amts.)
- Unemployment compensation (monthly amts.)
- Energy assistance
2014 SIPP: Content Highlights

Programs (Annual and Other):

- Free and reduced-price meal programs
- Lump-sum payments
- Disability income
- Retirement income
- Miscellaneous cash income
- Survivor benefits
- Child support and alimony received
- Support payments made
- EITC receipt and tax filing status
- Other training, food, clothing, cash, and housing assistance
2014 SIPP: Content Highlights

Assets – Income, Balances/Values, and Joint Holdings for:

- Checking and savings
- Money market accounts or funds
- Certificates of deposit
- Mutual funds
- Stocks
- Municipal and corporate bonds
- Government securities and savings bonds
- Royalties
- Rental property and rental property mortgage balance
- Other mortgages
- Miscellaneous investments
2014 SIPP: Content Highlights

Other Assets:

- Real estate owned, market value, and balance owed
- Vehicles (up to 3 per household):
  - Use
  - Make
  - Model
  - Year
  - Balance owed
- Business value and debt
- Retirement account balances
- Unsecured liabilities
2014 SIPP: Content Highlights

Health Care Utilization and Medical Expenditures:
- Overall health status
- Medical out-of-pocket expenses
- Health care utilization
- Hospitalization
- Sick days
- Doctor visits
- Dental visits/teeth lost
- Drug coverage
- Insurance premium payments
- Medical visits by uninsured
Disability:

- Everyone: Sensory disabilities (sight and hearing)
- Adults: Serious difficulty concentrating, remembering, making decisions, walking or climbing stairs, dressing or bathing, doing errands, finding a job or remaining employed, prevented from working
- Children: Serious difficulty concentrating, remembering, making decisions, walking or climbing stairs, dressing or bathing, playing with children of the same age, doing regular school work
- Young children: A developmental condition or delay that limits ordinary activity
- SSA: Detailed adult, child, and work disability sections in the 2014 supplement
2014 SIPP: Content Highlights

Fertility:
- Roster and ages for children birthed/fathered
- Identification of ‘other parent’ if not resident, enabling a multiple partner fertility measure
- Grandparent indicator

Parents’ Nativity and Mortality:
- Asked about biological parents
- Collected for both mother and father
- Month/year of birth
- Country of birth
- Mortality status
- Month/year of death
2014 SIPP: Content Highlights

Child Care:
- Type of arrangements used
- Which children used each
- Weekly amount paid for child care
- Assistance with costs
- Time lost from work related to child care

<table>
<thead>
<tr>
<th>EVENT HISTORY CALENDAR (EHC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Address verification</td>
</tr>
<tr>
<td>Sample address characteristics and coverage</td>
</tr>
<tr>
<td>Roster creation and demographics</td>
</tr>
<tr>
<td>Type 2 Roster/Info</td>
</tr>
<tr>
<td>Additional person level demographics</td>
</tr>
<tr>
<td>Income and program screener</td>
</tr>
<tr>
<td>Residence History</td>
</tr>
<tr>
<td>Marriage and Cohabitation</td>
</tr>
<tr>
<td>School Enrollment</td>
</tr>
<tr>
<td>Labor Force</td>
</tr>
<tr>
<td>(Jobs, Businesses, Contingent Work, Unemployment, Not In Labor Force, Commuting, Work Schedule)</td>
</tr>
<tr>
<td>Programs (SSI, FS, TANF, GA, WIC)</td>
</tr>
<tr>
<td>Health Insurance</td>
</tr>
<tr>
<td>(Private, Medicare, Medicaid, Military, Other)</td>
</tr>
<tr>
<td>Health insurance follow-up</td>
</tr>
<tr>
<td>Dependent care expenses</td>
</tr>
<tr>
<td>Social Insurance - Annual Programs - Other GI - Lump Sum - Child Support</td>
</tr>
<tr>
<td>Assets / Balances</td>
</tr>
<tr>
<td>Health (Medical Expenditures - Utilization, Disability)</td>
</tr>
<tr>
<td>Fertility, Child Care, Child Well-Being</td>
</tr>
<tr>
<td>Housing - Material Wellbeing, Food Security, Basic Needs</td>
</tr>
<tr>
<td>Wrap-up and Missing Follow-up visit/call Info</td>
</tr>
<tr>
<td>Interview closeout</td>
</tr>
<tr>
<td>Contact history instrument (CHI) with Neighborhood Observations Module</td>
</tr>
</tbody>
</table>
Child Well-Being:
- Eating dinner with parents
- Reading to and outings with children under 6
- School engagement, school effort
- Grade repetition, suspension, expulsion
- Gifted classes
- Activities: Sports, lessons, club participation, religious lessons

Adult/Material Well-Being:
- Environmental noise, trash, safety
- Problems with housing, pests, plumbing
- Ability to pay mortgage or utilities
2014 SIPP: Content Highlights

Food Security:

- Able to buy enough food
- Able to eat balanced meals
- Cut size of meals, and, if so, how often
- Defer food to children
- Hungry because not enough money for food
Processing Overview
New Processing System

• Besides a fully redesigned instrument, the 2014 data benefits from a fully redesigned processing system
• Rewrote and recoded all of the editing logic from scratch
• Moved programs from FORTRAN to SAS
• However, the basics of our processing strategy remain the same
Item-Level Imputation

When a respondent is in universe for a question but does not provide an answer, we impute the missing value via:

1. Logical imputation
2. Hot-deck imputation
3. Cold-deck imputation
Item-Level Imputation

• **Logical imputation** uses the respondent’s answers to related questions to derive the missing value

  – For example, if the respondent does not provide a state of birth for one child but an older and younger sibling were both born in Texas, we may set the missing state of birth to Texas.

• This method is also used to edit provided answers for logical consistency.
Item-Level Imputation

• **Hot-deck imputation** substitutes the answer provided by a similar respondent for the missing value
  
  – For example, if the respondent does not provide a tenure value but similar respondents own their residences, we may set the missing tenure to ‘Owned’.

• Donors usually determined by demographic characteristics such as age, race, educational attainment, etc.

• Reproduces distribution of reported data
Item-Level Imputation

• Cold-deck imputation substitutes a default answer for the missing value.
• We never want to use a cold-deck value, but it is there as a last resort.
• Usually the most common answer for the question.
Whole-Person Imputation

• In cases where we have no data at all for a person, previous SIPP panels used a similar person as a donor and substituted for the missing person.

• This is a fairly crude method of imputation.

• For 2014, replaced by...
Model-Based Imputation

• Replaces existing whole-person imputation strategy with parametric model-based approach
  – Allows inclusion of many more predictor variables
  – SIPP Synthetic Beta (SSB) file provides the methodological foundation for modelling
  – Use administrative data to mitigate problems caused when survey data are not missing at random
    • Draws in data from outside the pool of survey respondents
Data Overview
Data Resources

- SIPP website (http://www.census.gov/sipp)
- Census FTP site (http://thedataweb.rm.census.gov/ftp/sipp_ftp.html)
- SIPP Synthetic Beta (Accessible via SIPP website)
- NBER SIPP page (http://www.nber.org/data/sipp.html)
Technical Documentation

• SIPP website (http://www.census.gov/sipp) is your best resource

• Currently available:
  – Users’ Guide
  – Metadata
  – Release notes
  – User notes
  – Codebook
  – Crosswalks (for 2008-2014 and 2014-2008)
Data Tips

• Our datasets are provided in SAS or ASCII format
  – 2014 Wave 1 file also available in Stata format

• For Stata users, a good resource is Savastata
  – User-driven Stata command that saves SAS datasets as Stata datasets
  – \texttt{http://www.cpc.unc.edu/research/tools/data_analysis/sas_to_stata/transfer-tools/savastata.html}
  – A parallel command goes in the opposite direction

Adapted from Luke Shaefer (2015)
Data Management

• SIPP files have many variables for many observations
  – Can lead to serious memory limitations
• Check the capacity of your machine
  – Worth working on a well-equipped machine
• When you load in a dataset, keep only the observations and variables you need

Adapted from Luke Shaefer (2015)
SIPP Pros and Cons

• SIPP is good for:
  – Generating monthly national-level estimates
  – Generating annual estimates
  – Longitudinal analysis over the panel
  – Analyzing the income of the poor

• SIPP is not as good for:
  – Long-term longitudinal analyses
  – Analyzing the top of the income distribution

Adapted from Luke Shaefer (2015)
Use 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)

Adapted from Luke Shaefer (2015)
SIPP – General Research Areas

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
- Wages and inequality
- Medical expenses and disability
- Economic well-being

Academic Research
SIPP – Some Uses by Federal Agencies

**Survey of Income and Program Participation**

- **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
SIPP – Statistical Improvement

Survey of Income and Program Participation

-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.

-Joint imputation of missing data in household surveys. Jerry Reiter et al working with SIPP data at Duke NCRN.

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

Census.SIPP@census.gov
http://www.census.gov/sipp
(888) 245-3076