

# Demographic Survey Overview

Census Scientific Advisory Committee  
September 17-18, 2015 || U.S. Census Bureau

Eloise Parker

Assistant Director for Demographic Programs  
Survey Operations

==

Jason Fields

Survey Director, Survey of Income & Program Participation

# Overview

- Demographic Survey Management Structure
- Current Portfolio of Surveys
- Addressing Survey Challenges
  - Data-Driven Production Management
  - Managing Strategies to Improve Production
  - Research on Survey Redesign
- Case Study in Survey Redesign: Lessons Learned in Redesigning SIPP

# The Basics

- Demographic Survey Operations reorganized in 2013
- All household surveys managed under one roof
- Consolidates resources and expertise to facilitate efficiency and shared learning
- Matrixed management across survey teams and partnering divisions (e.g., Statistics, Research, IT, Field)
- Consistency in project management and quality standards
- Survey Operations Coordination Office

# Survey Portfolio

Survey	Sponsoring Agency
<b>Labor Force</b>	
Current Population Survey	Bureau of Labor Statistics/Census
National Survey of College Graduates	National Center for Science & Engineering Statistics
<b>Economic Well-Being</b>	
Consumer Expenditure Surveys	Bureau of Labor Statistics
Telephone Point of Purchase Survey	Bureau of Labor Statistics
Survey of Income & Program Participation	Census
American Time Use Survey	Bureau of Labor Statistics
<b>Crime</b>	
National Crime Victimization Survey	Bureau of Justice Statistics
<b>Health</b>	
National Health Interview Survey	National Center for Health Statistics
National Ambulatory Medical Care Surveys	National Center for Health Statistics
National Survey of Children's Health	Maternal & Child Health Bureau

# Survey Portfolio

Survey	Sponsoring Agency
<b>Housing</b>	
American Housing Survey	Housing & Urban Development
New York City Housing Vacancy Survey	City of New York
<b>Education</b>	
National Household Education Survey	National Center for Education Statistics
National Teacher and Principal Survey	National Center for Education Statistics
<b>Recreation</b>	
National Survey of Fishing, Hunting and Wildlife-Associated Recreation	Fish & Wildlife Service

- Multiple, revolving supplements move in and out of core surveys.

# Challenge & Opportunity

## Challenges

- Declining response rates translates to higher costs and impacts to data quality
- Respondents less willing to engage with traditional CAPI and CATI survey designs
- Produce more timely and relevant data for the same money, without adding respondent burden
- Mitigating 21<sup>st</sup> century disclosure risks
- Modernizing without impacting data series

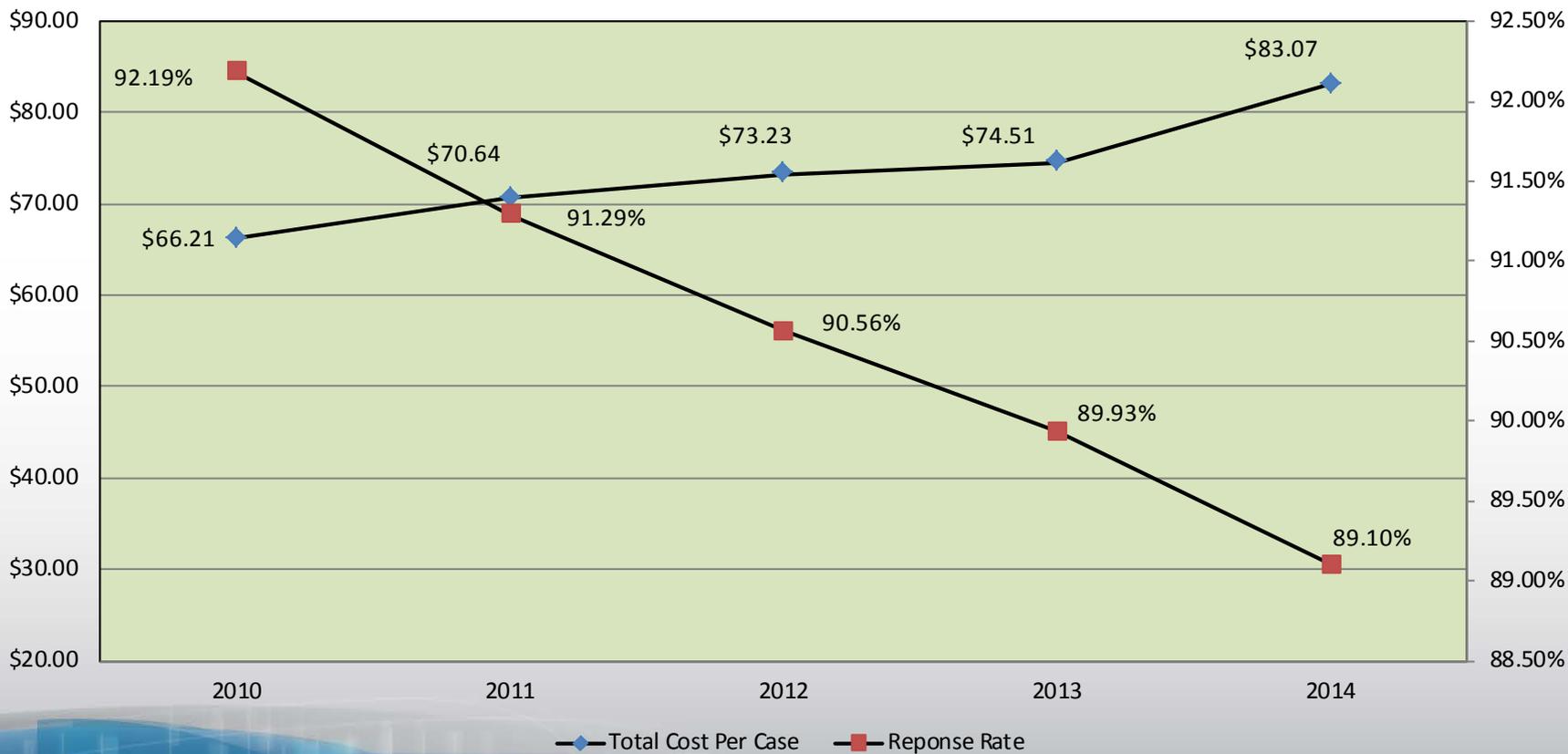


## Opportunities

- New Office of Survey & Census Analytics strengthening the Bureau's ability to produce actionable paradata analysis
- Expanding use of adaptive design techniques
- Exploring the potential of administrative data to enhance or replace survey data
- Collaborating with survey sponsors on design research and testing:
  - Effective Sampling
  - Questionnaire Restructuring
  - Mode Experimentation
  - Incentive Strategies
  - Respondent Communication
  - Disclosure

# Costs & Response: CPS

## CPS Cost Per Case and Response Rate



# Challenge & Opportunity

## Challenges

- Declining response rates translates to higher costs and impacts to data quality
- Respondents less willing to engage with traditional CAPI and CATI survey designs
- Produce more timely and relevant data for the same money, without adding respondent burden
- Mitigating 21<sup>st</sup> century disclosure risks
- Modernizing without impacting data series



## Opportunities

- New Office of Survey & Census Analytics strengthening the Bureau's ability to produce actionable paradata analysis
- Expanding use of adaptive design techniques
- Exploring the potential of administrative data to enhance or replace survey data
- Collaborating with survey sponsors on design research and testing:
  - Effective Sampling
  - Questionnaire Restructuring
  - Mode Experimentation
  - Incentive Strategies
  - Respondent Communication
  - Disclosure

# Office of Survey and Census Analytics (OSCA)

- Established this year in response to Bureau needs for:
  - Data-driven production tactics that could be implemented in real-time to manage rising costs and declining response
  - Support to Regional Office staff working to manage and make sense of multiple sources of paradata
  - Fully leverage the Bureau's new Unified Tracking System (UTS)
- Operates within the Field Directorate

# Metric Improvement

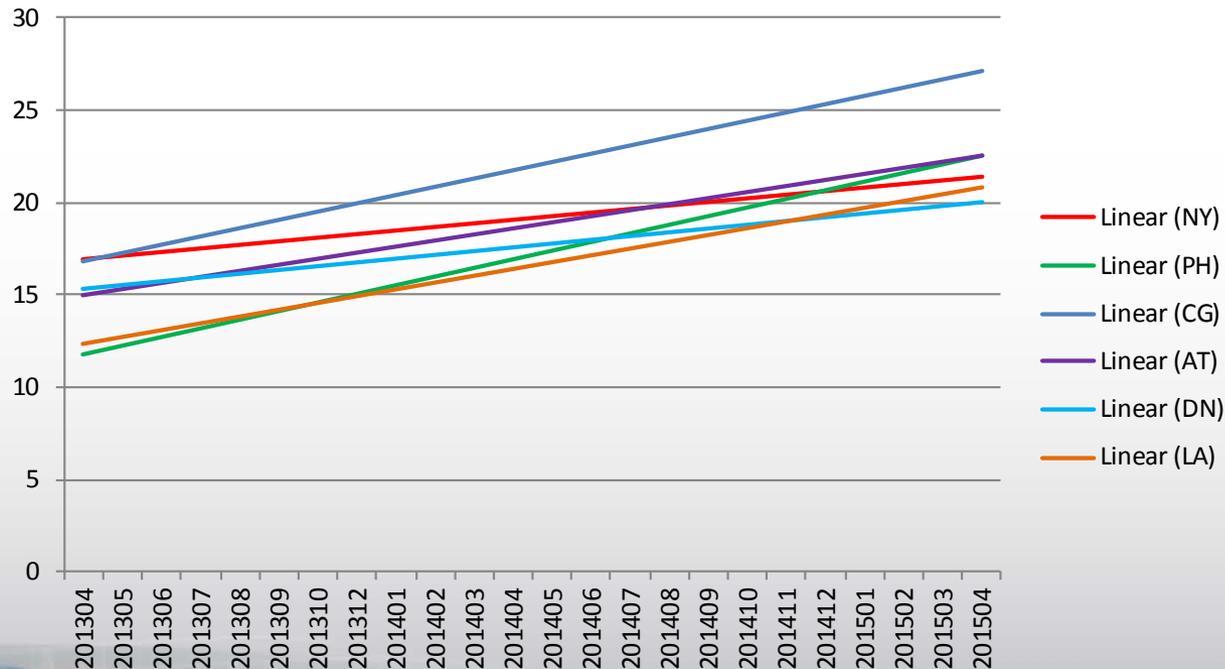
## Sample Report: Interviewing Progress, Hours, and Miles

Filtered View for FS 7181 - Carol (A79)					CAPI Workload		Hours		Miles	
SSF	FS	FR	Survey	Period	Current	Progress	Charged	Hours per Case	Charged	Miles per Case
71	7181	Abbey (T36)	ACP	201504	24	54%	14	0.58	84	3
71	7181	Abbey (T36)	CPS	201513	17	88%	11	0.64	66	4
71	7181	Beth (G57)	ACP	201504	20	70%	10	0.50	60	3
71	7181	Beth (G57)	CPS	201513	4	100%	2	0.50	12	3
71	7181	Beth (G57)	CPS	201534	2	0%	4	2.00	24	12
71	7181	Beth (G57)	CQR	201504	4	0%	0	0.00	20	5
71	7181	Cathy (J03)	ACP	201504	19	37%	8	0.42	48	3
71	7181	Cathy (J03)	SPR	201502	25	60%	10	0.40	60	2
71	7181	Daria (J02)	CPS	201513	20	40%	4	0.20	24	1
71	7181	Daria (J02)	CPS	201534	1	0%	0	0.00	0	0
71	7181	Ed (G45)	ACP	201504	12	50%	8	0.67	48	4
71	7181	Frank (G30)	CPS	201513	9	100%	5	0.56	30	3
71	7181	Frank (G30)	CQR	201504	6	0%	2	0.33	12	2
71	7181	Frank (G30)	HIS	2015204	14	50%	7	0.50	42	3

# Diagnostic Work

Varied analyses to identify actionable causes for response rate declines

Average Caseload of FRs on Survey with < 1 Year Census Interviewing Experience

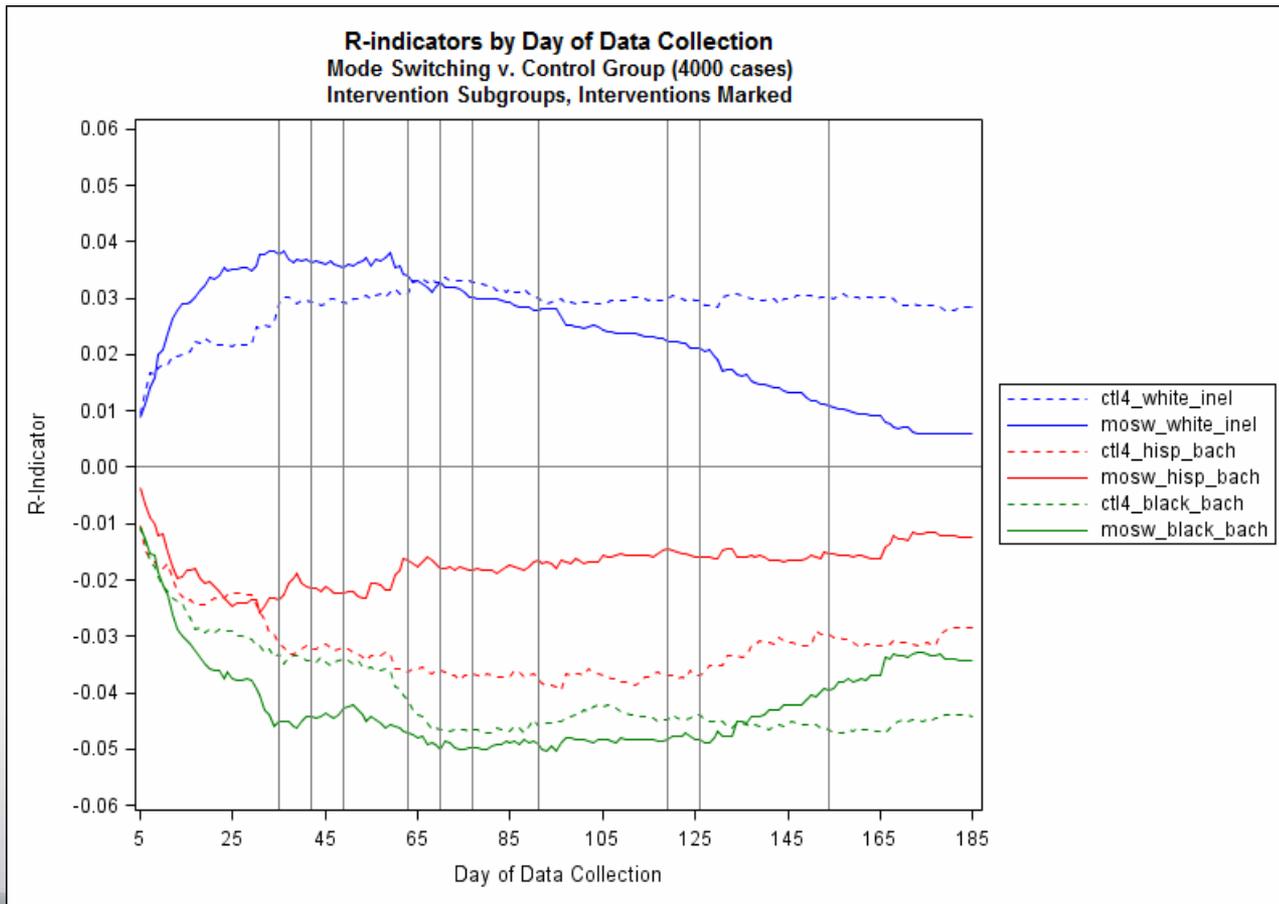


Source: ADRM Paradata  
April 2013 – April 2015

# Adaptive Design

- Adaptive design techniques used successfully on the 2013 National Survey of College Graduates
  - Focus on producing the highest quality data as cost-efficiently as possible, analyzing R-indicators and managing interventions in production
  - Reduction in response rate sustained in favor of yielding a more representative data product
  - Significant effort to develop case management and report monitoring capabilities

# Intervening in Data Collection



- All treatment groups had significant changes in trends after interventions
- Black highest degree B.A. improved after Full CATI

# Current Adaptive Design Work

- Working with the Center for Adaptive Design to build “universal” R-indicators across suite of surveys
- Census Bureau understanding of sponsors’ core research objectives and the needs of their data user communities is critical

# Administrative Data

Sponsor	Survey	Inquiry
Census	SIPP	Potential for enhancing, improving and/or replacing income data
HUD	AHS	Disclosure research to ensure continued confidentiality of PUFs
NCHS	N(H)AMCS	Use of Electronic Health Records (HER) to enhance/replace record abstraction
NCHS	NHIS	"Year Built" Designation to scope out ineligible housing units
MCHB	NSCH	Exploring ADREC use to enrich sample for households with children

# Survey Redesign Challenges

- Census and Sponsors acknowledge that survey redesign is needed, but challenged by:
  - Availability of funding and staff resources to conduct research and testing activities, and fully engage stakeholders
  - Balancing redesign activities with today's production needs
  - Concerns about maintaining data continuity / explaining impacts of methodological changes on data
  - Anticipating rapid advances in methods and technology over the course of a multi-year redesign effort

# Survey Redesign

Status	Survey(s)
Redesign in full production	SIPP, NHES, NSCG, NTPS (formerly SASS)
Design or design elements in field test	CE, NSCH
Research Planning Underway	NAMCS, NCVS, NHIS, TPOPS

# Outline: A Case Study - SIPP

- SIPP: A brief description
- Re-engineering
- Innovations
  - Event History Calendar with Dependent Data
  - Topic Model Imputation
  - Monitoring tools for quality and cost
- Progress and Milestones

# Survey of Income and Program Participation

- Nationally representative, longitudinal, multi-stage stratified sample
- Continuous data in 3-4 year panels from the 1980s through present
- Sample: Civilian, non-institutionalized U.S. households
- **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”*

# Key Design Changes and Benefits

- Annual interview
- 12-month reference period from 4-month
- Event History Calendar (EHC) methods - Facilitates respondent recall over longer reference period
- Reduced cost through annual administration

## Scope

- Similar to SIPP
- Broader than core / includes key topical module content in each wave

## Better integration of concepts

- EHC - integrates reporting across domains – incorporates dependent data
- Topics previously implemented as add-on modules now integrated

## Increased efficiency in processing and producing data products

Flexibility in administration (dynamic interview month and reference period)



# Innovations

- Focused use of dependent data - See informational attachment:
  - “Implementation of Dependent Interviewing in the SIPP Event-History-Calendar: Clear Benefit, Room for Improvement, Future Directions” presented at Panel Survey Methods Workshop in Ann Arbor, 2014
- Type-Z model-based imputation
  - informed by administrative records
  - operationalizing methods discussed in the early 1990's - sequential regression multiple imputation
- Monitoring
  - Integration of paradata streams for management and evaluation
  - Intensive interviewer training – many aspects to monitor
  - CARI – Audio Recorded Interviews

# Topic Model Imputation

## Problem:

- How to improve process for creating fully imputed data where whole people are missing from the household?
  - Previously relied on matching to donors and substituting prior to edits.
  - How to implement new imputation methods and still release data in a timely manner for a survey with 11,000 collected and 2,000 released variables?

## Solution

- Replace item-level hot deck with parametric model-based approach
  - Helps handle small hot deck cell size problems
  - Allows inclusion of many more predictor variable
  - SIPP SSB provides the methodological foundation for modelling
  - Use administrative data to mitigate problems caused when survey data are not “missing at random”
- Use topic flags as alternative to whole-record donation for cases where respondent did not complete the whole sections of the survey.
- Indicator variables for all the major topics covered by SIPP (See Ref. Sect. 1)
- Implement new methods only for these 40+ variables

# Results

## Overall Percentages for cases where SIPP respondent answered the first question about jobs held (94.5% of in-universe respondents)

Worked for pay in 2013?		W-2/Schedule C positive earnings in 2012?	
Yes	58.2	Yes	58.1
No	41.8	No	41.9

## Overall Percentages for cases where SIPP respondent DID NOT answer the first question about jobs held and TF was imputed (5.5% of in-universe respondents)

Worked for pay in 2013?		W-2/Schedule C positive earnings in 2012?	
Yes	61.5	Yes	60.4
No	38.5	No	39.6

# Topic Model

## Conclusions:

- Model-based imputation is feasible in a production environment for a large-scale survey
- Outside data sources (especially administrative data) are valuable:
  - Additional predictor variables in a model
  - Independent of survey non-response mechanism

## Next steps:

- Model respondent-reported earnings
- Model beginning and end of spells
  - Help mitigate seam bias
- Model more topics
  - Defined benefit pension contributions
- How to best take account of spouse/parent/sibling relationships in the data when modeling

# Monitoring for Quality and Cost

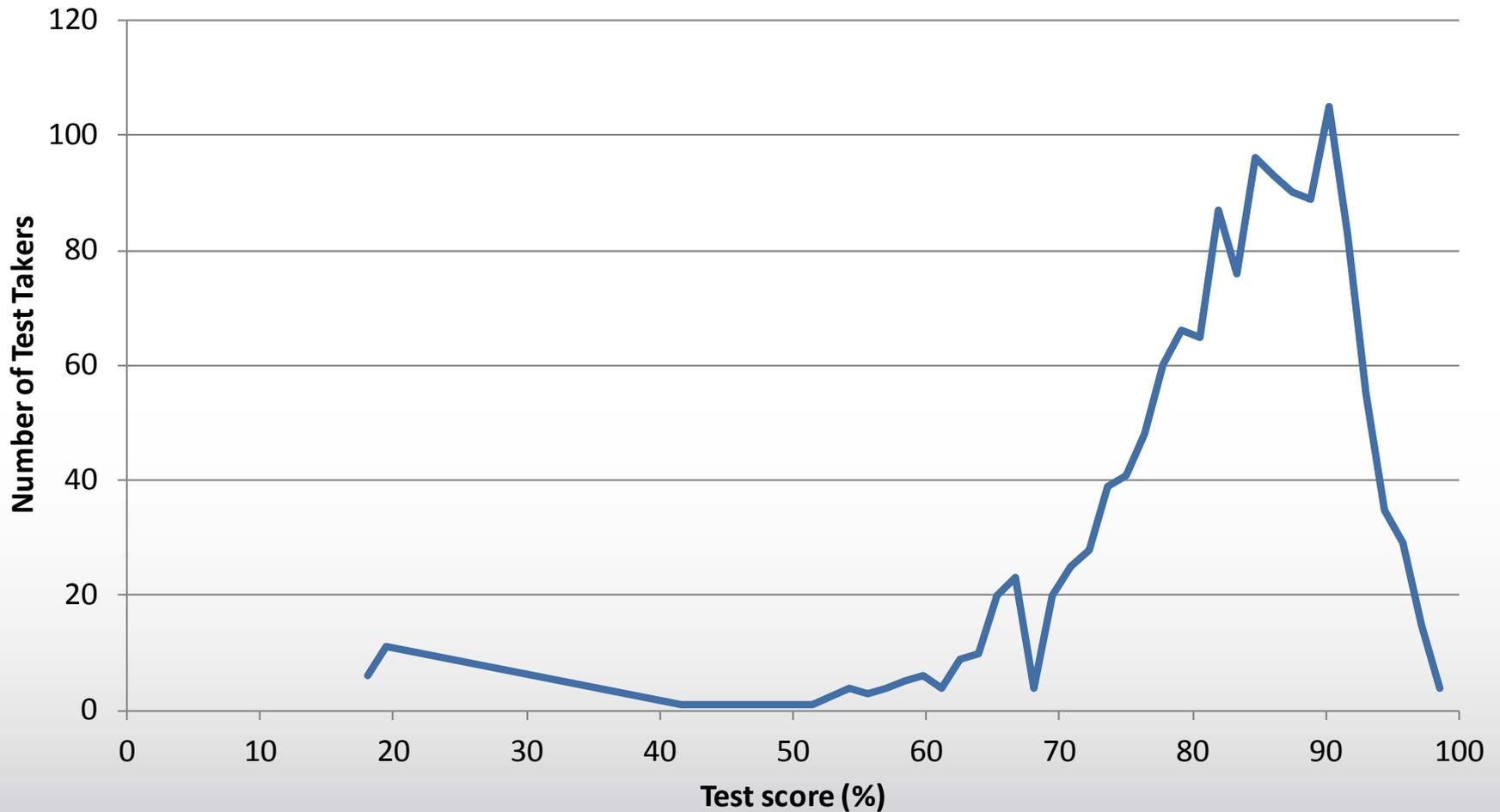
## - Interviewers (Lots of new hires)

	Wave 1 Staffing	Wave 2 Staffing
<b>Interviewing period</b>	February 1 – June 9, 2014	February 1 – May 31, 2015
<b>Hiring period</b>	Fall/winter 2013 (*significantly delayed by federal furlough in October 2013)	Fall/winter 2014
<b>Training period</b>	December 2013 – April 2014	December 2014 – March 2015
<b>Field representatives (FRs)</b>	1,198	1,140
<b>New hire field representatives</b>	423	310
<b>Sample Size</b>	Approx. 53,000 households	Approx. 30,000 households
<b>Average workload</b>	About 40-45 cases per interviewer	About 25-30 cases per interviewer
<b>Interviewing mode</b>	Interviews all started in-person with some telephone completion	Interviews mostly in-person but with some telephone on request
<b>Interviewed households</b>	Approx. 30,000 households	Approx. 23,000 households
<b>Response rate</b>	70.2%	74.2%

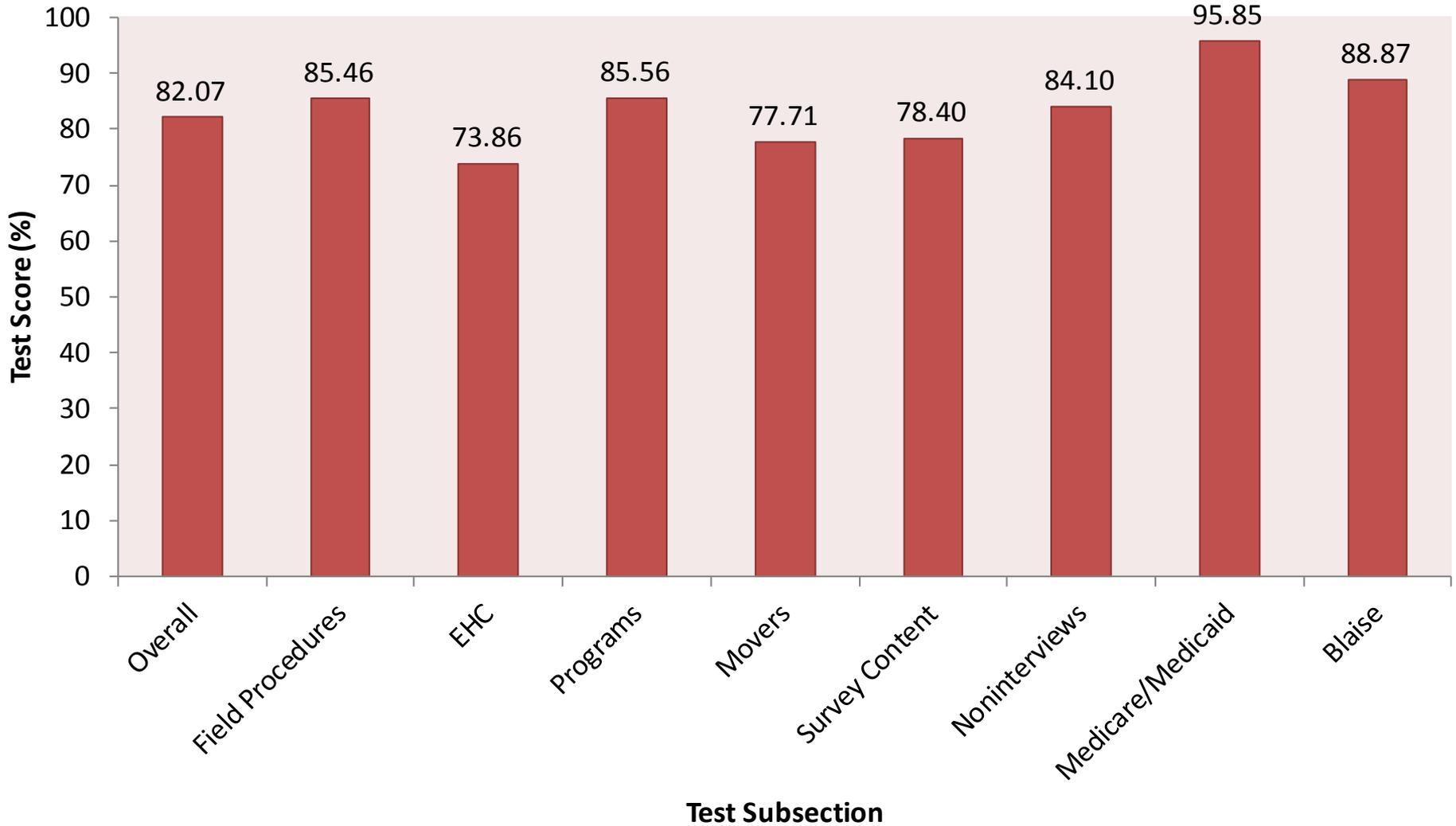
# SIPP 2014 Interviewer Training

- Decentralized training after centralized ‘Train-the-Trainer’ at Census HQ
- Two-day generic Census training
  - New hires only - Covers cross-survey skills
  - Communicating with respondents - Administrative training
- Four-day classroom training
  - All SIPP Interviewers (FRs) - Content specific to SIPP
  - Decentralized verbatim training - Daily quizzes
  - Paired-practices - Computer based training sequences
- Pre- and post-classroom self-study modules
- Ends with certification test
  - Required before fieldwork can be started

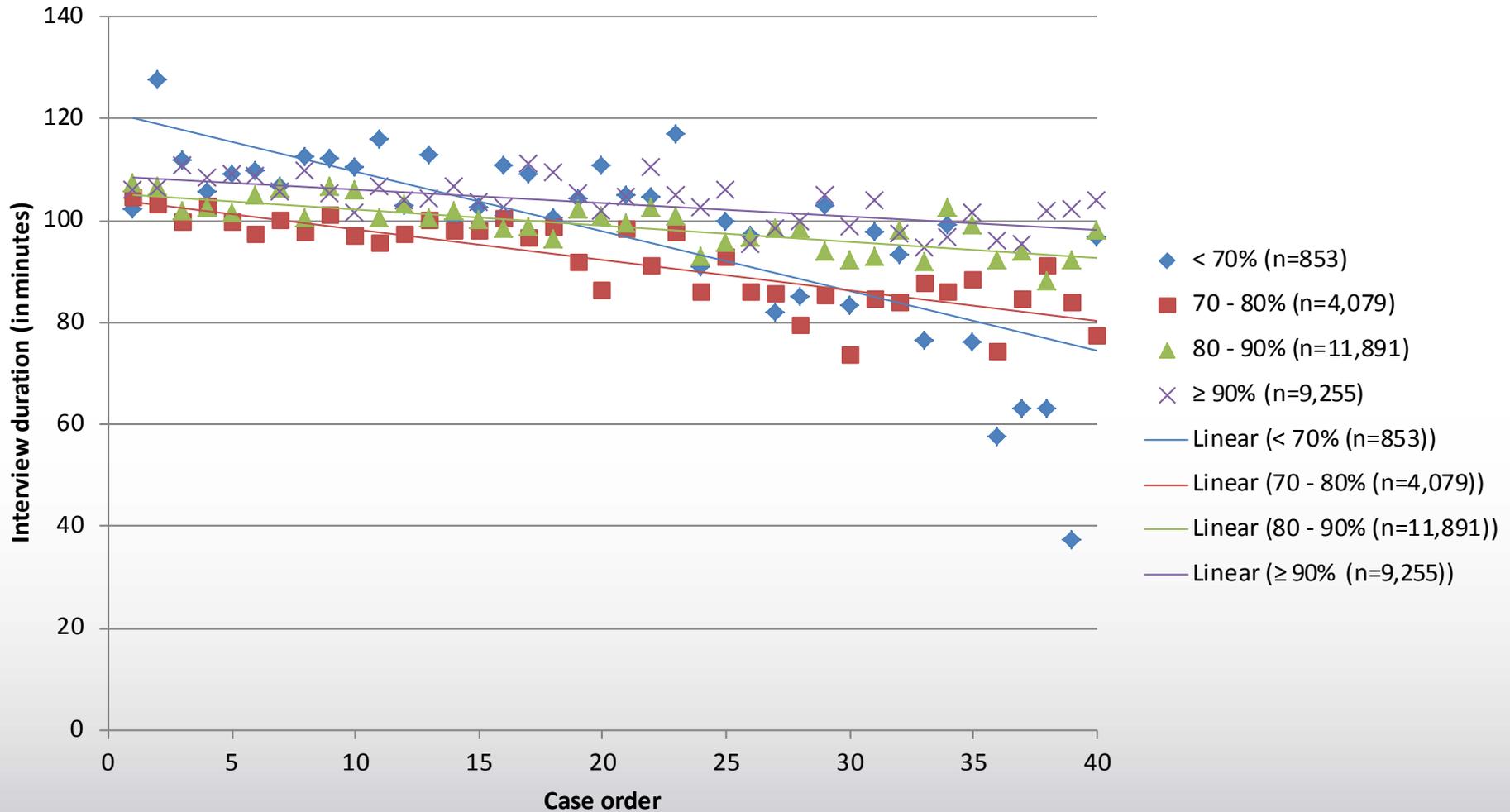
**Figure 1. Number of Test Takers by Certification Test Score, 2014 SIPP Wave 2 (n=1,362)**



**Figure 2. Average Certification Test Score for each Subsection, 2014 SIPP Wave 2  
(n=1,362)**



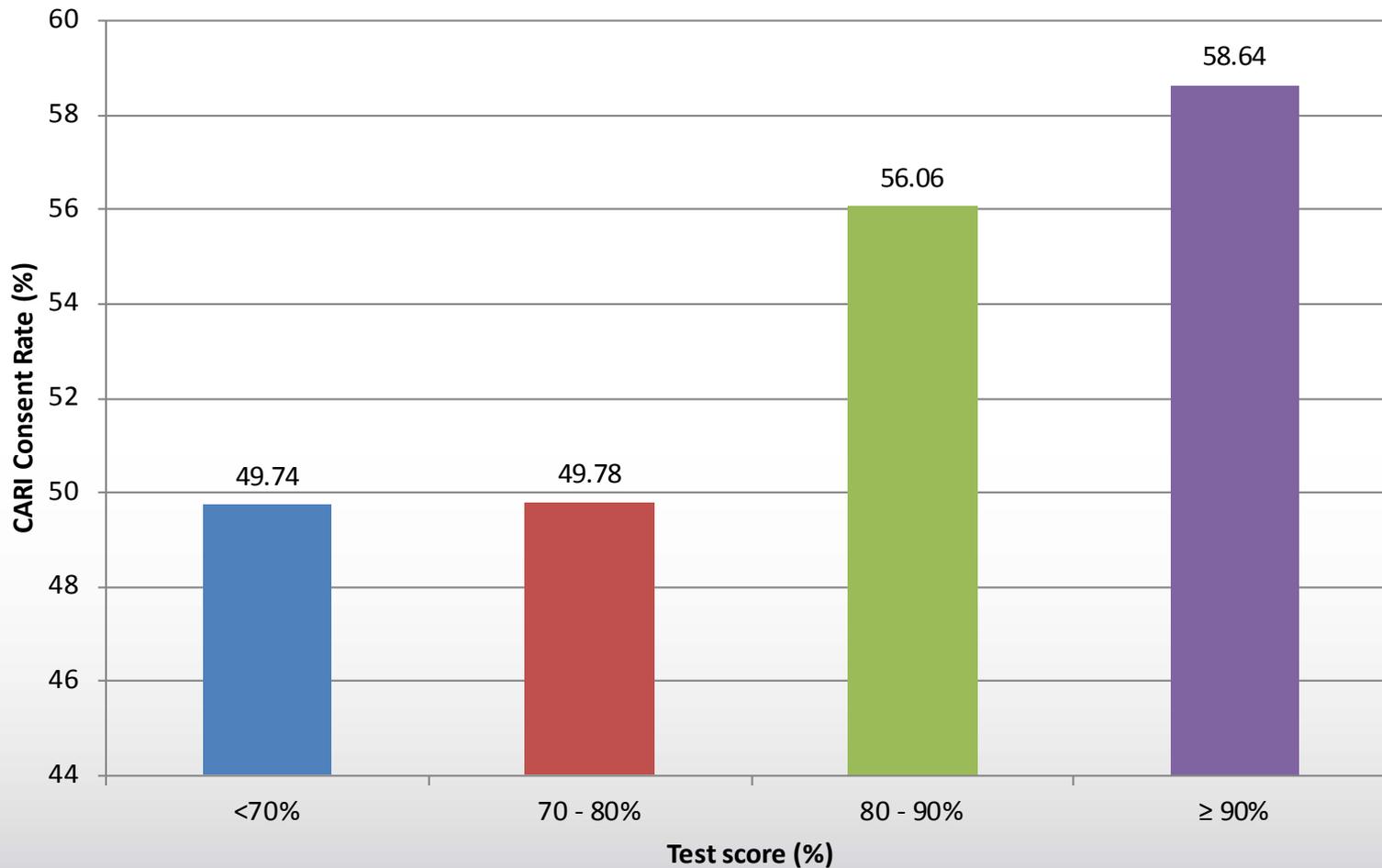
## Figure 5. Average Interview Duration across Caseload (First 40 Cases) by Certification Test Score, 2014 SIPP Wave 1



# Computer-Assisted Recorded Interviewing (CARI)

- FRs must obtain consent from each respondent to record the interview
- Records interactions between Field Representatives (FRs) and respondents
- The goal of CARI is to ensure the accuracy and quality of data collected
  - Improve the FR's performance
  - Identify difficult or problematic questions

**Figure 5. Mean CARI Consent Rate (Persons) by Interviewer Certification Test Score**



# CARI

- Helps ensure a focus on data quality and encourages professionalism
- Listen to recorded cases and code them for:
  - Authenticity (including consent to record)
  - Question administration
  - Behavioral conduct
- Coded Quality Assurance score will directly influence performance rating
- Completely in the control of the interviewer
- May increase non-response and will increase interviewing length

# Reengineering: Some Notes

- Integrated survey instrument that allows:
  - Data storage through a common Blaise data structure
  - Conversational EHC navigation
  - Dependent data incorporated into EHC for Wave 2 and beyond
  - Improved paradata monitoring including -Computer Assisted Recorded Interviewing (CARI)
- Development of a SAS-based data processing system
  - Contains all-new, ground-up edits - Includes model-based imputation
- Responsive and integral stakeholder involvement
- New SIPP and classic SIPP produce estimates that are not substantially different and corresponds with administrative data at least as well
- Transitions fall disproportionately on seams (now Dec-Jan) – continue to develop methods to minimize and adjust.

# Re-engineering Lessons Learned

- Rapid (Agile) development, complexities in design and implementation need:
  - Longer timeframe than expected - instrument change cycles with testing were 6-12 months with moderate changes
  - frequent reviews
  - revisions
  - benefit from prototyping
- Need early and continued stakeholder involvement
- Interview training and monitoring is critical. As is the importance of engaging field staff in the re-engineering process.

# Discussion

- Are there challenges or opportunities that the Census Bureau is not considering in its efforts to produce and deliver high-quality data to our sponsoring agencies?
- What strategies might we use to better understand the desired product output from a sponsor or sponsor stakeholder's point of view, so that we can make better design recommendations?
- Considering the trade-offs between increased effort to achieve response and options to focus on quality during survey management and redesign, how should Census utilize/prioritize resources available?
- Do you have specific questions or suggestions for SIPP?

# THANK YOU!

**Eloise Parker**

**Assistant Director for Demographic Programs – Survey Operations**

[Eloise.K.Parker@census.gov](mailto:Eloise.K.Parker@census.gov)

**Jason Fields**

**Survey Director, Survey of Income and Program Participation /**

**National Survey of Children's Health**

[Jason.M.Fields@census.gov](mailto:Jason.M.Fields@census.gov)

[www.census.gov](http://www.census.gov)

[www.census.gov/sipp](http://www.census.gov/sipp)

# List of Topic Flags in 2014 SIPP

## EHC topics:

- Education Enrollment
- Employment (job lines 1-7)
- General Assistance
- SNAP
- SSI
- TANF
- WIC
- Health insurance
  - Private
  - Medicaid
  - Medicare
  - Military
  - Other

## Non-EHC topics:

- Biological parent (fertility)
- Dependent care
- Disability - adult and child functional limitations (seeing, hearing, etc.)
- Disability (difficulty finding or keeping a job because of disability)
- Disability (not being able to work because of disability)
- Disability payments
- Energy assistance
- Lump sum payments
- Retirement
- Retirement payments
- Life insurance
- School lunch
- School breakfast
- Social Security- Adults
- Social Security- Kids
- Survivor payments
- Unemployment compensation
- Veterans affairs benefits
- Worker's compensation

# Paradata/Auxilliary Sources in Use

- Audit trail data from the Blaise/C# instrument
- Certification test for interviewer training
- Interviewer characteristics
  - Census experience
  - Prior SIPP experience
  - Supervisory status
  - Demographics
- Contact history instrument
- Mileage, case load, supervisor observation, hours billed
- Neighborhood observation
- Regional office progress management application data
- Interviewer debriefing

# Audit Trails

- Audit trail files are a record of all of the keystrokes entered by a field representative (FR) during an interview
- Audit trail files can be used to create paradata on such things as:
  - Section timers,
  - Don't know/refused counts,
  - Help screen calls,
  - Checks encountered,
  - Item-level notes left, and
  - FR navigation throughout the instrument