Moving from Decentralized to Centralized Reinterview

Alpha Savage
Demographic Statistical Methods Division, US Census Bureau*

INTRODUCTION

In May 2014, the U.S. Department of Commerce Office of Inspector General (OIG) issued a mandate to the Census Bureau to create an independent system to evaluate data collected by Field Staff interviewers and mitigate potential bias supervisors may introduce when reporting instances of falsified data by their staff. In response to this mandate, the Census Bureau embarked on implementing a Centralized Computer Assisted Telephone Interview (CATI) data collection process as part of its Quality Assurance (QA) program.

PURPOSE OF STUDY

The goal of the Centralization initiative, rooted in an adaptive design perspective, was to implement an additional data collection mode that allowed survey reinterviews (re-contacts) to be conducted from a centralized location separate from the Regional Offices (ROs) and outside of the interviewers' supervisory chain of command. Three major operational changes were made to facilitate the transition to a centralized reinterview environment, both of which proposed challenges in both scope and complexity:

I. New Technology for conducting personal interviews/CAPI Modernization: Computer Assisted Personal Interviewing (CAPI) is the process used to collect respondent data for a variety of mission critical demographic, economic, and decennial programs. The mobile computers used for conducting interviews prior to 2015 had become obsolete. The need to repair these devices was increasing and each repair was delaying data collection efforts. The CAPI Modernization program deployed in 2015 provided modern computing devices with mobile telecommunication and device management systems integrated into the Census Bureau enterprise architecture. The rationale behind this initiative was to incorporate shared services, using modern technology while reducing the cost to manage and secure the new devices.

II. Automation of mechanism for reporting QA findings: Form 11-163, officially known as The Field Representative Data Falsification Follow-Up and Quality Assurance Form is used to obtain information on the existence and prevalence of data falsification. To better accommodate an automated environment, enhance workflow, and facilitate completion, this tool used to track the quality of survey data collection operations began the transition to an automated system in March 2015.

III. In January 2015, the reinterview instrument was redesigned to trigger an automatic suspected falsification if any of the following three discrepancies were identified: 1) Household not contacted per reinterview respondent; 2) Classified Interview Type A unit as Type B/C; 3)Field Representative did not use laptop

Using the before and after centralization QA sample results from each survey over time, we investigate the impact of different data collection modes and operational changes on the following key reinterview measures:
a. Reinterview Completion Rate   b. Suspected Falsification Rate (Case Level)   c. Lag Time

PROFILE OF DEMOGRAPHIC SURVEYS

American Housing Survey (AHS): Provides current and continuous series of data on selected housing and demographic characteristics. Sponsored by the Department of Housing and Urban Development (HUD). This survey is conducted biennially in odd-numbered years.
Date of Centralization: July 2017

Consumer Expenditure Diary Interview Survey (CED): Provides data on expenditures, income, and demographic characteristics of consumers in the United States. Sponsored by the Bureau of Labor Statistics (BLS).
Date of Centralization: November 2015

Current Population Survey (CPS): Primary source of labor force statistics for the population of the United States. Sponsored jointly by the Census Bureau and the Bureau of Labor Statistics (BLS).
Date of Centralization: February 2015

National Crime Victimization Survey (NCVS): The nation's primary source of information on criminal victimization. Sponsored by the Bureau of Justice Statistics (BJS).
Date of Centralization: April 2016

National Health Interview Survey (NHIS)**: Principal source of information on the health of the civilian noninstitutionalized population of the United States. Sponsored by the National Center for Health Statistics (NCHS).
Date of Centralization: January 2016

Survey of Income and Program Participation (SIPP): Premier source of information for income and program participation. Collects and measures change for many topics including: economic well-being, family dynamics, education, assets, health insurance, childcare, and food security.
Date of Centralization: April 2016

REINTERVIEW COMPLETION RATE

The reinterview completion rate is the rate at which interviews are verified as correct or incorrect (misclassified and discrepancy). Misclassified and discrepancy cases are considered complete reinterviews. The reinterview completion rate is given by: $\left(\dfrac{\text{\# of complete reinterviews}}{\text{\# of eligible cases in reinterview}}\right)$

AHS Reinterview Completion Rate; CPS Reinterview Completion Rate; CED Reinterview Completion Rate; NCVS Reinterview Completion Rate; NHIS Reinterview Completion Rate; SIPP Reinterview Completion Rate
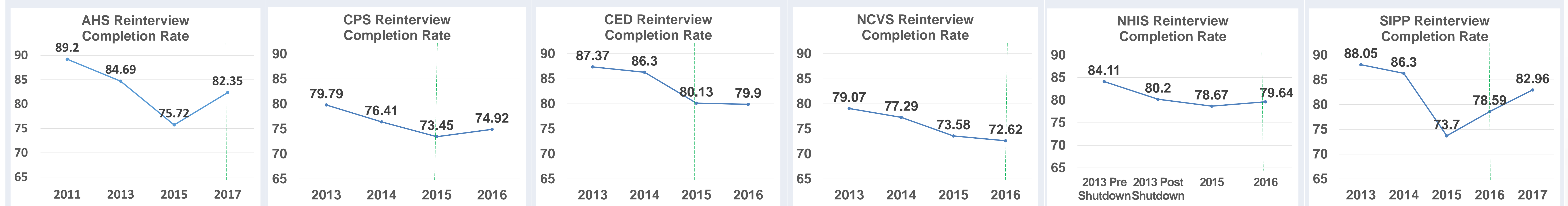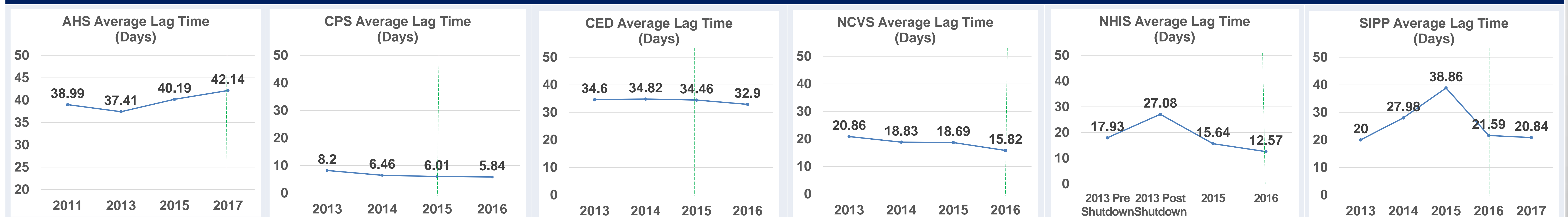
SUSPECTED FALSIFICATION RATE (WEIGHTED)

Falsification occurs when the interviewer makes an intentional departure from interviewing procedures. The suspected falsification rate is the percentage of cases in reinterview that were suspected of falsification as a result of the reinterview. It is given by: $\left(\dfrac{\text{\# of cases suspected of falsification}}{\text{\# of cases in reinterview}}\right)$

AHS Suspected Falsification Rate; CPS Suspected Falsification Rate; CED Suspected Falsification Rate; NCVS Suspected Falsification Rate; NHIS Suspected Falsification Rate; SIPP Suspected Falsification Rate
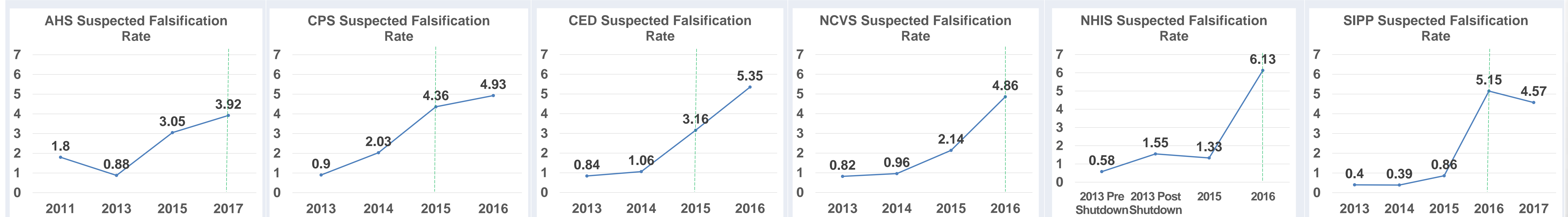
AVERAGE LAG TIME

Lag time is calculated as the time (in days) between the completion of the original interview and the completion of the reinterview. It is important to conduct reinterviews as soon as possible so that respondents can easily recall the original interview.

AHS Average Lag Time (Days); CPS Average Lag Time (Days); CED Average Lag Time (Days); NCVS Average Lag Time (Days); NHIS Average Lag Time (Days); SIPP Average Lag Time (Days)

FINDINGS

- All surveys experienced a significant reduction in reinterview completion rate in 2015 from the prior data collection period. The major reason for this decline was the 2015 implementation of the new technology for conducting personal interviews. The reinterview data collection process was challenged by FR's learning on the fly how to operate the new mobile computers, at the same time as logistics were being put in place for their operation. All surveys except NCVS maintained or experienced an improvement in completion rates in the following year as processes became more familiar and efficient. NCVS has experienced high field staff turnover over the last few years which has led to a struggle to meet reinterview completion goals.

- Excluding NHIS, suspected falsification rates increased significantly in 2015 when cases were automatically suspected of falsification due to any of the "Big 3" discrepancies.

- Excluding AHS, on average, lag times declined for surveys upon moving to centralization as the presence of call centers have increased the likelihood of re-contacting respondents by telephone in a shorter period of time. A significant reduction was identified for NCVS, NHIS, and SIPP. The AHS increase in lag time was not significant in the 2017 cycle and can be attributed to a temporary stop of CATI centralized reinterview operations in September 2017 caused by hurricane Irma.

**Archived NHIS data from 2014 was unable to be retrieved in time to include in this project.

CONCLUSION

- It is important to note that after a case is flagged as suspected, it goes through a supervisory review and investigation process to either confirm falsification, clear the interviewer, or propose additional FR training. Since the migration to centralization the investigation process continues to yield a weighted confirmed falsification rate under 1% on average across the demographic programs.

- Further analyses of current data have revealed that approximately 50% of the QA reinterview workload across surveys gets recycled from CATI to CAPI. Some of the reasons attributed to this recycle rate include:
- Incorrect telephone numbers
- Use of technology to block unwanted phone calls
- Some respondents' preference for personal interaction with interviewer

- Centralized Computer Assisted Telephone Interview addressed the OIG's concerns related to potential bias in field staff investigations, however resulted in early increases in suspected type 1 error rates due to call center interviewer inexperience, and as a result increased costs for field investigations. In an effort to reduce cost based on the existence of type 1 error, there is ongoing research to evaluate the addition of Computer Assisted Web Interviewing (CAWI) into the Census Bureau's QA reinterview program.

- Continued additional data monitoring will reveal further findings on the impact of the transition to a centralized reinterview environment on key reinterview measures.

U S C E N S U S B U R E A U
Helping You Make Informed Decisions