

What's the Chance? Interviewers' Expectations of Response in the 2010 SCF

Key words: paradata, expectations, nonresponse, interviewers

September 13, 2012

Arthur B. Kennickell
Assistant Director, Division of Research and Statistics
Board of Governors of the Federal Reserve System
Mail Stop 153, Washington, DC 20551
Arthur.Kennickell@frb.gov

Paper prepared for the Joint Statistical Meetings, San Diego, 2012

Abstract

Many household surveys collect and maintain process paradata. Such information has particular appeal, because variations in the level or particular patterns of effort might bear on understanding unit nonresponse or other aspects of survey participation. But for such information to be useful, it is necessary to understand the process that generates them. For surveys that lack a highly structured contact protocol, choices of survey managers or interviewers determine whether and when a record is generated; if that choice is not neutral with respect to characteristics of respondents, then the process data may not be directly usable to study such topics as response propensity, unless there is some means of controlling for choice. For such surveys, one a priori plausible control is the subjective likelihood that a given case can be successfully completed. Effort might be thought to be most likely to be applied to cases that are most likely to be completed. This paper examines data from the 2010 Survey of Consumer Finances (SCF), on the subjective evaluation of the likelihood of case completion that interviewers were required to complete for each effort on each case. In general, the data suggest that interviewers cannot predict the outcome of cases sufficiently reliably or precisely to be useful.

The analysis and conclusions set forth are those of the author and do not indicate concurrence by other members of the research staff or the Board of Governors. The author is grateful to Samuel Ackerman, Robert Argento, Jesse Bricker, Brian Bucks, Gerhard Fries, Kevin Moore, John Sabelhaus, Richard Windle and other colleagues then in the SCF group at the Federal Reserve Board; Susan Boemer, Adrian Dungan, Barry Johnson, David Paris, and Michael Parisi and other staff at the Statistics of Income Division of the I.R.S.; Catherine Haggerty, Micah Sjoblom and other outstanding central office staff at NORC; the SCF interviewers and other dedicated field staff; and the respondents for the survey.

For in-person surveys, the negotiation between a respondent and an interviewer is a game of chance where skill is important. *A priori*, it is generally not known whether any given case will be resolved as complete or incomplete, though there may be general knowledge about what “types” of respondents may tend to be relatively difficult to interview. Aside from the tools that may help to frame the interaction between the respondent and the interviewer (e.g., advance letters, brochures, monetary incentives, etc.), the interviewer must rely on his or her skill in listening to the respondent and providing information tailored to the concerns of questions of the respondent. Through selection over time, successful interviewers must inevitably have skills of this sort. The key question addressed in this paper is whether interviewers have information that allows them to assess the likely outcome of their work while they are in the midst of doing it. This question is of interest for a number of reasons, but the issue that motivated this investigation was the possibility that interviewers’ expectations (and those of their managers) might influence the patterns of work that lead to the ultimate distribution of outcomes and thereby distort straightforward analysis of those patterns.

Many household surveys now collect and maintain “call records” (process paradata), i.e., information on the sequences of efforts devoted toward gaining cooperation with respondents. These records may include date and time of actions, various dimensions of the actions, the outcome, as well as free-text commentary. Such information has great surface appeal, because variations in the level of effort or particular patterns of effort might bear on understanding unit nonresponse or other aspects of survey participation.

To understand the content of such records, however, it is important to consider the process that generates them. In many telephone surveys, contact attempts are driven by an automated calling scheduler, and in other surveys there are at least some elements of a programmatic design for contact attempts (e.g., at least three contact attempts must be made a different times and different days of the week). But in surveys where the programmatic element is more limited, the choices of survey managers or interviewers determine whether and when a record is generated.¹ If the choice to apply effort is not neutral with respect to the characteristics of respondents, then the process data may not be directly usable to study such topics as response propensity, unless there is some means of controlling for choice.

¹ See Biemer *et al.* 2011 for an interesting study of selectivity in the creation of call records.

Simple economic reasoning suggests that field staff who are rewarded for or otherwise encouraged to obtain certain levels of response should want to pursue first the sample cases that are most likely to be completed. If expectations of case completion are a causal factor in pursuing cases, then the process paradata cannot be interpreted without reference to those expectations. To gain insight into the plausibility of such influence, this paper uses data from the 2010 Survey of Consumer Finances (SCF) on subjective evaluations of the likelihood of case completion that interviewers were required to provide after each contact attempt.

The following section of the paper reviews the data used in the analysis. The second section provides various analyses of interviewers' expectations of cases completion. A final section concludes and points to further research.

I. The Data

The 2010 SCF (see Bricker *et al.* [2012]) is the latest in a series of cross-sectional household surveys that focus on wealth and other attributes relevant to the study of household finances. The survey is expected to provide sufficient information to describe both items that are highly concentrated in a small part of the population, as well as items that are more broadly distributed. To that end, the SCF employs a dual-frame sample design, including an area-probability (AP) design to provide broad national representation and a list sample designed to oversample wealthy households. Although the list sample expressly excludes members of the *Forbes* list of the 400 wealthiest people in the U.S., it includes many very wealthy households.²

The survey is detailed in its treatment of income, assets and liabilities. As a result, the median interview required about 75 minutes. Respondents with complicated finances might require several hours and multiple sessions to complete their interview. The subjects covered in the survey are generally considered to be very sensitive. In the 2010 survey, the AP sample included 7,313 in-scope observations, of which 5,012 were completed; the list sample included 4,522 in-scope observations, and across all strata was 1,480 were completed.³

The SCF has a contacting strategy that determines the general outlines of the first phases of work. The strategy was instituted in an attempt both to execute a sufficiently programmed set

² Although the LS oversamples wealthy households, it covers the full set of filers of federal income tax returns, including those who filed only to receive a refundable credit, such as the Earned Income Tax Credit. Because most of the results in this paper that refer to the LS use that sample as a whole, the results can be taken as largely indicative of the situation for relatively wealthy households.

³ In the AP sample, 1,338 cases were out of scope due to a demolition of conversion of a dwelling, a frame error or other reason. In the list, 53 cases were out of scope due to residence in another country, death or other reason. Cases deemed out of scope are excluded from all analyses presented in this paper.

of initial effort that such information could be used to investigate response propensities and to guarantee that all cases received a minimum amount of concentrated effort (see Kennickell [2005]). An initial in-person attempt is required (except in extraordinary circumstances), followed by a relatively small number of additional attempts to gain cooperation. After the required initial attempts have been made, the respondent is sent an informational package via express mail which is considered by field staff to be a very important tool. The package uses a variety of elements to gain the respondent's attention, to address key points of potential objection head on, and to positively motivate participation. Beyond this mailing, the effort devoted to each case is subject in varying degrees to the discretion of the interviewer and his or her manager.

Whenever an action of any sort is undertaken for a sample observation, a call record is required to be created to record the date and time of the action, the nature or outcome of the action, the mode of action (in person, telephone, e-mail, mail, fax, other contact, or no contact), the person contacted (if any), other information about the action, and comments about the action or descriptions of the next steps recommended. In addition at each such step, interviewers are required to assess the likelihood that the case will be completed;⁴ evaluation of this question over the course of interviewers' work is the core focus of the paper.

Not all call records are related to an action that has the potential of directly affecting the respondent. For example, some such records are administrative notes or requests for mailings. Nonetheless, the majority of call records for the SCF represent attempted contacts, which may or may not have resulted in an actual contact. One might choose a variety of indicators of progress or time within such records to use in framing the evolution of the subjective likelihood of getting a completed interview: for example, any sort of attempt, contact of any sort, contact only with a person, contact only with the respondent, or intervals defined in terms of elapsed time since first action on a case. There are reasons one might wish to choose any one of such concepts. In this paper, the unit chosen is contact of any sort, which includes voice contact with a person who may or may not be the respondent, voicemail messages and mailings. Some attempt was made in the analysis underlying this paper to verify the robustness of results to alternative measures of interaction with respondents.

⁴ The question asked is "What is the likelihood of completing this case?" and the response categories are "very unlikely," "somewhat unlikely," "neither unlikely nor likely," "somewhat likely" and "very likely."

II. Interviewers' Expectations

Very substantial effort is devoted in the SCF to locating and contacting respondents and persuading them to engage in an interview (table 1). The median number of contacts a case required to reach a final resolution was 9, and the mean was almost 11. One reflection of the intensity of the effort to work every case to an unambiguous resolution is the fact that the median number of attempts for cases that were ultimately not completed was more than twice that for cases that were completed; the difference was less marked for list sample cases than for AP sample cases, reflecting a higher overall level of effort devoted to the former set of cases. For cases that were ultimately accepted as not completed, the distribution of the number contacts is quite similar for the two sample groups. It should also be noted that substantial numbers of cases were completed with few contacts. For example, 10.4 percent of AP sample cases and 2.5 percent of list sample cases were completed on the first contact (data not shown in the table).

Table 1: Distribution of number of contacts, by sample type and ultimate disposition, 2010 SCF.

<i>Sample type and ultimate disposition</i>	<i>Mean</i>	<i>Percentile of number of contacts</i>						
		<i>P5</i>	<i>P10</i>	<i>P25</i>	<i>P50</i>	<i>P75</i>	<i>P90</i>	<i>P95</i>
All cases	10.9	1	2	4	9	15	22	28
All completed cases	7.9	1	1	3	6	11	18	22
All uncompleted cases	14.5	4	5	8	13	19	26	32
All AP sample cases	9.6	1	1	3	7	13	21	26
All completed AP cases	7.1	1	1	2	5	10	17	21
All uncompleted AP cases	14.0	3	5	8	12	18	26	32
All list sample cases	13.2	2	4	7	12	18	24	30
All completed list sample cases	10.4	1	2	4	8	14	21	26
All uncompleted list sample cases	15.0	5	6	9	13	19	26	32

As noted earlier, after each contact attempt on a case, the interviewer was asked to give an assessment of the likelihood of completing the case. The most striking finding as of the time of the first contact is the large fraction of cases (roughly half) for which the interviewers were equally uncertain whether or not a case would be completed, regardless of whether or not it was actually completed (table 2). If the cases are arranged by the number of contacts at which a final resolution was reached, the data show a general increase in this measure of uncertainty as the point of resolution becomes more distant—as might be expected. Except for cases resolved after more than 20 contacts, there was somewhat more uncertainty about cases that were

ultimately completed than about those that were not. Generally, there was more uncertainty about the outcome for list sample cases than for AP sample cases.

For cases where the interviewer was able to make a more definite subjective evaluation, correct assessments outweighed incorrect ones, but there was nonetheless a surprising amount of error. For example, among AP sample cases resolved after 2–to–5 contacts, about 6 percent of the cases actually completed were seen as very unlikely to be completed, and among those resolved as incomplete about 7 percent were seen as very likely to be completed. Moreover, this pattern is roughly the same across the groups of cases resolved after a larger numbers of contacts. For the list sample cases, there was a smaller fraction of false negatives and a larger proportion of false positives; the pattern is somewhat attenuated for cases resolved after larger numbers of contacts, largely because of the higher proportion of cases for which the interviewer did not make a definite estimate among such cases. Some of the error in assessment may be attributable to recording error, but the simplicity of the question and the means of recording the answers argue that such problems should be minimal among trained interviewers over the course of the field period. It seems more likely that the imprecision and the small change in assessments over the horizon to resolution indicate that interviewers simply have too little information after only one contact to make a reasonable assessment.

Table2: Percent distribution of subjective likelihood of completion as of the first contact, by final outcome, number of contacts to resolution, and sample type.

Likelihood	Number of contacts until resolutions									
	Area-probability sample cases									
	2-5 CONTACTS		6-10 CONTACTS		11-15 CONTACTS		16-20 CONTACTS		>=21 CONTACTS	
	Comp.	Incomp.	Comp.	Incomp.	Comp.	Incomp.	Comp.	Incomp.	Comp.	Incomp.
Very unlikely	5.6	28.6	5.0	22.1	3.8	14.4	2.5	12.1	5.9	6.5
Unlikely	4.3	14.1	5.5	13.7	6.8	15.7	5.4	9.4	5.9	7.7
Neither	51.9	41.6	55.3	45.5	59.9	52.4	67.4	60.1	59.5	63.2
Likely	15.8	8.6	17.5	12.6	14.8	10.0	14.1	11.2	18.0	12.0
Very likely	22.4	7.1	16.8	6.1	14.8	7.6	10.7	7.3	10.8	10.6
% of group	85.6	14.4	60.1	39.9	43.0	57.0	42.2	57.8	34.8	65.2
	List sample cases									
	2-5 CONTACTS		6-10 CONTACTS		11-15 CONTACTS		16-20 CONTACTS		>=21 CONTACTS	
	Comp.	Incomp.	Comp.	Incomp.	Comp.	Incomp.	Comp.	Incomp.	Comp.	Incomp.
Very unlikely	1.7	23.3	3.5	23.9	5.7	19.2	10.5	16.8	4.9	13.2
Unlikely	3.7	8.8	4.2	16.8	7.5	14.3	5.2	17.3	9.0	14.9
Neither	53.8	49.2	43.2	40.7	44.1	47.2	49.3	48.3	51.4	55.3
Likely	7.9	4.2	19.4	4.2	16.3	5.5	10.5	5.1	15.3	8.1
Very likely	33.0	14.5	29.7	14.5	26.4	13.8	24.6	12.5	19.4	8.6
% of group	64.8	35.2	40.5	59.5	28.2	71.8	26.3	73.7	26.0	74.0

Examining expectations at points beyond the first contact might be thought to show the effects of learning, and *a priori* it might seem natural to think that interviewers' judgments would become more precise (table 3). However, for the AP sample cases, the fraction of cases

for which the interviewer could not make a definite assessment increased both for cases that were ultimately completed and those that were not; for the list sample cases, there is a dip in the fraction of cases where the interviewer expressed a neutral expectation for cases resolved between 6–to–10 contacts, but that fraction increases subsequently. Moreover, the pattern of errors for the more definitive expectations shows no clear pattern of decline for either sample group or for either final resolution. One possibility is that systematic learning does take place over contacts, but the set of cases resolved at different points differ in ways that mask that relationship. To detect such change, it is necessary to look at the same group of cases at more than one point.

Table3: Percent distribution of subjective likelihood of completion as of the start of various intervals of contacts required for resolution, by final outcome, number of contacts to resolution, and sample type.

Likelihood	Number of contacts until resolutions									
	Area-probability sample cases									
	2-5 CONTACTS		6-10 CONTACTS		11-15 CONTACTS		16-20 CONTACTS		>=21 CONTACTS	
	Comp.	Incomp.	Comp.	Incomp.	Comp.	Incomp.	Comp.	Incomp.	Comp.	Incomp.
Very unlikely	5.6	28.6	8.2	34.1	6.5	21.9	7.4	18.7	6.8	13.5
Unlikely	4.3	14.1	11.4	15.2	13.0	21.7	13.6	23.3	12.2	19.7
Neither	51.9	41.6	46.5	39.2	50.6	45.9	52.5	46.5	56.3	51.9
Likely	15.8	8.6	18.5	8.4	17.5	7.2	15.7	7.6	15.8	10.6
Very likely	22.4	7.1	15.4	3.2	12.3	3.2	10.7	3.9	9.0	4.3
% of group	85.6	14.4	60.1	39.9	43.0	57.0	42.2	57.8	34.8	65.2
	List sample cases									
	2-5 CONTACTS		6-10 CONTACTS		11-15 CONTACTS		16-20 CONTACTS		>=21 CONTACTS	
	Comp.	Incomp.	Comp.	Incomp.	Comp.	Incomp.	Comp.	Incomp.	Comp.	Incomp.
	Very unlikely	1.7	23.3	2.4	17.7	3.5	10.9	3.7	10.4	1.4
Unlikely	3.7	8.8	3.7	8.9	3.5	8.1	1.5	7.2	3.5	6.4
Neither	53.8	49.2	54.4	52.3	62.1	57.2	68.7	64.5	65.3	69.9
Likely	7.9	4.2	12.2	3.4	6.2	7.4	3.7	4.0	9.7	7.8
Very likely	33.0	14.5	27.3	17.7	24.7	16.4	22.4	13.9	20.1	8.8
% of group	64.8	35.2	40.5	59.5	28.2	71.8	26.3	73.7	26.0	74.0

Table 4: Percent distribution of subjective likelihood of completion at first and fifth contacts or first and tenth contacts, for cases resolved between 6 and 10 contacts or between 11 and 15 contacts respectively, by final outcome and sample group.

Likelihood	Area-probability sample cases							
	Cases resolved in 6–to–10 contacts				Cases resolved in 11–to–15 contacts			
	First contact		Fifth contact		First contact		Tenth contact	
	Comp.	Incomp.	Comp.	Incomp.	Comp.	Incomp.	Comp.	Incomp.
	Very unlikely	5.0	22.0	8.4	33.8	3.7	14.9	6.1
Unlikely	5.6	13.7	11.5	15.4	6.6	15.5	13.0	21.6
Neither	55.5	45.5	46.3	39.4	60.2	52.1	50.9	45.4
Likely	17.3	12.5	18.2	8.2	14.9	10.1	17.6	7.1
Very likely	16.5	6.3	15.7	3.2	14.7	7.4	12.5	3.2
% of group	60.6	39.4	60.6	39.4	43.8	56.2	43.8	56.2
Likelihood	List sample cases							
	Cases resolved in 6–to–10 contacts				Cases resolved in 11–to–15 contacts			
	First contact		Fifth contact		First contact		Tenth contact	
	Comp.	Incomp.	Comp.	Incomp.	Comp.	Incomp.	Comp.	Incomp.
	Very unlikely	2.7	21.7	3.3	24.8	3.5	14.0	6.1
Unlikely	3.5	10.1	4.3	16.8	3.0	8.5	6.9	13.2
Neither	53.1	49.7	43.1	42.0	59.7	55.9	46.8	49.1
Likely	11.7	3.7	18.7	3.9	7.4	6.9	14.3	6.1
Very likely	29.0	14.9	30.6	12.6	26.4	14.7	26.0	13.2
% of group	41.7	58.4	41.7	58.4	28.6	71.5	28.6	71.5

Holding the set of cases fixed to include those resolved in either 6–to-10 contacts or 11–to-15 contacts, the proportion of AP sample and list sample cases for which the interviewer was unable to make a definite assessment declines consistently for both sets of final resolutions, but the amount of the decline is relatively small (table 4). At the same time, the fraction of cases in the two sample groups that were assessed as unlikely or very unlikely to be completed increased—both for cases that were ultimately completed and those that were not. Positive assessment of AP and list sample cases ultimately completed changed little, but the error rate for those not ultimately completed declined.

These results on interviewers’ predictions are based on decisions across all cases. It is possible that the experience is different for individual interviewers. A first cut at looking at this is to examine the distribution across interviewers of outcomes relative to interviewers’ expectations at different numbers of contacts for the same set of cases (table 5). The median percent of cases for which the interviewer did not hazard an appraisal of the likelihood of case completion declined about 10 percentage points from the first to the fifth contact; lower percentiles also declined, but the fraction rose substantially at higher points in the distribution. The percentage of correct appraisals rose modestly at least across the central half of the distribution of correct appraisals, but the distribution of incorrect appraisals shifted more broadly to a higher level.

Table 5: Percentiles of distributions over interviewers of actual versus expected outcome as of first and fifth contacts, AP and list sample cases that required at least 6 contacts to resolve.

<i>Percentile of interviewers</i>	Indifferent		Correct		Incorrect	
	1 st contact	5 th contact	1 st contact	5 th contact	1 st contact	5 th contact
P10	38.7	30.0	13.3	11.5	0.0	4.3
P25	50.0	41.7	20.0	23.1	5.6	7.7
P50	61.0	51.8	27.3	31.6	10.0	13.7
P75	69.2	60.6	35.6	40.0	15.9	22.2
P90	100.0	70.0	91.7	50.0	23.9	27.3

At the level of individual cases, there was also some flux in interviewers’ assessments of the likelihood of completion, but the accuracy of the assessments did not change for most cases (table 6). For about 68 percent of the cases for which the interviewer was not able to hazard an assessment at the time of the first contact, that situation held at the time of the fifth contact; for about 10 percent, the interviewer made a correct assessment at the time of the fifth contact and for about 22 percent there was an incorrect assessment. For the cases with a correct initial assessment, about 60 percent remained so at the time of the fifth contact, about 16 percent became incorrect and about a quarter moved into the middle, uncertain category; the pattern for

cases with an initially incorrect assessment is parallel. Some of the changes in assessments undoubtedly reflect new information, but it would be hard to argue that there is a strong force toward greater accuracy in interviewers' assessments. Extending the horizon to ten contacts does not change the general picture.

Table 6: Change in accuracy of interviewers' assessments of case completion from the first contact to the fifth or tenth contact, cases requiring at least six or eleven contacts respectively to resolve; area-probability and list sample cases; percent of cases in baseline group.

Assessment as of first contact	Assessment as of fifth contact				Assessment as of tenth contact			
	Indifferent	Correct	Incorrect	All	Indifferent	Correct	Incorrect	All
Indifferent	67.5	10.4	22.2	56.8	64.5	11.7	23.8	60.2
Correct	26.5	57.9	15.6	14.9	28.5	51.6	20.0	15.7
Incorrect	26.7	7.6	65.7	28.3	34.1	9.4	56.5	24.1
All	49.8	16.7	33.5	100.0	51.5	17.4	31.1	100.0

The data show considerable variability in this pattern across interviewers (table 7). Nonetheless, the impression of the persistence of the baseline assessment is maintained. Of course, the tails of the distribution (not shown)

Table 7: Variability across interviewers in change in accuracy of assessments of case completion from the first contact to the fifth contact, cases requiring at least six contacts to resolve by interviewers completing more than four cases; area-probability and list sample cases; percent of cases in baseline group.

Assessment as of first contact Percentile of distribution across interviewers	Assessment as of fifth contact		
	Indifferent	Correct	Incorrect
Indifferent			
P25	55.6	10.0	14.3
P50	66.7	14.8	23.5
P75	80.0	22.2	33.3
Correct			
P25	25.0	50.0	14.3
P50	40.0	66.7	20.3
P75	52.2	100.0	34.9
Incorrect			
P25	25.0	7.2	50.0
P50	33.3	14.3	66.5
P75	44.4	20.0	80.0

Using a variety of observed characteristics of sample cases, a number of probit models of aspects of interviewers' assessments were estimated and the results are summarized here. For every observation, information is available from the sample frame about geographic location, characteristics matched to the case for the geographic cases, and the final interviewer of record (possibly not the same as the original interviewer); the interviewer identification was used to create a measure of the number of interviews completed, as an indicator of interviewers' abilities to persuade respondents. Of course, for cases that ultimately participated, there is also the information collected in the main interview as well as paradata about the respondents' degree of engagement with the interview.

For cases that required more than one contact to resolve, at the time of the first contact interviewers were more likely to hazard a view of the final resolution of the case for the cases in the higher strata of the list sample and for cases in Census regions bordering the Atlantic; there was no significant difference by the final resolution of the case. For cases requiring more than five contacts to resolve, some differences remain for list sample cases and some geographic regions. If an indicator of whether the interviewer had hazarded a view at the time of the first contact is included, that variable is highly significant and no other variable has significant explanatory power; this result suggest that variation in prediction over contacts may have a substantial random component.

There is some tendency for interviewers to be less likely to be mistaken in their predictions at the time of the first contact for list sample cases and cases that were actually completed. By the fifth contact, the result for the list sample cases is sustained, but there is no significant effect of the final resolution of the case. Among cases that ultimate participated, interviewers were more likely to err for respondents aged 65 or older, those who expressed suspicion, those who showed a low level of interest and those who seemed to understand the interview least well. Generally, the patterns of error indicate that interviewers do not appear to have an obvious ability to make predictions about case outcomes. Directly modeling the interviewers' predictions yields similar results.

III. Conclusion

This paper presents evidence on interviewers' expectations of the final resolution of cases in the 2010 Survey of Consumer Finances. Interviewers were required to make such an assessment after each contact. This series of data provides a means of examining whether interviewers have reliable "private knowledge" from their interactions with the respondent and that person's surroundings. Such knowledge, if it is sufficiently reliable and precise, might be helpful in managing field work; it knowledge might also be helpful in addressing some potential biases in the analysis of call records, which are generated by a choice-based process that may be affected by perceptions of the likelihood that a case will be completed.

The analysis of the data suggests that any such private knowledge interviewers have is very noisy and not something that has obvious potential to be exploited either for managing cases in the field or for addressing the choice-based nature of the generation of call records. However, it is possible that the underlying question asking interviewers to evaluation the

likelihood of case completion was insufficiently aligned with what interviewers actually know and can report systematically. Moreover, because no effort was made in the interviewer training to highlight the importance of the questions, its importance may have been insufficiently salient. Ongoing work for the 2013 SCF will attempt to address these concerns and provide a further basis for exploration of this topic.

Bibliography

- Biemer, Paul P., Patrick Chen and Kevin Wang [2011] “Errors in the Recorded Number of Call Attempts and the Their Effect on Nonresponse Adjustments Using Callback Models,” mimeo, RTI International, Research Triangle, NC.
- Bricker, Jesse, Arthur B. Kennickell, Kevin B. Moore and John Sabelhaus [2012] “Changes in U.S. Family Finances from 2007 to 2010: Evidence from the Survey of Consumer Finances,” *Federal Reserve Bulletin*, Board of Governors of the Federal Reserve System, Washington, DC.
- Kennickell, Arthur B. [2005] “Darkness Made Visible: Field Management and Nonresponse in the SCF,” *Proceedings of the Section on Survey Research Methods*, 2005 Joint Statistical Meetings, Minneapolis, MN (or www.federalreserve.gov/econresdata/scf/files/asa20055.pdf).