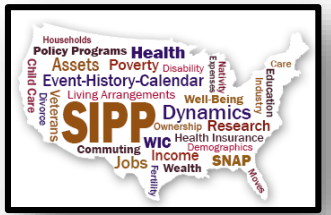
From the Ashes: Re-envisioning and Re-building the Survey of Income and Program Participation

Jason Fields
SIPP Survey Director
U.S. Census Bureau

December 6-7, 2016
Federal Committee on Statistical Methodology
Washington DC

United States Census Bureau
U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov



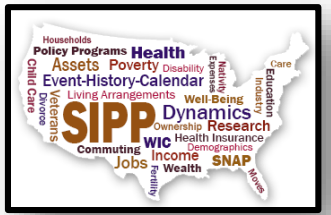
Outline

- 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**



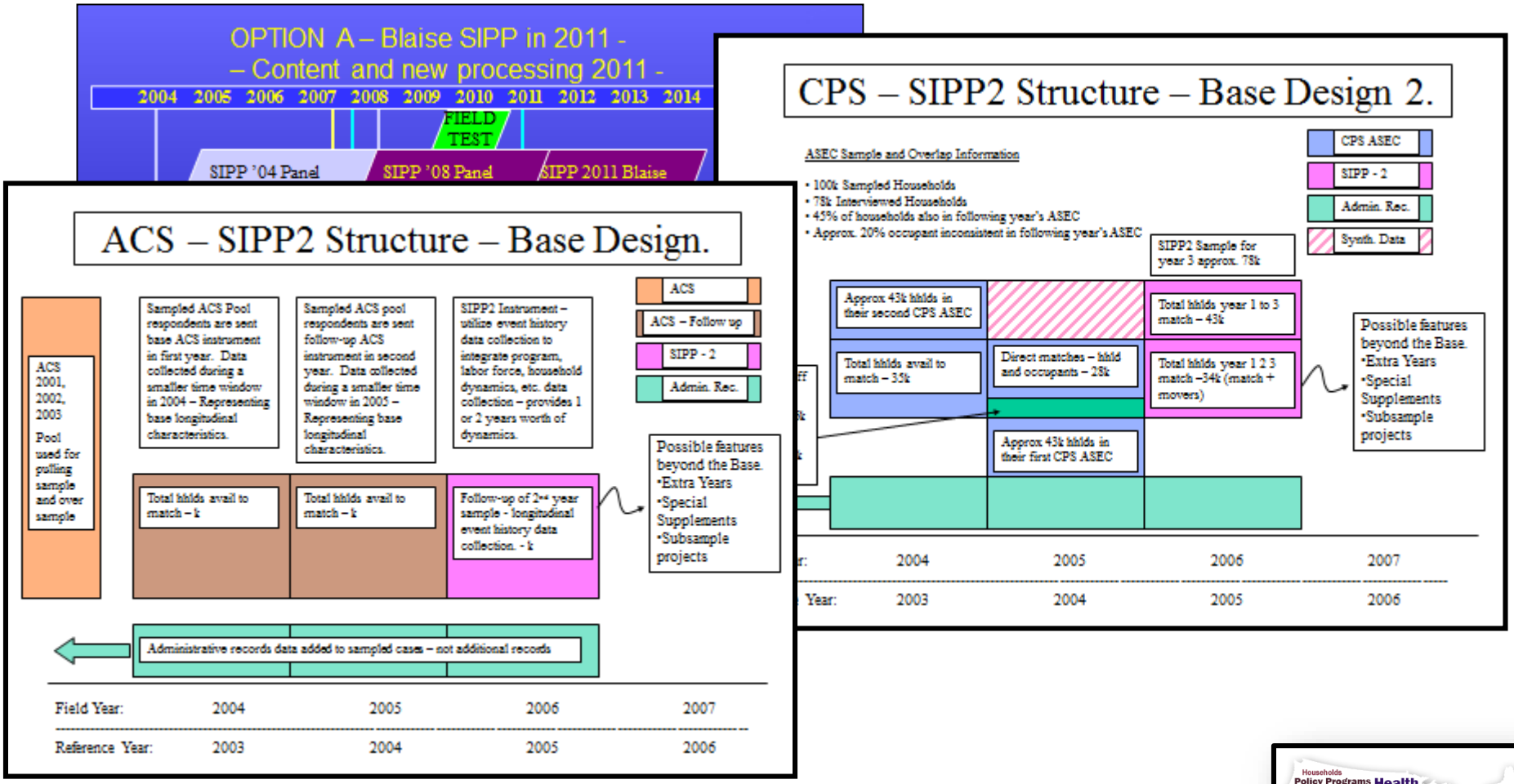
The SIPP Mission

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.

March 2006

We began brainstorming options



Ask for & value input embrace transparency

- Reach out to stakeholders and concerned advisors
  - Federal Agencies
  - CNSTAT
  - Associations and policy organizations
  - Academic researchers

- *"Save the SIPP"* (still Thank You!)

- Adjust to changing priorities

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

12

Administrative records successes

Medicare Enrollment Database (MEDB)	Medicare health insur
Master Beneficiary Record (MBR)	Work disability and sc income (e.g., Social S
IRS 1040, 1099-INT, and 1099-DIV	Asset ownership, inco a job, profit from a bu unemployment compe of alimony
Census Numident	Demographic charact age, race, Hispanic or
SSA SSR	Receipt of Federal/sta
HUD-TRACS	Public housing and re of rent subsidies

USCENSUSBUREAU
Helping You Make Informed Decisions

Questions

- What aspects of the current SIPP do you feel are the most important for measuring economic well-being? 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

22

Stakeholder Matrix Meetings

Variable Name (SIPP 1993 Longitudinal File)

Current SIPP Variable	Description	Section	Annual/Monthly	Critical Stakeholder Variable (Y/N)
ASST140	Royalties - ISS Code 140	Assets		
ASST150	Other financial investments, other - ISS Code 150	Assets		
ATT-SCHL	School enrollment - full time/part time	Demographics		
BUS-ID1	Check Item S1 - First business ID number	Labor Force		
BUS-ID2	Check Item S1 - Second business ID number	Labor Force		
CAIDCOV	Medicaid coverage	Health Insurance		
CARECOV	Medicare coverage	Health Insurance		
CHAMP	Health insurance this month (military)	Health Insurance		
CLSSWRK1	Employee of which type of organization - first job	Labor Force		
CLSSWRK2	Employee of which type of organization - second job	Labor Force		
DISAB	Disability that limits type of work	Health Insurance		
ED-FINAN	Educational financing: GI Bill, Pell Grant, SEOG, et. al.	Education		

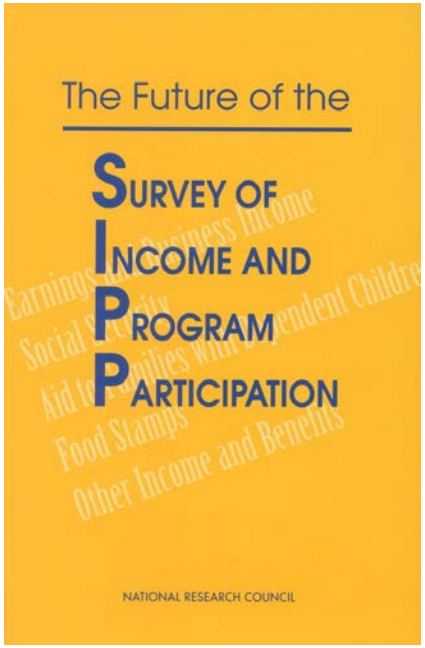
Goals for SIPP Re-engineering

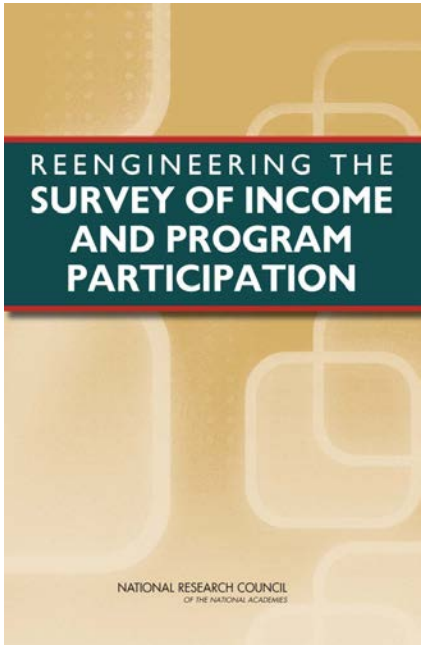
- Include a new household survey data collection
- Modernize the data collection instrument
- Reduce respondent burden
- Integrate survey data and administrative records data
- Require fewer resources than the current SIPP program
- Improve processing efficiency
- Be releasable to the public in a timely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

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

CNSTAT Reports on SIPP





Evaluation
Of 2014 SIPP

In development

1993 **2009** **2017**

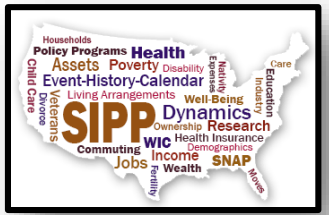
Recommendations from the 2009 NAS Report

- 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

Recommendations from the 2009 NAS Report

Recommendations in process

- 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

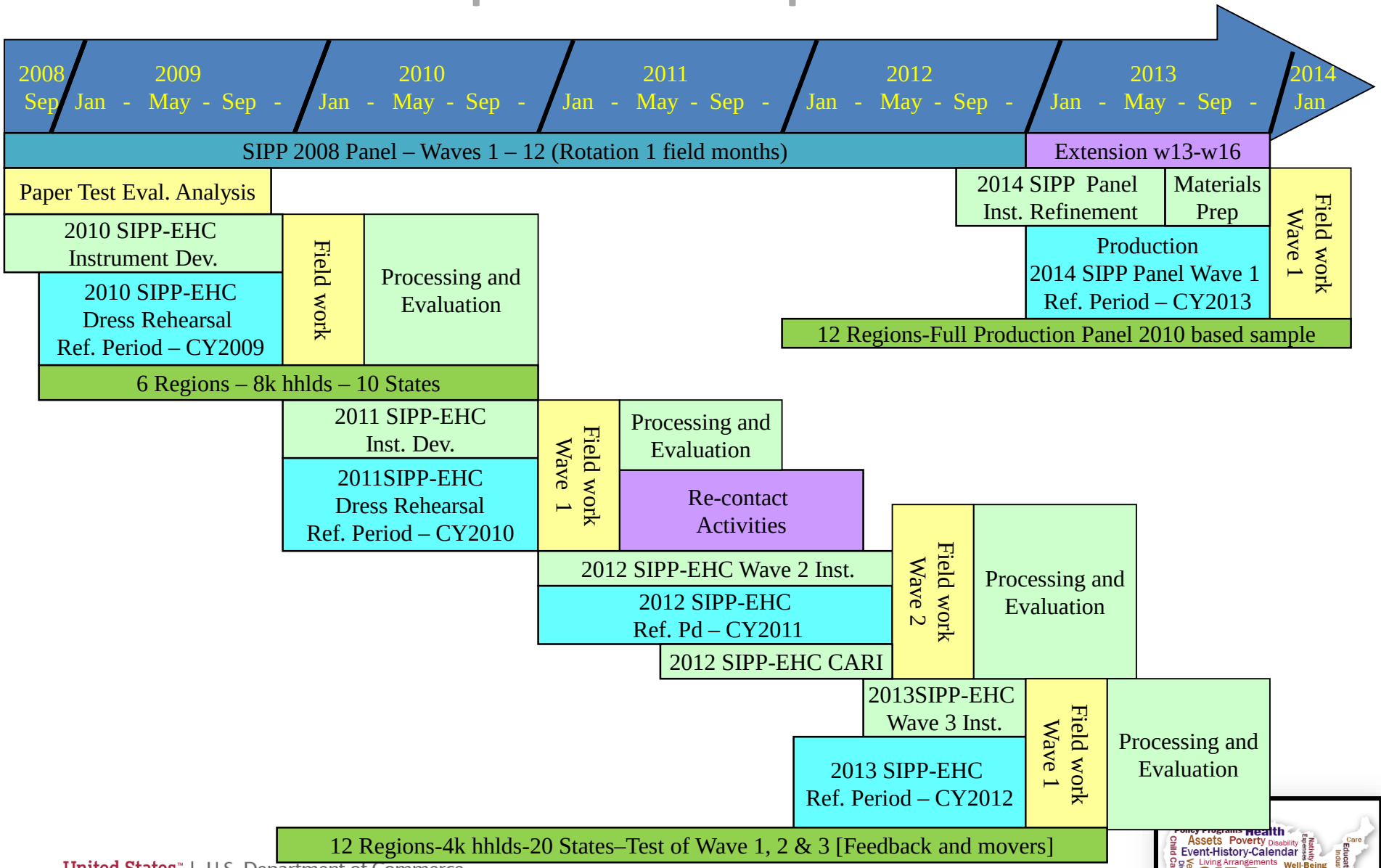


Recommendations from the 2009 NAS Report

Recommendations in process
Recommendations on a longer timeline

- 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

SIPP-EHC Development and Implementation for 2014



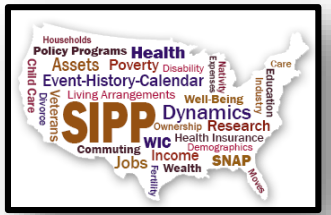
United States Census Bureau | U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

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

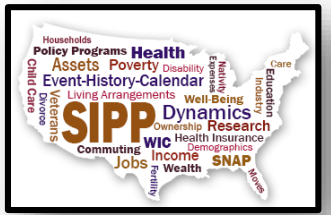
Challenges

- 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



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 – early 2017
  - Available for RDC use independently of Research File – early 2017

- 2014 Social Security Administration Supplement on Supplement on Retirement, Pensions and Related Content
  - Telephone interview follow-up
  - By June 2017
  - Dependent on Wave 1 edited input

- Wave 2 (Collected February-May 2015)
  - Full public use release – by end of September 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.M.Fields@Census.Gov

HTTP://www.census.gov/SIPP