From the Ashes: Re-envisioning and Re-building the Survey of Income and Program Participation

Jason Fields
SIPP Survey Director
U.S. Census Bureau

August 29, 2017 Joint Statistical Meeting – American Statistical Association Baltimore, MD





Outline

- A brief look back to 2006
- Ask for and value input embrace transparency
- Challenges and the reality of a full re-engineering
- Many successes and lessons learned





Purpose of SIPP

- "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

Paper SIPP Interviewing (1984-1993)

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 Panel Sizes and Collection Periods

E	lig	ib	le
---	-----	----	----

Elibroic					
Panel	Number of Waves	Households	Date of First Interview	Date of Last Interview	Notes
2014	. 4	42,348	Feb. 2014	May 2017	(1)
2008	16	52,031	Sept. 2008	Dec. 2013	(2)
2004	12	51,379	Feb. 2004	Jan. 2008	
2001	. 9	50,500	Feb. 2001	Jan. 2004	
1996	12	40,188	Apr. 1996	Mar. 2000	(3)
1993	9	21,823	Feb. 1993	Jan. 1996	
1992	10	21,577	Feb. 1992	May 1995	
1991	. 8	15,626	Feb. 1991	Sept. 1993	
1990	8	19,800	Feb. 1990	Sept. 1992	
1989	3	12,867	Feb. 1989	Jan. 1990	
1988	6	12,725	Feb. 1988	Jan. 1990	
1987	7	12,527	Feb. 1987	May 1989	
1986	7	12,425	Feb. 1986	Apr. 1988	
1985	8	14,306	Feb. 1985	Aug. 1987	
1984	. 9	20,897	Oct. 1983	Jul. 1986	

⁽¹⁾ The 2014 Panel is the first EHC panel with annual interviewing.

⁽³⁾ This is the first CAPI SIPP panel, and first of the non-overlapping panels.





⁽²⁾ The 2008 Panel start was delayed due to budget and extended into 2013 to overlap with the 2014 Panel

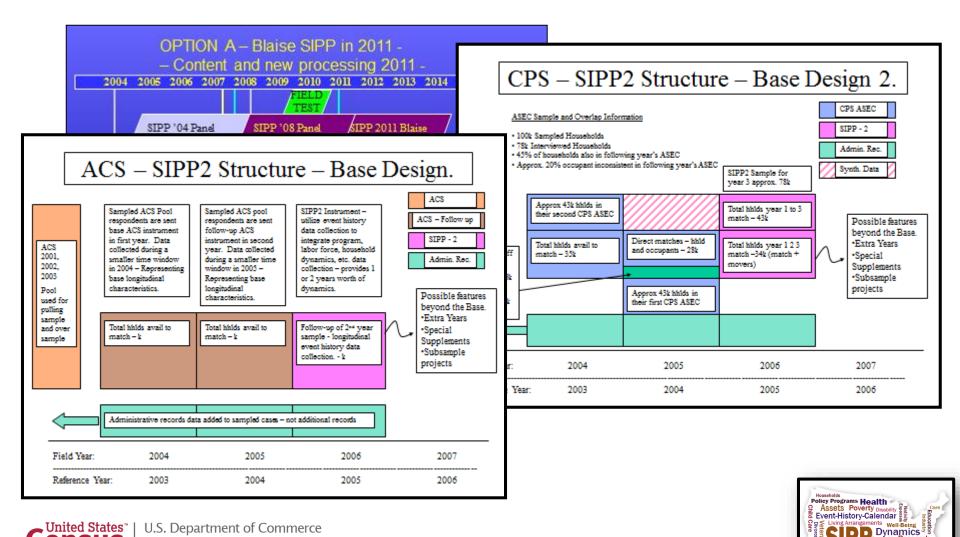
March 2006







We began brainstorming options





Ownership Research
WIC Health Insurance

June 2006 @ Brookings

Possible sample sources

- · There are several options for generating a sampling frame:
 - sampling from the Master Address File;
 - sampling directly from ACS interviewed cases;
 - sampling directly from administrative data;

Possible modes of data collection

Mail out/mail back



Administrative records successes

Medicare Enrollment Database (MEDB)

Master Beneficiary Record (MBR)

IRS 1040, 1099-INT, and 1099-DIV

Census Numident

SSA SSR

HUD-TRACS

Medicare health insur-

Work disability and so income (e.g., Social S

Asset ownership, inco a job, profit from a buunemployment compe of alimony

Demographic charact age, race, Hispanic or

Receipt of Federal/sta

Public housing and re of rent subsidies

Questions

- What aspects of the current SIPP do you feel are the most important for measuring economic wellbeing? And in particular, what would you be willing to give up?
- Do you have examples of research (or products) that you regard as "best practices" in the use of SIPP data?
- · Do you have needs for both the cross-sectional and longitudinal aspects of the data?

USCENSUSBUREAU

Helping You Make Informed Decisions

United States™

USCENSUSBUREAU

Helping You Make Informed Decisions

12



Goals for SIPP Re-engineering

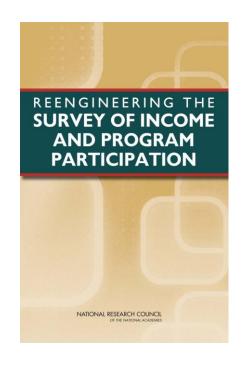
- Include a new household survey data collection
- Modernize the data collection instrument
- Reduce respondent burden
- Integrate survey data and administrative records data
- Require fewer resources than the current SIPP program
- Improve processing efficiency
- Be releasable to the public in a timely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





CNSTAT Reports on SIPP





Evaluation Of 2014 SIPP

In development

1993

2009

2017





Recommendations from the 2009 NAS Report

- Rec 2.1 Goal is short-run dynamics
- Rec 2.2 Evaluate all innovations
- Rec 3.1 Acquire more admin data from Federal sources
- Rec 3.2 Develop plan to obtain admin data from States
- Rec 3.3 Evaluate data quality and reporting errors
- Rec 3.4 Evaluate imputation methods
- Rec 3.5 Have OMB set-up SIPP advisory group
- Rec 3.6 In short run focus on indirect uses of admin data
- Rec 3.7 Evaluate possible direct uses of admin data
- Rec 3.8 Develop methods to create public data and data access
- Rec 4.1 Develop intensive plan to evaluate EHC
- Rec 4.2 Create a bridge between EHC and current SIPP
- Rec 4.3 Don't rush implementation (shoot for 2012)
- Rec 4.4 Evaluate trade-offs with data quality and respondent burden
- Rec 4.5 Establish SIPP advisory group
- Rec 4.6 Release data within one year of collection







Challenges & reality of a full re-engineering

- When can you have it done?
 - First thoughts new data in 2009
 - Quickly determined the need for thoughtful and more comprehensive redesign
- Make your decisions and move forward
- Work issues thoroughly from beginning to end
- Innovate!
 - Biggest gains and what direction?
 - Respondent burden
 - Administrative data
 - Modeling
 - Monitoring
 - Data quality
- Challenges
 - New processing system
 - Field staffing, training, and monitoring
 - More for less





Challenges

Instrument design

- Blaise and C# integration
- Ability to allow conversational collection and navigation

Fieldwork

- Hiring
- Training
- Retention

Data processing

- Create in SAS from new comprehensive specs
- Changing data structure and content through development
- Once file structure available reconciling the timeline to develop, test, correct

Expectations



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





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





What is the SIPP Good For?

- Estimates of the income for the majority of the population
- Focus is on <u>eligibility and take-up of public transfer and</u> <u>assistance programs</u>
- Focus on inter-related topics and the <u>complexity of</u> <u>messy questions</u>
- You want to conduct longitudinal analyses over relatively short periods (month-to-month; annualized, up to 4 years)
 - Classic SIPP and current SIPP Pay attention to recall issues and seams
 - Estimates must be adjusted for sample design





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



Lessons learned (and still learning)

Where is the time?

- Instrument design (iterative)
- Processing development (need to work from stable platform to avoid rework)

Successes

- Flexible data collection stable instrument
- Training and evaluation
- Supporting stakeholders with integrated and updated content
- Administrative data integration

Lessons

- Timeline expectations
- Response, cost, quality, and burden
- Holistic data editing many decisions need to be made up front
- Opportunity for innovation and a blank sheet





Enduring strategies and principles

Plan for iterative development

- Nimble sprints and plan for cascading changes
- Final collection and format decisions need to be as early as possible
- Consider the impact of rework to instruments and processing

Embrace innovation

- Foundation for "what's next"
- For SIPP Focus on data quality and providing a platform for innovation



Enduring strategies and principles

- Extend your vision
 - Impact to hiring and collection represent ongoing commitments
 - Data editing and processing development can't easily be done up front
 - Take the time to outline the minimum and ideal development once the final goal has been identified
 - Redeveloping an existing product is different than creating from scratch



Enduring strategies and principles

- Leverage administrative data opportunities
 - Can be integrated into production modelling absorbs data lags
 - Validation during development against test data can be planned
 - Explore indirect uses locating, imputation adjustments, frame and sampling

- Keep communicating and working with stakeholders and users
 - Provide early access to data
 - Invite input and review





Release Plans

- Wave 1 (Collected February-June 2014)
 - Research file releases so far February, July, and November 2016
 - Full public use release March 2017
 - Available for RDC use independently of Research File March 2017
- 2014 Social Security Administration Supplement on Supplement on Retirement,
 Pensions and Related Content
 - Telephone interview follow-up
 - June 2017
 - Dependent on Wave 1 edited input
- Wave 2 (Collected February-May 2015)
 - Full public use release by end of 2017
- Wave 3 (Collected April-July 2016)
 - Full public use release by end of September 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/sipp

