A Data Quality Program for the Survey of Consumer Finances

Catherine Haggerty, Micah Sjoblom and Steven Pedlow
NORC, University of Chicago, 55 East Monroe, Suite 2000, Chicago, IL 60603

Abstract

NORC engaged in data quality improvement activities throughout the most recent ten month data collection period of the Survey of Consumer Finances. These activities included weekly data review and ongoing instruction to interviewers. The data review activities included reviewing data monitoring, data evaluation and reviewing three potential sources of explanation for data problems: call record entries, reviewing commentary recorded during the interview, and debriefing notes recorded after the interview. The ongoing instruction to interviewers was delivered in a variety of forms: weekly memos, targeted interactive self-directed lessons delivered in electronic format, ad hoc group meetings held via-teleconference; and during weekly one-on-one meetings with a supervisor. We will describe the processes we used to identify data quality issues, our data quality improvement protocols and tools, and the data quality measures and data we examined over time.

Introduction

This summary of our presentation at the Joint Statistical Meetings (JSM) in August 2014 is a follow-up to the presentation we made at the 2013 JSM Conference. Last year we presented data based on just the first 12 weeks of the 2013 Survey of Consumer Finances data collection. In this paper we present data collected and evaluated over the entire field period.

Background

A fundamental focus of the efforts across all of our survey processes is quality and continuous improvement. For detailed information about the variety of ways in which we focus on quality see Haggerty et.al. 2013. The following list includes the activities we described in our presentation and which summarizes a subset of the activities we engage in to ensure high data quality:

- Validation

- Monitoring

- Evaluation

- Feedback

- Remedial Training/Continuous Learning

Validation: At least ten percent of each interviewer's completed interviews are re-contacted to verify that the interview was conducted with the respondent. Several factual questions are asked, such as the address where the interview was completed, whether or not the interview was completed by telephone or in-person, the kinds of questions the respondent was asked and the approximate length of the interview. We also confirm that an incentive was paid to those respondents flagged to receive an incentive. The answers to these questions help us to be sure the data were actually collected from the right household and person within the household.

Data Quality Monitoring: We use multiple data monitoring techniques to monitor data quality. These techniques include: 1) applying Benford's Law to our data, 2) counting missing data in particular questions, 3) measuring the length of time the interview takes to complete, 4) looking for a dollar value of "1", and 5) counting character key strokes in selected questions.

All questionnaire data is compared to Benford's Law, also called the First-Digit Law, which checks the frequency distribution of digits in many real-life sources of data. This distribution expects the number "1" to occur as the leading digit about 30% of the time, while larger numbers occur in that position less frequently: e.g. 9 occurs as the first digit less than 5% of the time. We look for interviewers whose work does not conform to the expected distribution in Benford's Law and very closely examine those cases so as to identify potential data falsification as early as possible.

Respondents are often off-put by the exact detail we require about their household finances. Interviewers must constantly encourage and persuade respondents to provide answers to these questions; cases with too many missing values are flagged for review and often do not contain enough data to be acceptable for FRB analysis. The weekly examination of these data help us to identify interviewers who need coaching to ensure they hone their persuasiveness skills so that the many highly sensitive questions which require the capture of dollar amounts actually have valid data.

Interviews taking fewer than 35 minutes to complete are considered executed too fast to allow the collection of high quality data; these cases are closely examined, always validated, and sometimes deemed unacceptable by the FRB.

Interviewers sometimes incorrectly record a response or a respondent misunderstands a question and gives an incorrect answer with leads interviewers down a path in the instrument that is not applicable to the respondent. When this happens we typically see dollar value answers of "1" as interviewers put in an answer until they get out of a section of the questionnaire they should not be in. When we find this occurring we review with the interviewer the proper procedure to follow when they discover they are in a section that does not fit the respondent's financial situation.

The debriefing questions require the interviewer to clarify extraordinary responses, provide a synopsis of the family's finances, and to explain or clarify unusual family situations. If the keystroke count is low we suspect the explanation is inadequate and the case is flagged for a more thorough review. We also identify interviewers who need coaching to ensure they improve their debriefing notes.

Data Evaluation. Traditionally, the FRB have exclusively reviewed all questionnaire data; in 2004 the FRB began to provide case level feedback to interviewers. While FRB economists continue to engage in a full review of the questionnaire data and all commentary associated with each case, the 2013 SCF marked the first triennial in which NORC engaged in a substantive review of a subset of the data interviewers collected. SAS code programmed by the FRB checked for 36 potential problems associated with 15 questions and triggered a subset of the data collected each week to be reviewed by survey staff. Examples of the kinds of problems are: interviewing the wrong household member, unknown homeowner status, incorrect dollar value(s), and households with zero income.

The notes contained in the call records, interviewer commentary recorded inside the instrument and debriefing documentation were reviewed for explanations of the triggered flags. Often there are valid reasons for the unexpected data recorded but when an explanation does not exist we ask the interviewer if they recall any details regarding the unexpected responses. When we are unable to resolve the potential issues prior to data delivery, we notify the FRB to prioritize the review of cases with unexplained potential problems. Of the 6,034 cases completed, NORC staff reviewed substantive data associated with 2,511 cases and the narrative data written in the call records, interviewer commentary in the questionnaire and in the debriefing section for all 6,034 cases.

Observations from Ongoing Data Monitoring and Evaluation Activities

Data Monitoring: Benford's Law

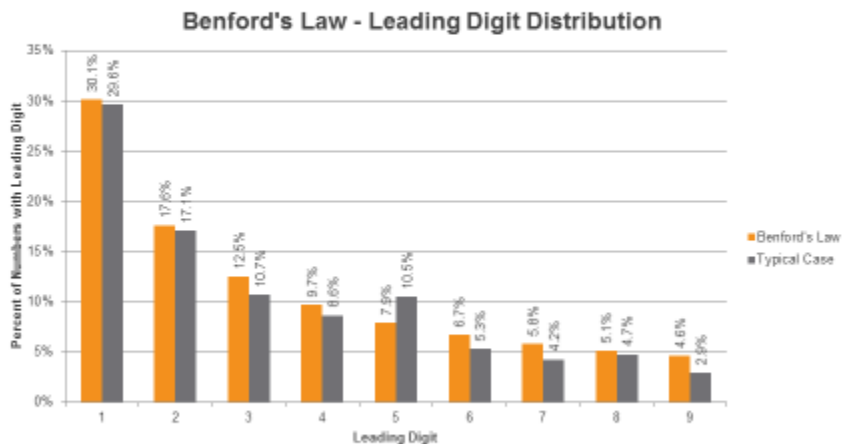The following graph depicts the leading digit distribution typically found in nature.
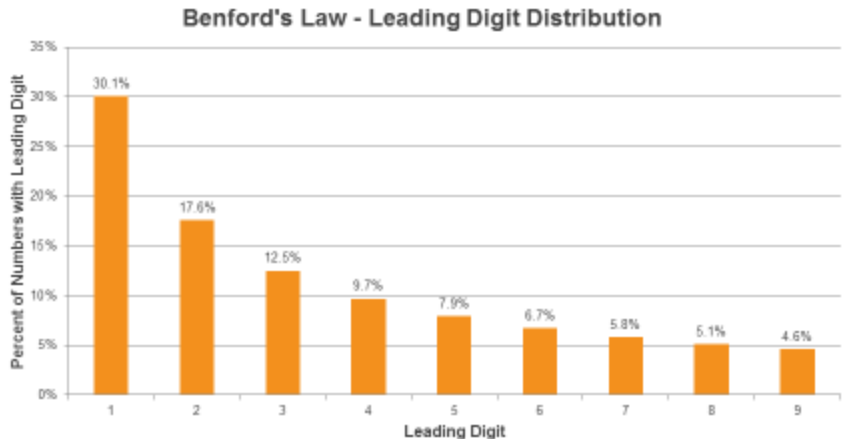
The following is an SCF case with a first digit distribution that is typical of most cases. Note that the digit five appears to be out of sync with what occurs naturally. This is due to the rounding that people often do when presenting numbers.

The following is an SCF case with a first digit distribution that is not typical of most cases; this is a case that we would closely examine to look for potential data falsification.
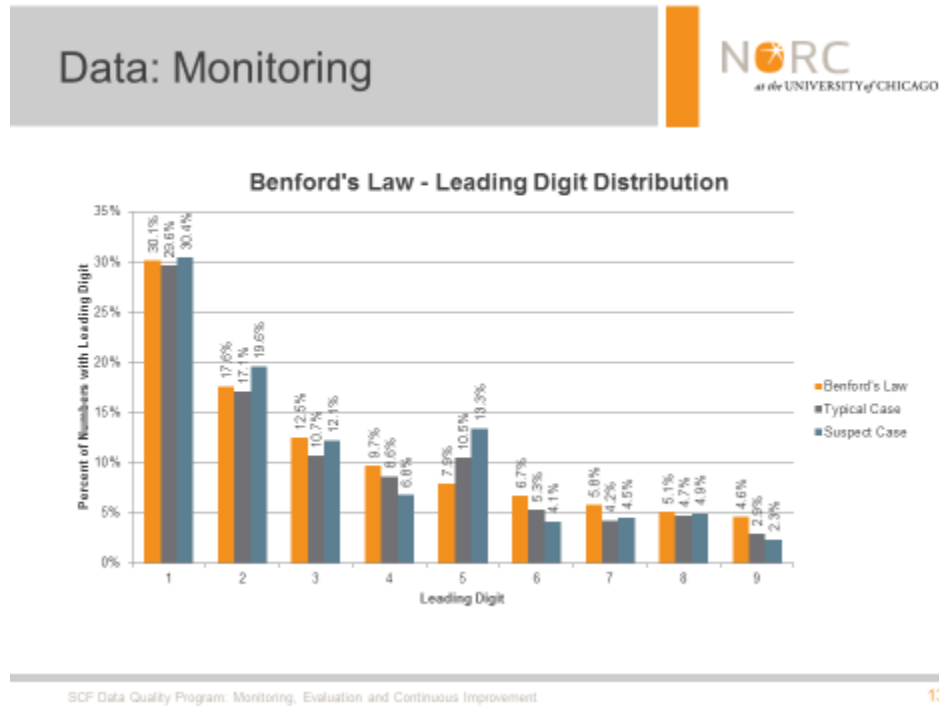
Data Monitoring: "Don't know" and "Refuse" Responses

When we look at the distribution of interviewers with more than 10 percent ("too many") don't know and refuse responses we see a spike in October. The spike is coincident with an increase in the amount of respondent incentives offered, therefore, we may have interviewed some respondents who were uncomfortable disclosing some of their information. After October we returned to the level we were at previously. We note that even as the cases were getting progressively more difficult interviewers were keeping missing data largely constant.
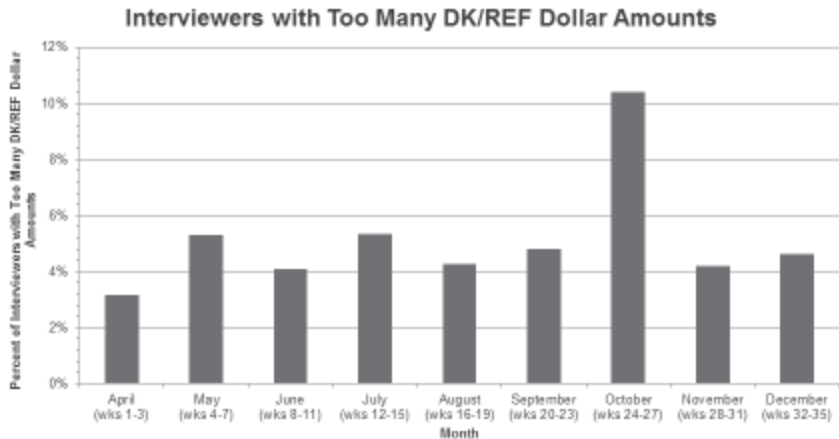
Data Monitoring: Questionnaire Length

It is difficult to complete an SCF interview in fewer than 35 minutes. We had 41 of these and we closely examined these cases. As it turned out, all 41 were accepted by the FRB as acceptable completed interviews. Note: 73 cases have an elapsed time of greater than 5 hours and are not shown on this chart and 41 cases have an elapsed time of fewer than 35 minutes.
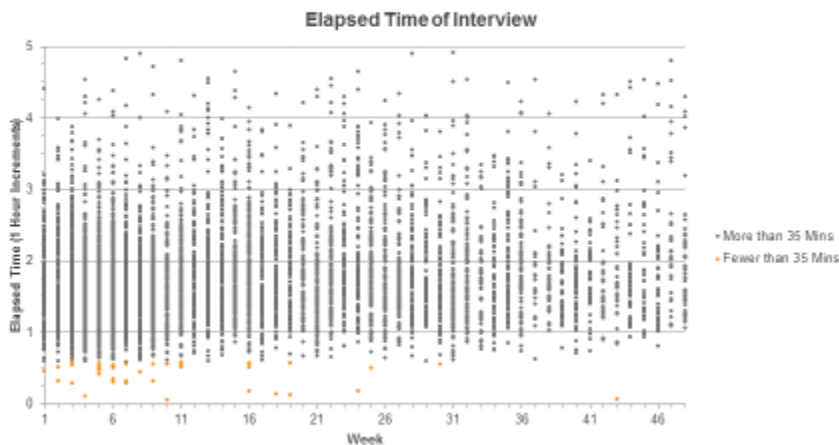
Data Evaluation

In this section we select and present data associated with data problems that occurred throughout the field period. These problems include:

- Dollar value = 1 (one of 15 questions had to have a value of "1" for this flag to be triggered)
- Institution flag is triggered when the interviewer codes the institution as an "other specify" instead of selecting the proper code
- The income flag is triggered if one of three questions indicating income = "0"
- The respondent flags are triggered when there is an indication that we may not have interviewed the right person in the household

The following graph shows these four flags over the data collection period. We see some variability in the distribution of these errors; we took a closer look at these errors to see if there was variation by interviewer type.
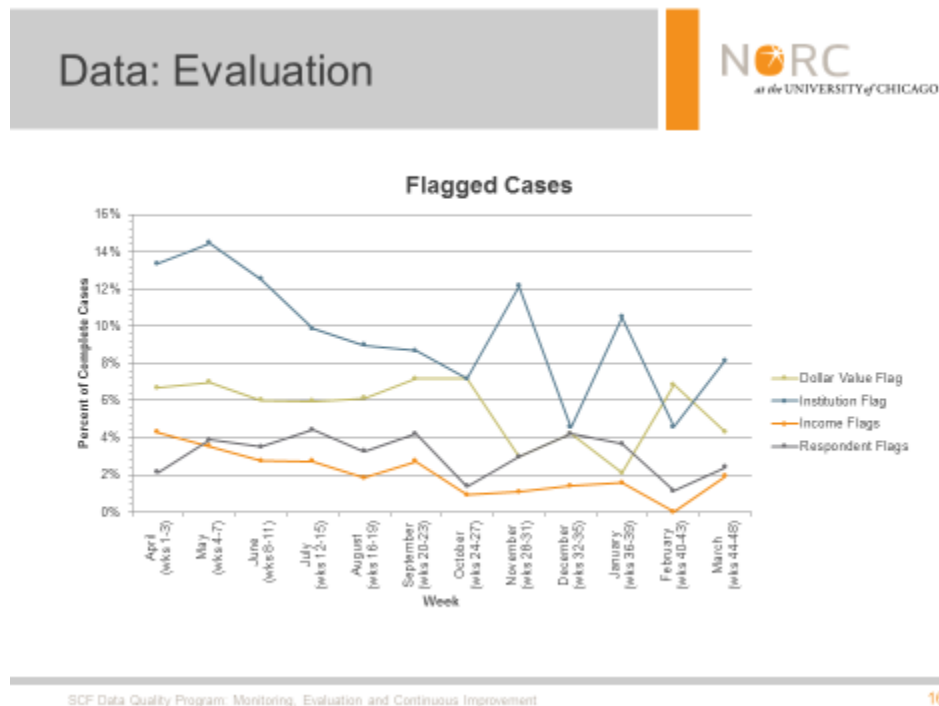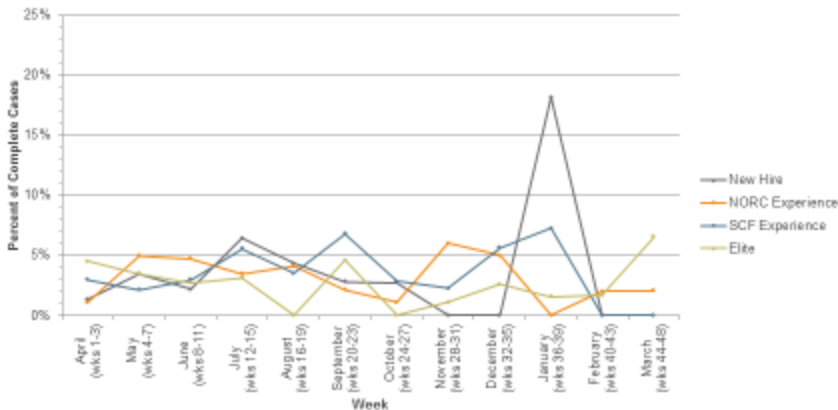

In the following graph we see one spike which is associated with just couple of new hires we still had working at that point in time and that one suspect case caused this spike. What we see here is that even as the cases get harder potential errors in interviewing the wrong respondent are low.
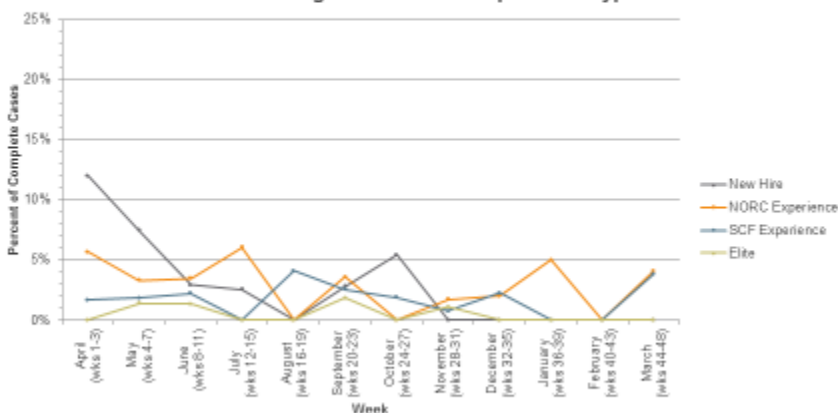
Similarly, in the graph below we see that cases with zero sources of income are few throughout the field period.
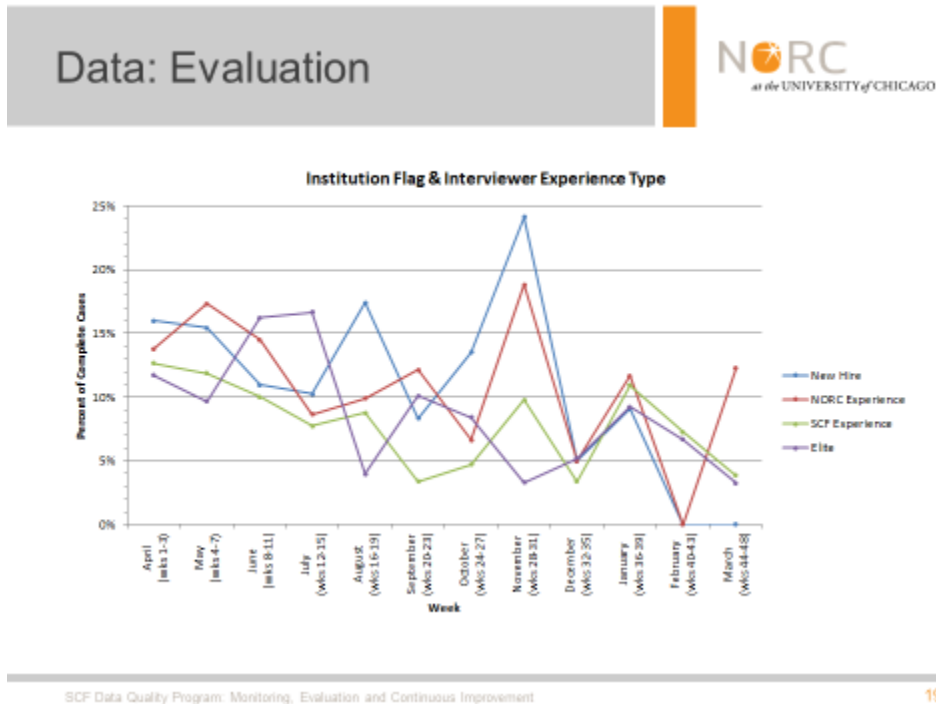
The institution flag is triggered when the interviewer codes "other specify" rather than classifying the type of institution. In the following graph we see a spike in the month we raised incentives. We need to continue to work with interviewers on classifying institutions.


Dollar flag is triggered when the interviewer codes "1" in a dollar field. The following graph shows we still have some work to do to get this error closer to zero.
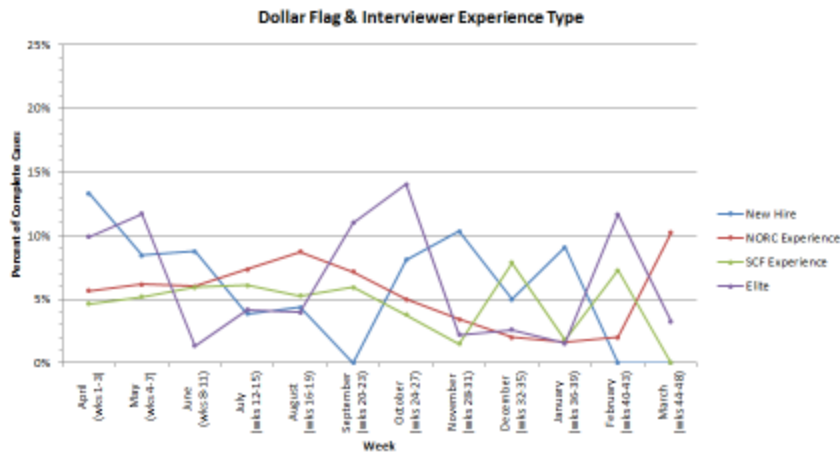
The following three graphs were prepared by Richard Windle (2014) of the Federal Reserve Board. These three graphs show the improvement in data quality across the last three SCF triennial data collection rounds.
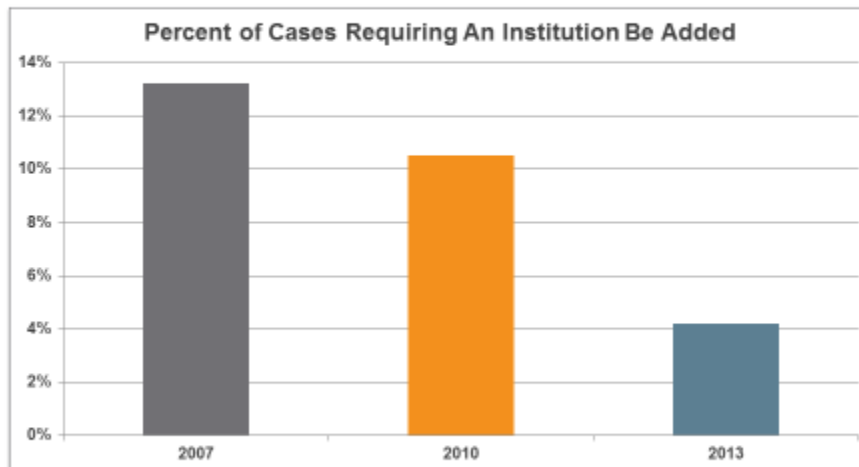
The Federal Reserve Board also reported that over time, there has been an increase in the number of cases requiring no change, a decrease in the cases requiring a edit to decrease income and a decrease in the cases requiring an edit to increase income.
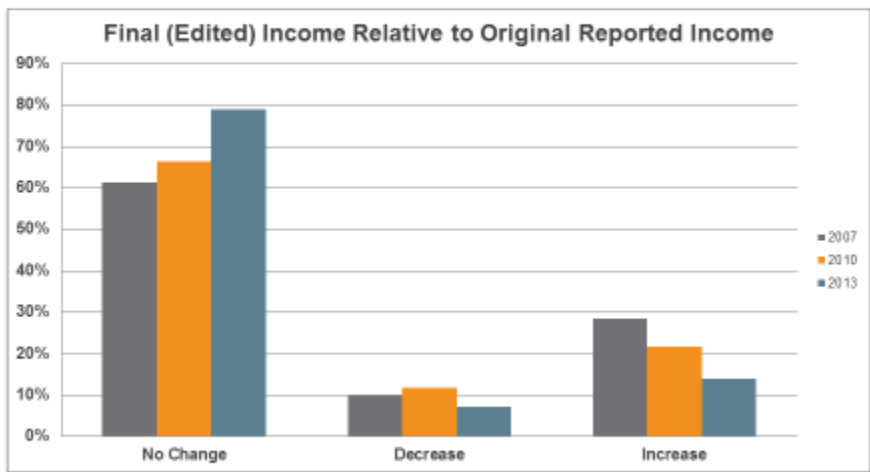
The Federal Reserve Board also reported a decrease over time in the cases requiring edits to the pension section.
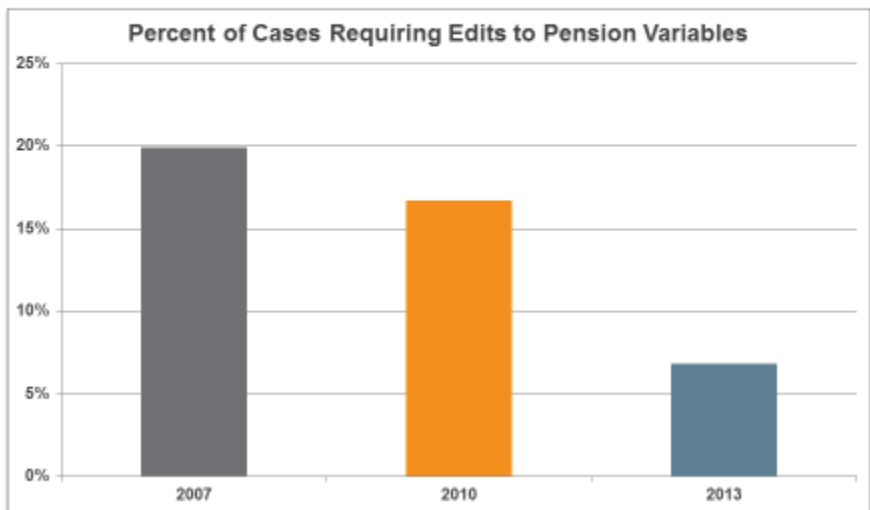
Data Problem Remedies

Activities to improve data quality were varied and custom tailored to the needs of the individual interviewer. The tools we used to improve data quality included the following:

Weekly two page memos which present three to four issues using simple language and explanations of ways to deal with problems encountered while collecting data. The memos used large font, bright colors, icons and cartoon-like characters and short and simple messages to help to ensure the memos were read and the content absorbed. We produced 42 memos.

Moodle, our on-line learning tool, is an open-source online classroom software used by many colleges and universities. Moodle allows for efficient development and delivery of online learning content. Both the aforementioned memos and specially prepared videos were accessed using this tool. Each topic was accompanied by a quiz that interviewers took to be sure they learned the key points of the lesson.

Twenty three videos, one complete mock interview, and all the PowerPoint slides presented at the in-person training were available to interviewers to re-visit throughout the field period. Supervisors asked interviewers to review training videos or slides if their work indicated they needed remedial help. Several dozen, short, self-directed, electronic training videos designed to address the most common data collection errors were produced prior to the start of data collection based on our repeated experience with the SCF. All interviewers were directed to view these videos, one per week, starting at the beginning of data collection. The supervisor follows-up with each interviewer regarding the video lesson during weekly one-on-one meetings. Fifteen additional videos were produced for interviewers earmarked for advanced learning to prepare them to handle collecting data from respondents with complex finances. The videos address known strategies for working with respondents to properly collect complex financial information within the existing CAPI instrument.

Discussion and Recommendations

While our data show that we have been improving both across and within periods of data collection, there is still room for more improvement. The key is to design a data quality program that honors and respects the interviewers. Measure, evaluate, and have a means to improve data quality, and do so in a non-threatening and constructive manner. A non-punitive focus on data problems results in interviewing staff eager for feedback about their interviews and staff earnestly working to correct errors (Haggerty and Kennickell, 2012).

This paper presented the ways in which we examined the data throughout the field period and the corrective action in the form of continuous learning designed to reduce errors. We encourage all surveys to carry forward what is learned about patterns of error and to seek effective preventive and remedial strategies so that future surveys are improved upon.

Bibliography

Haggerty, C. and A. Kennickell. 2012. The Survey of Consumer Finances: Creating a Consistent Culture of Quality. In *JSM Proceedings*, Techniques and Methodologies Section. Alexandria, VA: American Statistical Association.

Haggerty, C. C., M. R. Sjoblom, and S. Pedlow. 2013. Identifying Data Problems and Improving Data Quality in the Survey of Consumer Finances. In *JSM Proceedings,* Survey Research Methods Section, Alexandria, VA: American Statistical Association. 3679-3693.

Windle, R. 2014. 2013 SCF Data Quality Review. Presentation at the SCF Interviewer Debriefing, NORC, Chicago, IL. February.