UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001

December 14, 2016

2016 AMERICAN COMMUNITY SURVEY RESEARCH AND EVALUATION REPORT
MEMORANDUM SERIES #ACS16- RER-22

MEMORANDUM FOR   ACS Research and Evaluation Advisory Group

From: Victoria Velkoff
                     Chief, American Community Survey Office

Prepared by: Gregory J. Mills
                     Survey Analytics Measures Branch
                     American Community Survey Office

Subject: Simulated Effects of Changing Calling Parameters and Workload
                     Size on Computer Assisted Telephone Interview Productivity in
                     the American Community Survey

Attached is the final American Community Survey Research and Evaluation report "Simulated Effects of Changing Calling Parameters and Workload Size on Computer Assisted Telephone Interview Productivity in the American Community Survey." The purpose of this paper was to better understand the ACS CATI operation in terms of how calling efficiency varies both across rounds of callback attempts and across the ranked workload. Using six months of CATI paradata, the analysis first shows that the proportion of the CATI workload that receives a completed interview has fallen considerably in recent years. Contrarily, the proportion of cases that are closed out due to never reaching the sample unit have substantially increased. Secondly, we found that cases where a CATI interviewer made contact in earlier rounds of calling had considerably higher completion rates than cases that did not make early contact. Finally, simulations showed that several operational changes could improve calling efficiency, including: 1) lowering the maximum number of callback attempts, 2) decreasing the size of the CATI workload, or 3) by using a combination of both methods. However, based on the data used in the simulations, both reducing the workload size or using a combination of both methods appear to provide a greater efficiency boost than by only lowering the maximum number of callback attempts.

If you have any questions about this report, please contact David Raglin at 301-763-4226 or Gregory Mills at 301-763-6172.

Attachment

Page intentionally left blank.

Simulated Effects of Changing Calling Parameters and Workload Size on Computer Assisted Telephone Interview Productivity in the American Community Survey

FINAL REPORT

Gregory J. Mills, American Community Survey Office

Page intentionally left blank.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

The American Community Survey (ACS) is a multi-modal survey carried out by the U.S. Census Bureau. Nonresponse followup is first attempted with a telephone interview (CATI), and then with a personal interview (CAPI). For several years, the CATI mode has been losing efficiency to the point that by 2016, the cost of a CATI interview was approaching that of a CAPI interview. This purpose of this paper is to research ways to gain efficiency in the CATI mode by either cutting back on callback attempts, restricting the CATI workload to the most likely address-phone number matches, or by a combination of both methods.

The first part of the analysis compared the distribution of CATI outcomes in 2016 to those in 2012, and it is clear that both productivity and efficiency have fallen considerably. The proportion of cases that result in a complete interview has fallen, while all other outcomes have increased, particularly the proportion of cases that are closed out due to the inability to reach the sample unit. The next part of the analysis examined CATI case final dispositions based on early-round callback attempt outcomes. Cases where interviewers make contact with the sampled household, even if that contact is a refusal or immediate hangup, are more likely to close out with a completed interview. Cases that are successfully contacted are also more likely to result in a complete self-response late mail return (LMR), indicating the potential for CATI to push active cases to self-respond.

The final part of the analysis simulated outcomes and outcome distributions based on 1) cutting back callback attempts, 2) restricting the size of the CATI workload, and 3) by utilizing a combination of both of those methods. Based on these simulations, it appears that restricting the workload or the combination of cutting callback attempts and restricting the CATI workload are both better methods to increase efficiency than only cutting callback attempts. CATI efficiency is measured by the number of interviews conducted per hour – the more interviews per hour, the more efficient the operation. We use this rate to measure how efficiently we are obtaining completed CATI interviews in terms of the average amount of interviewer time expended. For example, using a goal of retaining about 80 percent of the interviews, we examined the CATI efficiency achieved by the various methods. By cutting callback attempts only, the 80 percent goal was accomplished at about $0.52 – 0.53$ interviews per hour. By cutting the workload only, the goal was accomplished with a $0.61 – 0.63$ interviews per hour rate. When combining the two methods, the interviews per hour rate was between 0.59 and 0.61.

1. INTRODUCTION

The American Community Survey (ACS) is a multi-modal survey carried out by the U.S. Census Bureau. The yearly sample is split into 12 roughly equal-sized panels, and each panel initiates the first mode of data collection at the beginning of each calendar month throughout the year. During each panel's first data collection month, respondents are encouraged to self-respond on either an Internet or paper instrument. After the self-response month, the first nonresponse followup mode is a computer-assisted telephone interview (CATI), which initiates at the beginning of the second data collection month.[1]

Before CATI data collection begins, ACS receives phone numbers from the Census Bureau's Center for Administrative Records Research and Applications (CARRA), who acquires telephone numbers from commercial vendors. Attached to each phone number is a model-based score reflective of the likelihood of having a correct address-phone number match. After removing successful self-response cases from the workload, ACS sends the 95,000 top ranked best-score cases (reduced from 110,000 in March 2016) to the Census Bureau's CATI Contact Centers (CCs) each month, along with second-best numbers for selected cases that also have a second-best number with a score above a set quality threshold. At the end of the CATI month, a subsample of cases for which there is still no completed interview is selected for a computer-assisted personal interview (CAPI).

In both the CATI and CAPI operations, respondent burden is a particularly important consideration, as both modes involve interviewers attempting to make direct contact with the sampled household, sometimes repeatedly, over the course of two months. Thus, a major goal of the ACS is to continually monitor, and if possible to reduce, respondent burden in the nonresponse followup CATI and CAPI modes of data collection.

This study focuses specifically on the CATI operation, with the two overarching research goals of finding ways to increase program efficiency and decrease respondent burden. CATI efficiency has been decreasing considerably over the past few years, which indeed is an almost universal phenomenon for telephone surveys in the past decade (Dutwin and Lavrakas 2016, Groves and Couper 2012). In fact, efficiency has decreased to such an extent that the program cost of a CATI interview is approaching that of a CAPI interview. However, the ACS is in the advantageous position compared to many other surveys in that there is the CAPI mode after CATI to followup with non-respondents.

This research first examines the potential for increasing CATI productivity by cutting callback attempts to telephone numbers that do not result in early contact. In an era where telephone surveys in general are dealing with higher nonresponse, it is possible that cases that receive

---

[1] For the telephone mode, cellular phone numbers are excluded and only landlines are used.

numerous, consecutive unproductive calls have a particularly elevated nonresponse rate. Thus, closing these cases sooner than the 12 callback attempt maximum that currently exists in the CATI program might improve overall efficiency. A second way to make CATI more efficient might be to restrict calling efforts to those cases that have higher-ranked match scores from CARRA, and thus are more likely to have a correctly matched phone number. Lastly, we examine the potential gains of simultaneously using both methods of cutting back callback attempts and reducing the CATI workload. If we are able to decrease the total number of calls per household or concentrate attention on households that are more likely to finish with an interview, the operation could decrease respondent burden and program costs considerably while maintaining a sizable share of CATI interviews from willing respondents. With these simulations, we will be able to estimate the associated effects in terms of completed interviews, labor hours, and labor efficiency relative to current parameter and workload settings.

2. RESEARCH QUESTIONS AND ANALYSIS METHODOLOGY

- Research Question 1. What are the general CATI productivity trends by the number of call attempts? Specifically, what is the distribution of final CATI case outcomes by the number of call attempts? What is the distribution of call attempt totals within CATI outcomes?

The analyses from the first research question provides a general layout of CATI outcomes. The results show the distribution of dispositions across call attempt frequencies. For example, for households that required six call attempts to reach a final disposition, what percent ended in a completed interview or sufficient partial interview? What percent reached the case-level limit for refusals? The analysis will also show the call attempt frequency distribution for each final outcome disposition category. For example, what percent of the total interviews were obtained after two call attempts? How many were obtained after five call attempts? If a relatively high percentage of interviews were obtained in early and middle rounds of calling, it might be feasible to cut the total number of call attempts without a sizable negative impact on the number of interviews.

To answer this research question, I use the CATI transaction file for the months of March – August 2016.[2] The CATI transaction file contains paradata about each occurrence of a case receiving a callback attempt, including the outcome of each callback attempt and cumulative case status. Thus, for each callback attempt within each case we have data about the callback attempt outcome, as well as the point-in-time final outcome if calling were stopped. The CATI transaction file contains an entry for every call placed in the CATI month by the CCs, with associated case identification variables, call outcome, other call details, and cumulative case

---

disposition. To sort cases into disposition categories, I utilized a hierarchical coding scheme similar to Zelenak and Davis (2013), which used a mix of final and intermediate disposition patterns to sort cases into eight larger disposition categories. These are detailed in Table 1. "Interview" includes cases that had a fully completed interview or a sufficient partial interview. Among non-interviews, a case is classified as a "Refusal: Hangup Maximum" if three immediate hangups are recorded, and a "Refusal: Other" if any implicit or explicit refusal occurs during the month without reaching the three-hangup threshold. Among cases that are neither a complete interview or refusal, a case is classified as one of the following: 1) reaching the unproductive[3] call maximum of 12 call attempts or 15 total calls if contact is made in the first 12 call attempts, 2) other noninterview, or 3) ineligible noninterviews or unconfirmed/disconnected telephone numbers. Cases that were never attempted or reveiwed a late mail return (LMR) during the CATI month are deleted from the analysis in this research question.

Table 1. 2016 CATI Case Disposition Categories

Disposition Category:	Includes:
Interview	Fully complete interview or
	Sufficient partial interview
Refusal: Hangup Maximum	Three immediate hangup call outcomes
Refusal: Other	Non-interview with at least one implicit or explicit refusal, but fewer than three immediate hangups
Call Maximum	Reached unproductive callback attempt maximum
Other, Noninterview	Sample unit eligible but unavailable
	Sample unit not found/unreached/eligibility uncertain
	Congressional case
	Insufficient partial
	Language barrier
	Hearing barrier
	Privacy detector
	Never contacted - confirmed number
Other, Ineligible	Sample unit ineligible, out of scope
	Never contacted – unconfirmed/disconnected number

To show how things have changed over time, I compare the data from 2016 to the 2012 data in Zelenak and Davis (2013).

---

[3] In this paper, an unproductive callback attempt is one that fails to make contact with the sample unit, such as a ring with no answer.

- Research Question 2. What are the residual interview completion, LMR, and refusal rates for the subgroups of cases in CATI that are still active after callback attempt numbers 2 through 14? How do these residual rates change when further subsetting the cases by those that made verified contact, those that had at least one refusal, or those that had at least one immediate hangup?

To answer this question, I use the CATI transaction file and consider only cases that are active after each given callback attempt round. For each round, I calculate the residual interview completion rate, LMR rate, and refusal rate for cases that are still unreconciled; as well as subgroups of cases that had at least one previous verified contact, at least one refusal, and at least one immediate hangup. By considering active cases after each round of callback attempts as a universe, we might see a more feasible threshold to use for closing unproductive cases. For example, if after the seventh round of calling the interview completion rates for cases still open falls dramatically, it would signify a drop in callback efficiency and potentially a reasonable place to cutback callback attempts. Another scenario might be that the residual interview rate falls for all cases, but stays relatively stable for cases that have made prior contact. This might indicate a particular point where prior contact is a reasonable predictor of interview completion, warranting additional callback attempts only for cases that make contact. If, on the other hand, the overall completion rate stays relatively steady for each round of calling, then more research might be needed before making recommendations.

The second part of this research question involves looking at subsets of cases that have received at least one verified contact, one refusal, or one immediate hangup in prior rounds of calling to see how the residual interview completion rates change. Currently, the ACS CATI program extends additional rounds of calling to cases that received at least one verified contact. This makes sense based on prior research, such as Martonik (2016), which found that cases that received at least one verified contact in early rounds of calling had much higher completion rates than cases that did not have a contact. ACS staff are additionally interested in seeing how cases that receive a refusal or hangup perform in later callback attempt rounds.

- Research Question 3. How would overall efficiency be affected by implementing a lower maximum to callback attempts, in terms of interview completion rates and interviewer time spent calling? What is the added efficiency by allowing additional callback attempts for cases that had at least one prior contact?

The third research question looks at the effect on interviews and efficiency from implementing cuts in callback attempts. This can be measured by dividing the total number of complete interviews after each round of callback attempts by the aggregated interviewer time spent calling. We then were able to compare the simulated productivity measures after each round to an average full CATI month.

Data for this analysis also comes from the CATI transaction file. With the results from this research question, we would know the change in the number of interviews, total labor time spent calling, and therefore interview efficiency based on cuts in callback attempts. Similar to the second research question, this analysis was extended by also examining the effect of allowing additional callback attempts past the initial cut for cases with prior verified contact. Similar productivity measures were calculated for the additional rounds of calls and cumulatively.

- Research Question 4. How would efficiency and productivity be affected by reducing the CATI workload from the top-ranked 95,000 cases? For each additional workload cut and cumulatively, what is the associated simulated completion rate and efficiency?

This question takes a similar methodology as Research Question 3, and applies it to investigates the effect of a stricter cut-off criteria for determining the CATI workload. Towards the end of the self-response month when the CATI workload is created, CARRA attaches a match score to each telephone number, and ACS produces a file with the top 95,000 ranked cases that have not self-responded. In essence, the scores are less important than the relative ranking of the scores in terms of being selected for the CATI workload. Within the highest-scored 95,000 cases, if a case has a second-best phone number with a score over a certain threshold, the second best telephone number is also included for additional calling. All months included in this analysis had 1.0 as this threshold, giving a relatively equal number of cases per month that were sent with a second telephone number.

- Research Question 5. How would efficiency be affected by using a combination of reducing callback attempts and the CATI workload?

This research question examines the effect of simultaneously cutting callback attempts and stricter cut-off criteria for determining CATI workload. Looking at the interaction of reducing callback attempts and the CATI workload may give us insight into the best method or combination of methods for making the CATI operation more efficient. For example, reducing the CATI workload to 60,000 might have only a slight impact compared to reducing the workload to 70,000, as does reducing callback attempts to six from eight. However, reducing the workload to 70,000 in combination with reducing callback attempts to eight might make the largest efficiency gain, while retaining the largest number of interviews.

3. LIMITATIONS

One limitation to this study is that the results of the simulations do not take into account changes to contact center (CC) staffing schedules and potentially changed callback planning. With a reduction in CATI workload or callback attempts, the overall volume of work for the CCs would decrease. In this scenario, it is possible for example that the CCs will stretch the callback

5

attempts over more time or alter callback scheduling in some other way. This might make it more difficult to accurately simulate productivity after a parameter or workload change based on figures prior to the change.

An additional limitation is that the simulations are based on only six months of CATI data from March through September 2016. CATI workloads are dependent on the self-response rates for the particular month, which have substantial seasonal effects throughout the calendar year. As such, it would be ideal to use at least 12 consecutive panels of data to run simulations in order to capture a full cycle, and using six panels of data may not fully control for these seasonal effects. However, ACS did not begin to receive CARRA match scores until the February panel for March CATI calling. We also know that CATI efficiency and productivity were altered with the change to using CARRA-provided telephone numbers (Mills 2016). Thus, we believe that past efficiency averages will be most accurately simulated restricting data to the panels subsequent to receiving CARRA-provided telephone numbers, which includes the six CATI months of March – August 2016.

4. LITERATURE REVIEW

For some time, response rates in the ACS CATI mode have fallen precipitously—a trend seen across telephone surveys (Dutwin and Lavrakas 2016, Groves and Couper 2012). The ACS CATI operation has the added obstacle of only calling landline telephone numbers, which are shown to have smaller population coverage each year. As of the end of 2015, the National Center for Health Statistics estimates that almost 50 percent of U.S. adults and 60 percent of U.S. children live in households without landline telephone coverage (Blumberg and Luke 2016). These figures are up from merely 10 percent in 2005. And among those that do maintain a landline, refusal and no answer rates have increased (Dutwin and Lavrakas 2016).

Additionally, the ACS recently began receiving matched phone numbers from the CARRA, which seems to be having a slightly negative effect on the number of interviews collected in the CATI month, as well as CC productivity (Mills 2016). Thus, we undertook this research to better understand CATI calling efficiency and how productivity might change through manipulation of the calling parameters and/or the CATI universe. One particular concern for telephone surveys when decisions about the sample size or number of callbacks are being considered is nonresponse bias (Biemer and Link 2007, Lin and Schaeffer 1995). However, the ACS CATI mode is in an advantageous position compared to other telephone surveys in that this can be addressed in the CAPI mode, which occurs after CATI.

Over the past few years, the ACS has researched and implemented several changes to the CATI operation to both reduce respondent burden and increase efficiency (Griffin 2013, Griffin and Hughes 2013, Zelenak and Davis 2013). Based on ACS research performed in 2013 (Slud and Erdman 2013, Zelenak and Davis 2013), the maximum number of calls was reduced from 25 to

15, and from 20 to 12 for unproductive calls. The maximum number of immediate hangups was also reduced from four to three. While these changes reduced the average number of interviews received in the CATI operation by about 11 percent, Griffin (2013) outlines several measures of calling efficiency and public burden that improved considerably in the CATI month immediately following the implemented parameter changes. Total call attempts dropped by almost 25 percent, call attempts per eligible case by 30 percent, and total login hours at the CCs by almost 18 percent (Griffin 2013).

Thus, the current research is well informed from this past work in considering ways to manipulate the parameters to reduce respondent burden and improve efficiency. We additionally put considerable focus on attempting to predict what will have the largest impact on the probability of completing an interview, based on the outcomes from early rounds of call attempts. This has been done with some success by some more recent telephone surveys, in that response rates were much higher for cases that had contact in early calling rounds compared to those with no verifiable contact in early rounds of calling (Martonik 2016). Our paradata allow for a similar detailed analysis, in hopes of retaining groups of likely responders if we follow through with decreasing the number of callback attempts.

5. RESULTS

In the following section, we analyze the data to answer the research questions posed in Section 2. Unweighted data are used in all analysis, as we are interested in the effects on the CATI operation rather than the effects on the population in general.

5.1.   What are the general CATI productivity trends by the number of call attempts? Specifically, what is the distribution of final CATI case outcomes by the number of call attempts? What is the distribution of call attempt totals within CATI outcomes?

Table 2 shows the distribution of final CATI case outcomes by the total number of calls made to resolve the case, while Table 3 shows the distribution of call attempts by CATI outcomes. Both compare Zelenak and Davis' (2013) figures from 2011-2012 (labeled "2012" in Tables 2 and 3) to data from the test period included in the current study (labeled "2016").[4] More specifically, for Table 2 we are interested in the distribution of outcomes for cases that closed out with a specific number of callback attempts. For example, the right column shows that in 2016, an average of 27.5 percent of cases were closed out in just one call attempt. Within those cases, 6.8 percent resulted in a completed interview. To compare with 2012, 12.6 percent of cases that closed with one call attempt obtained a completed interview. Among all cases that closed out after two call

---

[4] The figures presented in this paper differ slightly from Zelenak and Davis (2013) because: (1) they removed cases that self-responded during the Computer-Assisted Personal Interview (CAPI) month, and (2) I removed call attempts with a "busy signal" outcome where an immediate additional attempt was placed, as the call-counter also ignores these calls.

attempts, the 2016 data show that 23.2 percent had successfully completed an interview, compared with 38.8 percent from the 2011-2012 era. From 2012 to 2016, we see that the decrease in interviews in these cases was partially made up for by an increase in the proportion of cases that closed out with an ineligible telephone number, which are out of service or otherwise unable to be connected. The other columns and how they are calculated are detailed in the previous section.

A similar trend is seen throughout Table 2, with the share of cases closing out with a completed interview in each callback attempt dropping from 2012, and the share of non-interviews or ineligible telephone numbers increasing. For example, cases resolved with between four and 11 call attempts had completion rates ranging from 17.1 to 22.8 percent in 2016, compared with 36.4 to 46.8 percent in the earlier period. The share of cases that were a non-interview or ineligible telephone number increased considerably across these rounds. Reflected in the "call max" columns is where the unproductive callback attempt maximum was reduced from 20 to 12 attempts.

Table 3 shows the cumulative distribution of number of callback attempts within case-level outcomes. For example, 18.2 percent of all cases that eventually became completed interview were closed out after the first callback attempt. This compares with 16.2 percent for the earlier timer period. Likewise, after five call attempts 66.2 percent of the total interviews typically received across the test period had been completed. In the 2011-2012 era, 60.9 percent of the total interviews had been completed. Thus, if we were to set the maximum number of callback attempts to five, over one-third of completed CATI interviews that we typically receive would be lost.

We see aggregated outcomes across in the bottom row of Table 3. The 2016 panels completed less than half of the proportion of cases as compared to the 2011-2012 panels, dropping from 24.0 percent to 10.2 percent. The category that had the greatest increase was cases that reached the maximum number of unproductive call attempts, which more than doubled. This makes some intuitive sense, as the unproductive call maximum decreased from 20 to 12, likely closing out many cases earlier than would have occurred with more callback attempts. The proportion of cases that reached the maximum number of immediate hangups or result in a refusal increased modestly. Over half of all cases are coded as ineligible for an interview, and almost half of these are resolved with just a single call attempt, and over 90 percent of ineligible cases require six or fewer call attempts. Likewise, of the 17.4 percent of cases that received the maximum of unproductive calls, 91.4 percent of these were resolved after 12 calling attempts.

Table 2. Final Outcome by Total Number of CATI Callback Attempts

Call #	Int.[5] 2012	Int. 2016	Hangup 2012	Hangup 2016	Ref. 2012	Ref. 2016	Call Max 2012	Call Max 2016	Non-int. 2012	Non-int. 2016	Inelig. 2012	Inelig. 2016	Cumul. Total 2012	Cumul. Total 2016
1	12.6	6.8	--	--	--	--	--	--	0.3	0.4	87.1	92.9	30.9	27.5
2	38.8	23.2	--	--	3.1	6.1	--	--	1.7	3.5	56.5	67.2	41.6	35.6
3	26.1	8.7	--	0.6	4.0	4.1	--	--	3.9	3.6	66.0	83.0	51.9	50.0
4	37.6	17.1	0.6	1.8	7.2	9.8	--	--	7.1	10.7	47.4	60.7	57.5	55.7
5	44.8	22.5	1.4	3.2	9.2	14.0	--	--	8.3	14.2	36.4	46.3	61.4	59.3
6	36.4	17.2	1.5	2.6	9.0	12.1	--	--	6.4	11.2	46.8	56.8	65.2	62.9
7	44.2	20.2	2.2	3.5	12.0	14.1	--	--	7.1	13.7	34.5	48.5	67.8	65.5
8	45.5	22.8	3.2	4.4	13.4	16.7	--	--	6.6	14.5	31.4	41.6	69.9	67.5
9	46.8	21.9	3.7	5.1	14.6	18.5	--	--	6.5	13.6	28.4	40.8	71.6	69.2
10	37.4	22.2	3.4	5.7	12.8	20.3	--	--	4.9	14.8	41.5	37.0	73.5	70.6
11	42.8	20.8	4.1	5.0	15.8	20.4	--	--	5.9	14.7	31.4	39.1	74.9	72.0
12	40.7	1.5	4.5	0.4	16.2	1.8	--	93.3	7.5	0.8	31.0	2.3	76.2	89.0
13	41.4	8.6	4.7	2.8	17.8	17.5	--	53.8	10.0	6.1	26.1	11.1	77.3	91.1
14	39.9	13.1	4.6	3.6	20.2	35.5	--	18.0	11.9	12.9	23.4	16.9	78.3	92.2
15	35.0	3.8	3.7	0.5	20.4	49.7	--	2.2	11.2	32.7	29.6	11.0	79.4	100.0
16	26.4	--	3.1	--	18.8	--	--	--	12.2	--	39.5	--	80.6	--
17	23.3	--	3.2	--	19.6	--	--	--	13.9	--	40.0	--	81.9	--
18	20.9	--	3.0	--	19.4	--	--	--	14.8	--	41.9	--	83.1	--
19	20.8	--	3.3	--	21.0	--	--	--	16.3	--	38.6	--	84.2	--
20	2.6	--	0.4	--	3.0	--	89.7	--	1.7	--	2.5	--	91.7	--
21	9.6	--	1.7	--	13.9	--	59.6	--	8.4	--	6.9	--	93.4	--
22	16.2	--	2.9	--	28.0	--	24.3	--	16.4	--	12.2	--	94.2	--
23	17.4	--	3.9	--	32.3	--	14.8	--	21.3	--	10.3	--	94.8	--
24	17.6	--	3.4	--	36.6	--	7.9	--	23.0	--	11.4	--	95.3	--
25	5.1	--	0.4	--	49.0	--	0.5	--	37.1	--	7.9	--	100.0	--

Source: US Census Bureau, February 2011 - March 2012 and March - August 2016 ACS CATI Paradata

---

[5] In Tables 2 and 3: Int. = Interview; Ref. = Refusal; Non-int.= Non-interview; Inelig = Ineligible

Table 3. Cumulative Distribution of Total Number of Callback Attempts by Final Outcome

Call #	Int. 2012	Int. 2016	Hangup 2012	Hangup 2016	Ref. 2012	Ref. 2016	Call Max 2012	Call Max 2016	Non-int. 2012	Non-int. 2016	Inelig. 2012	Inelig. 2016
1	16.2	18.2	--	--	--	--	--	--	1.6	1.5	50.8	46.0
2	33.5	36.6	--	--	4.0	5.4	--	--	4.8	5.8	62.2	55.7
3	44.8	48.8	--	8.3	8.9	11.8	--	--	11.8	13.5	75.2	77.2
4	53.5	58.4	3.6	18.5	13.8	18.0	--	--	18.7	22.7	80.2	83.4
5	60.9	66.2	9.4	29.6	18.1	23.5	--	--	24.3	30.3	82.8	86.4
6	66.7	72.4	15.4	39.0	22.3	28.3	--	--	28.5	36.4	86.2	90.1
7	71.5	77.5	21.6	48.0	26.0	32.3	--	--	31.7	41.7	87.9	92.4
8	75.4	81.9	28.7	56.5	29.4	36.0	--	--	34.1	46.0	89.1	93.8
9	78.8	85.5	35.6	65.1	32.4	39.4	--	--	36.0	49.5	90.1	95.1
10	81.7	88.6	42.3	73.2	35.3	42.6	--	--	37.6	52.7	91.5	96.0
11	84.2	91.5	48.4	80.2	38.0	45.8	--	--	39.1	55.8	92.4	97.0
12	86.4	93.9	54.6	86.2	40.5	49.2	--	91.4	40.7	57.8	93.1	97.7
13	88.4	95.7	60.3	92.0	43.0	53.1	--	97.8	42.7	59.7	93.7	98.1
14	90.1	97.1	65.3	96.0	45.5	57.5	--	99.0	44.8	61.9	94.1	98.5
15	91.6	100.0	69.5	100.0	48.1	100.0	--	100.0	46.9	100.0	94.7	100.0
16	93.0	--	73.5	--	50.8	--	--	--	49.4	--	95.6	--
17	94.1	--	77.7	--	53.7	--	--	--	52.4	--	96.6	--
18	95.2	--	81.5	--	56.6	--	--	--	55.5	--	97.5	--
19	96.1	--	85.3	--	59.3	--	--	--	58.6	--	98.3	--
20	97.0	--	88.3	--	62.1	--	83.8	--	60.9	--	98.7	--
21	97.6	--	91.2	--	64.9	--	95.8	--	63.3	--	98.9	--
22	98.2	--	93.6	--	67.5	--	98.1	--	65.5	--	99.1	--
23	98.6	--	96.1	--	69.9	--	99.2	--	67.7	--	99.2	--
24	99.0	--	98.1	--	72.3	--	99.7	--	69.9	--	99.3	--
25	100.0	--	100.0	--	100.0	--	100.0	--	100.0	--	100.0	--
Total	24.0	10.2	0.9	1.0	8.3	9.1	8.1	17.4	5.8	6.7	52.9	55.6

Source: US Census Bureau, February 2011 - March 2012 and March - August 2016 ACS CATI Paradata

5.2. **What are the residual interview completion, late mail return (LMR) and refusal rates after rounds 2-14 of calling for those still active in CATI? How do these residual rates change when subsetting the by those that made verified contact, those that had at least one refusal, or at least one immediate hangup? Do cases that make prior contact complete an interview at a higher rate than cases in general?**

The second research question inquires about residual measures of productivity across attempted callback rounds. Table 4 shows the results of this analysis. The three columns in each section of the table denote a distinct final outcome of interest: (1) interview, (2) LMR, and (3) refusal. The table is split into four sections. The first from the left includes all cases still active after a given callback number. The other three represent subsets of cases, which from left to right are: (1) cases that received at least one prior contact, (2) cases the received at least one prior refusal, and (3) cases the received at least one immediate hangup. Importantly, each column is a smaller subset of the preceding column. Cases that received contact are a subset of all cases, a refusal is a type of contact, and an immediate hangup is a type of refusal. We examine these categories of cases due to specific interest from ACS staff.

For all cases still active after the first callback attempt, there is on average a 10.4 percent interview completion rate. This means that 10.4 percent of all cases still active after one callback round eventually close with a completed interview. For cases that received at least one prior verified contact and were still active after one callback attempt, the completion rate is 17.8 percent. Comparing these two columns shows that cases with prior contact have higher completion rates and are on average more productive in later rounds than all cases in general. Having a prior contact is associated with elevated completion rates well into later rounds of callback attempts. Cases with a refusal (which is also a verified contact) or an immediate hangup (which is also both a contact and a refusal) also had higher completion rates than all cases in general. This is an important finding, as we hypothesized that refusals (and especially hangups) might have particularly low response rates. This finding suggests that our current parameter of allowing cases that make contact (including refusals or hangups) additional callback attempts adds to overall efficiency.

Cases that make verified contact also have on average nominally higher LMR rates than cases in general. Cases that received a refusal or immediate hangup appear to have lower LMR rates. Due to the nested nature of the categories, this also suggests that cases that make a non-refusal contact have particularly elevated LMR rates, and although the data are not shown, this is indeed the case. For example, 10 percent of cases that are still open with at least one non-refusal contact after three callback attempts become an LMR, compared to 8.1 percent of cases with any contact. Overall, the data suggests that making contact in the CATI mode pushes respondents to self-respond.

11

Table 4. Interview, LMR, and Refusal Rates for All Cases and Selected Subsets of Cases that are Active After Each Round of Calling

Call #	Interview	LMR	Refusal	Interview, with prior contact	LMR, with prior contact	Refusal, with prior contact	Interview, with prior refusal	LMR, with prior refusal	Refusal, with prior refusal	Interview, with prior hangup	LMR, with prior hangup	Refusal, with prior hangup
1	10.4	8.5	12.8	17.8	9.9	37.2	12.6	5.3	66.2	12.0	5.6	65.0
2	9.1	7.1	13.9	16.7	9.1	38.6	12.4	5.1	67.5	11.5	5.1	67.1
3	9.5	7.1	17.0	15.2	8.1	39.9	11.8	4.7	69.3	11.0	4.6	69.3
4	8.6	6.1	17.8	13.9	7.2	41.0	11.1	4.4	71.3	10.3	4.3	71.5
5	7.5	5.1	18.1	12.6	6.4	42.0	10.3	4.0	73.4	9.5	3.9	73.9
6	6.6	4.3	18.6	11.8	5.9	45.2	9.4	3.6	75.6	8.7	3.4	76.3
7	5.7	3.5	18.9	10.6	5.1	47.0	8.5	3.1	77.9	7.8	2.9	78.7
8	4.8	2.8	19.0	9.4	4.4	48.7	7.6	2.7	80.3	6.9	2.5	81.0
9	3.9	2.1	18.9	8.1	3.6	50.1	6.6	2.2	82.7	6.0	2.1	83.5
10	3.0	1.4	18.7	6.8	2.8	51.4	5.6	1.8	85.1	5.1	1.6	85.8
11	2.2	0.8	18.4	5.5	2.1	52.5	4.6	1.4	87.5	4.1	1.3	88.2
12	3.8	1.3	48.1	4.2	1.4	53.6	3.6	0.9	90.1	3.3	0.8	90.5
13	2.6	0.7	52.5	2.8	0.7	56.2	2.4	0.4	92.8	2.2	0.4	93.1
14	1.4	0.0	53.9	1.4	0.0	57.4	1.3	0.0	95.5	1.2	0.0	95.6

Source: U.S. Census Bureau, March – August 2016 ACS CATI Paradata

5.3.   How would overall efficiency be affected by implementing a lower maximum to callback attempts, in terms of interview completion rates, and interviewer time spent calling? Does allowing cases that have made prior contact additional callback attempts add to overall CATI efficiency on average?

Table 5 is a simulation table that shows our best prediction of the effect of reducing the unproductive maximum callback attempts, as well as well as the additional effects of allowing verified contacts a certain number of additional callback attempts.[6] Effects are measured in terms of the number of interviews, the number of interviewer hours, interviews per hour of interviewer labor time, and the percentage cumulative completes received and interviewer hours spent in relation to an average CATI month. The bottom row shows that during the six-month test period, the CATI operation received an average of 8,336 interviews, utilizing 17,472 hours of interviewer labor, with a completes per hour ratio of 0.48.

For the simulations, the table shows the efficiency measures both by call and cumulatively. For example, the fifth callback round obtained 0.49 interviews per hour, but cumulatively (includes callback rounds 1-5, as if we were to make the callback maximum five attempts), there were 0.53 interviews obtained per hour over the first five callback attempts. This would amount to 68.5 percent of the typical number of interviews received, while expending 61.2 percent of the average hours interviewing. Directly under this row are up to six simulated additional callback attempts to cases with a prior verified contact. Interestingly, each of the additional six attempts are more productive (0.66, 0.61, 0.61, 0.58, 0.52, and 0.52 interviews per hour, respectively, for rounds 6-11 for verified contacts) than the fifth callback attempt to all cases. Setting the callback maximum at five attempts, with six additional attempts to verified contacts, would amount to 83.5 percent of the original number of CATI interviews while expending 73.2 percent of the original labor hours. Thus, it seems there may be some savings associated with this approach to changing the parameters. This difference diminishes somewhat as the simulated callback maximum increases. By the tenth callback round, 90.6 percent of original interviews are obtained, with 88.3 percent of original labor hours.

---

[6] The interview percentages in Table 5 are slightly different from those in Table 3 for a couple reasons. First, Table 3 excludes cases that closed out as an LMR, whereas Table 5 includes them. Also, Table 5 considers a sufficient partial as closed out after the first callback attempt for which the given case became classified as such, even though these cases typically receive several more callback attempts to complete the interview. Table 3 disregards when a case became a sufficient partial, and only considers the total number of callback attempts. Therefore, sufficient partials appear to have more callback attempts in Table 3, lending to slightly higher complete interview percentages in earlier rounds of Table 5. The simulations in Tables 6 and 7 are similar to Table 5 in this way.

Table 5. Simulation of Parameter Changes with a 95,000 CATI Workload

Call #	Extra Calls	Completes	Hours	Completes per Hour	Cumulative Completes per Month	Cumulative Hours per Month	% of Completes	% of Hours	Cumulative Completes per Hour
01	00	1,596	3,225	0.50	1,596	3,225	19.1	18.5	0.50
02	00	1,595	2,527	0.63	3,191	5,751	38.3	32.9	0.55
03	00	1,050	1,996	0.53	4,241	7,748	50.9	44.3	0.55
04	00	813	1,588	0.51	5,054	9,335	60.6	53.4	0.54
05	00	658	1,357	0.49	5,712	10,692	68.5	61.2	0.53
05	01	350	526	0.66	6,062	11,218	72.7	64.2	0.54
05	02	259	423	0.61	6,321	11,641	75.8	66.6	0.54
05	03	215	356	0.61	6,536	11,997	78.4	68.7	0.54
05	04	176	304	0.58	6,712	12,301	80.5	70.4	0.55
05	05	135	259	0.52	6,847	12,560	82.1	71.9	0.55
05	06	117	226	0.52	6,964	12,786	83.5	73.2	0.54
06	00	517	1,177	0.44	6,229	11,869	74.7	67.9	0.52
06	01	287	462	0.62	6,516	12,331	78.2	70.6	0.53
06	02	231	385	0.60	6,746	12,716	80.9	72.8	0.53
06	03	188	329	0.57	6,934	13,044	83.2	74.7	0.53
06	04	146	281	0.52	7,080	13,326	84.9	76.3	0.53
06	05	127	246	0.52	7,206	13,571	86.4	77.7	0.53
06	06	100	211	0.47	7,306	13,782	87.6	78.9	0.53
07	00	421	1,035	0.41	6,650	12,904	79.8	73.9	0.52
07	01	250	415	0.60	6,900	13,318	82.8	76.2	0.52
07	02	199	353	0.56	7,100	13,671	85.2	78.2	0.52
07	03	155	302	0.51	7,254	13,973	87.0	80.0	0.52
07	04	135	264	0.51	7,389	14,237	88.6	81.5	0.52
07	05	107	228	0.47	7,496	14,464	89.9	82.8	0.52
08	00	354	922	0.38	7,004	13,825	84.0	79.1	0.51
08	01	215	378	0.57	7,219	14,203	86.6	81.3	0.51
08	02	163	319	0.51	7,382	14,522	88.5	83.1	0.51
08	03	142	279	0.51	7,523	14,802	90.2	84.7	0.51
08	04	112	241	0.46	7,635	15,043	91.6	86.1	0.51
09	00	300	838	0.36	7,304	14,663	87.6	83.9	0.50
09	01	176	340	0.52	7,480	15,003	89.7	85.9	0.50
09	02	149	294	0.51	7,630	15,297	91.5	87.6	0.50
09	03	119	255	0.47	7,748	15,552	92.9	89.0	0.50
10	00	251	766	0.33	7,555	15,429	90.6	88.3	0.49
10	01	158	309	0.51	7,713	15,738	92.5	90.1	0.49
10	02	123	266	0.46	7,836	16,004	94.0	91.6	0.49
XX[7]	--	--	--	--	8,336	17,472	100.0	100.0	0.48

Source: US Census Bureau, March - August 2016 ACS CATI Paradata

---

[7] The bottom rows in Tables 5 and 7 (marked "XX") refer to the unmodified, present-day parameters.

5.4. **How would efficiency and productivity be affected by reducing the CATI workload from the top-ranked 95,000 cases? For each cut and cumulatively, what is the associated simulated completion rate and efficiency?**

We now ask how the CARRA model itself would function as a call parameter, while leaving in tact the current callback maximum parameters. The guiding hypothesis to this research question is that by restricting the CATI workload to the highest ranked cases via the CARRA scoring model, we might be able to increase average workload efficiency by omitting the potentially least-productive cases from the CATI workload. Table 6 shows the extent to which this would occur. In each month included in the analysis, ACS sent the top 95,000 scored cases, so I used these rankings to make simulated cuts. Similar to Table 5, Table 6 shows efficiency measures by group and cumulatively, beginning with the first cut of the first 40,000 cases. On average, the top-ranked 40,000 of 95,000 cases produced 72.6 percent of the total interviews each CATI month while utilizing 53.5 percent of the labor hours, netting 0.65 interviews per hour.

Descending within the table, we see the added efficiency of each additional 5,000 cases both as a group and cumulatively. Productivity descends somewhat rapidly, as the cases ranked 40,001 - 45,000 obtained 0.46 interviews per hour on average, followed by 0.42 for cases 45,001 - 50,000, and 0.35 for 50,001-55,000. The best-ranked 60,000 cases on averaged received 89.4 percent of the interviews of a 95,000-case CATI workload, and used 74.1 percent of the labor hours. Comparing Tables 5 and 6 gives us some clue as to the comparative efficiency of changing the calling parameters and reducing the CATI workload. For example, the top-ranked 60,000 cases produce nearly the same number of interviews as if the callback maximum were reduced to seven attempts with five additional callbacks to productive numbers. This would result in 89.9 percent of the interviews, but use 82.8 percent of the labor (versus 74.1 percent for the top 60,000 cases). In other words, to get 90 percent of the interviews of a typical CATI month under the current callback maximum parameters, we need the top ranked 60,000 - 65,000 cases, which would net an efficiency measure of 0.56 - 0.58 interviews per hour. Contrarily, by adjusting the parameters as in Table 5, we reach 90 percent of the interviews with an efficiency of 0.51. Thus, in terms of efficiency, these data suggest that cutting the workload is a somewhat more efficient way to increase efficiency than changing the parameters.

Table 6. Efficiency and Cumulative Productivity by Average Ranked CATI Workload Cut

Ranked Workload Cut	Interviews per Month	Hours per Month	Completes per Hour	Cumulative Interviews per Month	Cumulative Hours per Month	% of Completes	% of Hours	Cumulative Completes per Hour	Cumulative Calls Per Month	Cumulative Minutes Per Call
1 - 40,000	6,049	9,350	0.65	6,049	9,350	72.6	53.5	0.65	237,196	2.37
40,001 - 45,000	463	1,002	0.46	6,512	10,352	78.1	59.2	0.63	266,214	2.33
45,001 - 50,000	397	939	0.42	6,909	11,291	82.9	64.6	0.61	294,636	2.30
50,001 - 55,000	305	860	0.35	7,214	12,151	86.5	69.5	0.59	322,114	2.26
55,001 - 60,000	236	791	0.30	7,450	12,942	89.4	74.1	0.58	348,571	2.23
60,001 - 65,000	177	702	0.25	7,627	13,644	91.5	78.1	0.56	373,018	2.19
65,001 - 70,000	186	707	0.26	7,813	14,352	93.7	82.1	0.54	397,398	2.17
70,001 - 75,000	154	668	0.23	7,968	15,020	95.6	86.0	0.53	421,117	2.14
75,001 - 80,000	123	647	0.19	8,091	15,667	97.1	89.7	0.52	444,628	2.11
80,001 - 85,000	89	620	0.14	8,179	16,286	98.1	93.2	0.50	467,954	2.09
85,001 - 90,000	98	636	0.15	8,278	16,923	99.3	96.9	0.49	491,796	2.06
90,001 - 95,000	58	549	0.11	8,336	17,472	100.0	100.0	0.48	512,820	2.04

Source: US Census Bureau, March - August 2016 ACS CATI Paradata

5.5. How would efficiency be affected by using a combination of reducing callback attempts and the CATI workload?

In Table 5, we saw a simulation of how productivity would be affected if we changed the maximum callback parameters, while Table 6 showed potential effects of cutting the CATI workload to the highest quality cases. In Table 7, we perform joint simulations of combining both methods, whereby the maximum callback parameter is manipulated while cutting the CATI workload to the top-ranked 40, 50, 60, and 70 thousand cases. Included in each workload simulation is the percentage of complete interviews, labor hours expended, and complete interviews per hour compared to the workload averages during the test period.

It is helpful in comprehending the simulations in Table 7 to draw some efficiency comparisons between a given CATI workload cut and the parameter changes that would be required to get a similar percentage of interviews from a larger workload. For example, from Table 6 we saw that cutting the workload to 40,000 cases with no parameter change would result in 72.6 percent of typical CATI interviews at a 0.65 interviews per hour rate. In Table 7 we see that a similar percentage of interviews with a 50,000 case workload could be accomplished with six or seven callback attempts and extra attempts for verified contacts, with a 0.66 – 0.68 interviews per hour rate. However, if we cut the workload to 60,000 cases, a similar number of interviews can be obtained at a 0.64 - 0.65 interviews per hour rate, with five or six callback attempts plus extra for verified contacts. Contrarily, in Table 5, we saw that to get 70 - 75 percent of the interviews with only changing the callback attempts, it would bring efficiency to around 0.52 - 0.54 interviews per hour, which is considerably lower than when only cutting the workload or combining cutting the workload and callback attempts. Thus, it appears that obtaining about three-fourths of the average number of interviews per month would be most efficiently accomplished with a 50,000 case workload and some parameter changes.

Cutting to the workload to 50,000 cases with no parameter changes would net 82.9 percent of the typical CATI interviews in a month, at a rate of 0.61 interviews per hour. This percentage of interviews could also be accomplished by simultaneously cutting the workload to 60,000 and instituting a maximum of eight or nine callback attempts with extra calls to verified contacts, with 0.60 – 0.61 interviews per hour; or with a 70,000 case workload with seven callback attempts, resulting in 0.59 interviews per hour. As shown in Table 5, only cutting callback attempts could net a similar percentage of interviews with an efficiency of around 0.51 – 0.55 interviews per hour, depending on the specific callback parameter selected. Thus, achieving 80 percent of the average number of interviews per CATI month could be more efficiently accomplished with either cutting the workload with no parameter changes or a combiniation of cuts to the workload and parameter changes than by only cutting callback attempts. In fact, achieving 70 or 60 percent of typical CATI interviews shows a similar trend of cutting only callback attempts being the least efficient method, and cutting the workload as being slightly

Table 7. Simulation of Parameter Changes with a 40,000, 50,000, 60,000, and 70,000 Case Workloads

--	--	40,000	40,000	40,000	50,000	50,000	50,000	60,000	60,000	60,000	70,000	70,000	70,000
	Extra	% of	% of		% of	% of		% of	% of		% of	% of	
Call #	Calls	Completes	Hours	Efficiency	Completes	Hours	Efficiency	Completes	Hours	Efficiency	Completes	Hours	Efficiency
01	00	14.4	9.5	0.72	16.1	11.4	0.68	17.3	13.1	0.63	18.0	14.7	0.58
02	00	28.6	17.6	0.78	32.2	21.0	0.73	34.5	24.0	0.69	36.0	26.7	0.64
03	00	37.8	23.7	0.76	42.8	28.4	0.72	45.8	32.5	0.67	47.8	36.1	0.63
04	00	44.8	28.7	0.75	50.7	34.4	0.70	54.5	39.3	0.66	57.0	43.7	0.62
05	00	50.5	32.9	0.73	57.3	39.5	0.69	61.6	45.2	0.65	64.4	50.1	0.61
05	01	53.5	34.7	0.74	60.7	41.6	0.70	65.3	47.6	0.65	68.3	52.7	0.62
05	02	55.5	36.1	0.73	63.2	43.3	0.70	68.0	49.5	0.66	71.2	54.8	0.62
05	03	57.3	37.3	0.73	65.2	44.8	0.70	70.3	51.1	0.66	73.6	56.6	0.62
05	04	58.7	38.3	0.73	66.9	46.0	0.69	72.1	52.5	0.66	75.6	58.1	0.62
05	05	59.9	39.2	0.73	68.3	47.1	0.69	73.5	53.7	0.65	77.1	59.4	0.62
06	00	54.9	36.5	0.72	62.3	43.9	0.68	67.1	50.2	0.64	70.2	55.7	0.60
06	01	57.2	38.1	0.72	65.1	45.8	0.68	70.1	52.3	0.64	73.4	58.0	0.60
06	02	59.1	39.4	0.72	67.3	47.3	0.68	72.5	54.1	0.64	76.0	59.9	0.61
06	03	60.7	40.5	0.72	69.1	48.6	0.68	74.5	55.6	0.64	78.1	61.5	0.61
06	04	61.9	41.4	0.71	70.6	49.8	0.68	76.1	56.8	0.64	79.7	62.9	0.60
06	05	62.9	42.2	0.71	71.8	50.8	0.68	77.4	58.0	0.64	81.1	64.1	0.60
07	00	58.4	39.7	0.70	66.4	47.7	0.66	71.5	54.6	0.62	74.9	60.6	0.59
07	01	60.5	41.1	0.70	68.8	49.4	0.66	74.2	56.5	0.63	77.7	62.6	0.59
07	02	62.1	42.3	0.70	70.8	50.8	0.66	76.3	58.1	0.63	79.9	64.4	0.59
07	03	63.4	43.3	0.70	72.3	52.1	0.66	77.9	59.5	0.62	81.6	65.9	0.59
07	04	64.5	44.2	0.70	73.7	53.1	0.66	79.4	60.7	0.62	83.1	67.2	0.59
07	05	65.4	44.9	0.69	74.7	54.0	0.66	80.5	61.7	0.62	84.4	68.3	0.59
08	00	61.4	42.5	0.69	69.9	51.2	0.65	75.3	58.6	0.61	78.9	64.9	0.58
08	01	63.1	43.7	0.69	72.0	52.7	0.65	77.5	60.3	0.61	81.3	66.8	0.58
08	02	64.5	44.8	0.69	73.6	54.0	0.65	79.3	61.7	0.61	83.1	68.4	0.58
08	03	65.7	45.8	0.69	75.0	55.1	0.65	80.8	63.0	0.61	84.7	69.7	0.58
08	04	66.6	46.5	0.68	76.0	56.0	0.65	82.0	64.1	0.61	85.9	70.9	0.58
09	00	63.9	45.0	0.68	72.8	54.2	0.64	78.4	62.1	0.60	82.2	68.9	0.57
09	01	65.3	46.1	0.68	74.5	55.6	0.64	80.3	63.6	0.60	84.2	70.5	0.57
09	02	66.6	47.1	0.67	76.0	56.8	0.64	81.9	65.0	0.60	85.9	72.0	0.57
09	03	67.6	48.0	0.67	77.1	57.8	0.64	83.2	66.1	0.60	87.2	73.2	0.57
XX	--	72.6	53.5	0.65	82.9	64.6	0.61	89.4	74.1	0.58	93.7	82.1	0.54

Source: US Census Bureau, March - August 2016 ACS CATI Paradata

advantageous to using the combination of methods. However, for choosing a specific direction for the ACS CATI operation, these results can be used to weigh the effects of various scenarios in terms of costs, productivity, CC staffing, and longer-term CATI planning.

6. CONCLUSION

The ACS CATI operation has been losing efficiency for several years. By 2016, the program cost of obtaining an interview in the CATI mode had approached the cost of obtaining an interview in the CAPI mode. The purpose of this paper was to research ways to increase the efficiency of the CATI operation in terms of the number of interviews received and the calling time required to complete an interview. Using the first six months of data after changing telephone number sources to CARRA, this paper used CATI paradata to analyze efficiency across both callback attempts and CARRA-modeled rankings, and then uses past outcomes to simulate manipulating the callback parameters and workload.

It is clear by looking at Tables 2 and 3 that both productivity and efficiency have fallen since a similar, earlier study by Zelenak and Davis (2013). The proportion of the CATI workload that receives a completed interview has fallen considerably, while all other outcomes have increased, particularly the proportion of cases that are closed out due to never reaching the sample unit. Table 4 shows that even though efficiency has fallen, it is still somewhat possible to predict final outcomes based on early callback attempts. Cases that made contact in early rounds of calling had higher completion rates, even when that contact was a refusal or hangup. Cases that made contact also seem to have a higher propensity to result in a self-response.

A number of interesting observations came from the parameter and workload simulations. First, manipulating the callback maximums, the workload, or both the callback maximums and workload simultaneously resulted in efficiency gains. However, based on Tables 5 - 7, it appears that when choosing between these three methods, either only cutting the workload or combining workload and callback attempt cuts are both better methods to increase efficiency than by merely cutting the callback attempt maximum. In terms of choosing an optimal method, the data presented here suggest that cutting the CATI workload is slightly more efficient than combining both methods. However, ACS staff will likely need to use the data presented here, along with other research, to weigh various scenarios and staffing options for moving forward with the CATI program.

19

7. REFERENCES

Biemer, P. P. and Link, M. W. (2007) Evaluating and Modeling Early Cooperator Effects in RDD Surveys, in Advances in Telephone Survey Methodology (eds J. M. Lepkowski, C. Tucker, J. M. Brick, E. D. d. Leeuw, L. Japec, P. J. Lavrakas, M. W. Link and R. L. Sangster), John Wiley & Sons, Inc., Hoboken, NJ

Blumberg, S. J. and Luke, J. V. (2016), "Wireless Substitution: Early Release of Estimates From the National Health Interview Survey, July-December 2015." National Center for Health Statistics.

Dutwin, D. and Lavrakas, P. J. (2016), "Trends in Telephone Outcomes, 2008-2015," *Survey Practice*, Vol. 9(2).

Griffin, D. (2013) Effect of Changing Call Parameters in the American Community Survey's Computer Assisted Telephone Interviewing Operation. 2013 American Community Survey Research and Evaluation Report Memorandum Series #ACS13-RER-17.

Griffin, D., and Hughes, T. R. (2013) Analysis of Alternative Call Parameters in the American Community Survey's Computer Assisted Telephone Interviewing. 2013 American Community Survey Research and Evaluation Report Memorandum Series #ACS13-RER-11.

Groves, R. M., and Couper, M. P. (2012) *Nonresponse in household interview surveys*. John Wiley & Sons.

Lin, I. F., and Schaeffer, N. C. (1995) "Using Survey Participants to Estimate the Impact of Nonparticipation," *Public Opinion Quarterly* Vol. 59:236-258.

Martonik, R. (2016) Managing Efficiency in Telephone Surveys: Insights from Survey Paradata Trends 2010-2015, Paper Presented at the Annual Meeting for the American Association for Public Opinion Research in Austin, TX.

Mills, G. (2016) Evaluation of Transitioning Telephone Number Sources for the American Community Survey. 2016 American Community Survey Research and Evaluation Report Memorandum Series #ACS16-RER-12.

Slud, E. and Erdman, C. (2013). Adaptive Curtailment of Survey Followup Based on Contact History Data, Paper Presented at the Proceedings of the Federal Committee on Statistical Methodology (FCSM) Research Conference in Washington, D.C.

Zelenak, M. and Davis, M. (2013) Impact of Multiple Contacts by Computer-Assisted Telephone Interview and Computer-Assisted Personal Interview on Final Interview Outcome in the American Community Survey. 2013 American Community Survey Research and Evaluation Report Memorandum Series #ACS13-RER-08.