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

September 21, 2016

2016 AMERICAN COMMUNITY SURVEY RESEARCH AND EVALUATION REPORT
MEMORANDUM SERIES #ACS16-RER-12

MEMORANDUM FOR   ACS Research and Evaluation Advisory Group

From: Victoria Velkoff *(signed Victoria Velkoff)*
                       Chief, American Community Survey Office

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

Subject: Evaluation of Transitioning Telephone Number Sources for the
                       American Community Survey

Attached is the final American Community Survey Research and Evaluation report "Evaluation of Transitioning Telephone Number Sources for the American Community Survey." This report summarizes the effect on the Computer Assisted Telephone Interview (CATI) program resulting from changing the source of telephone numbers from third-party vendors to the Census Bureau's Center for Administrative Records Research and Applications (CARRA). The research examines changes in completed interviews, telephone contactability, and calling center productivity. Additionally, this report provides an initial look into the relationship between CARRA's telephone number matching algorithm and contactability. While most measures examined after the transition to CARRA show a decline in telephone number quality and contactability, it is difficult to separate these changes from the general downward trajectory in CATI efficiency that has been observed for some time. Additionally, there is a positive relationship between the CARRA telephone number-scoring model and contactability. It will be important to continue to monitor these trends in the coming months to verify these patterns as well as work to improve the CARRA modeling.

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.

September 21, 2016

Evaluation of Transitioning Telephone Number Sources for the American Community Survey

FINAL REPORT

Gregory J. Mills

United States™
Census
Bureau

American Community Survey Office

Page intentionally left blank.

Table of Contents

EXECUTIVE SUMMARY

At the beginning of 2016, the ACS implemented a design change for the Computer-Assisted Telephone Interview (CATI) data collection mode. Prior to 2016, the ACS had been acquiring telephone number data for sampled households from several third-party vendors. In a cost-saving effort, the ACS decided to let these contracts expire and begin using internal resources to receive telephone number-address matches from the Center for Administrative Records Research and Applications (CARRA). In addition to phone numbers, CARRA agreed to provide an associated score and rank for each telephone number within each panel, to enable ACS to target the numbers with the highest likelihood of successful contact. This design change was fully implemented with the January 2016 panel for February 2016 CATI calling. The purpose of this research project is to analyze the results of this change to the ACS CATI program in terms of interview responses, the relationship between CARRA model scores and contactability, and contact center (CC) labor productivity, for the first two months of CATI calling after the CARRA switch. In this analysis, we use unweighted data, as we are interested in the effects on the CATI operation rather than the population in general.

Both completion rates and contactability were in substantial decline throughout 2015. These changes seem in line with difficulty experienced by other telephone-based surveys. Completion rates and contactability seem to have declined further after switching to CARRA-provided telephone numbers; however, because of the previous downward trajectory, it is difficult to estimate how much of the decline was caused by the change in telephone number sources. The CARRA modeling strategy appears to function with some success. Higher telephone number scores are positively related with contactability. This relationship also held for the group of second-best numbers that ACS sent in March, suggesting that sending higher quality second-best numbers is a reasonable strategy. In a separate analysis of the lowest 10,000 scores for each month, however, completion rates were rather low and the relationship between CARRA scores and contactability was weak or nonexistent, suggesting limits to expanding the CATI workload. Finally, after the change to CARRA telephone numbers, efficiency has also fallen somewhat. However, it is difficult to estimate how much of this decrease is attributable to the transition to CARRA without further analysis.

The next steps in this line of research include continuing to monitor trends in contactability over time in the CATI program and its relationship with CARRA telephone number scores. This is particularly important as CARRA is expected to make small to moderate changes in its modeling strategy over time and as new or updated sources of telephone numbers are obtained. The purpose of these actions is to improve the production and efficiency of the CATI operation.

1. INTRODUCTION

The American Community Survey (ACS) is an ongoing multi-modal statistical survey conducted by the U.S. Census Bureau. The sample for each year is split into twelve roughly equal-sized panels that loosely correlate to the twelve calendar months of the year. During each panel's first month, respondents are encouraged to self-respond via an internet instrument, and later in the month via a paper questionnaire. If a response is not received and a landline telephone is matched to the address, an interviewer attempts to contact the respondent via the telephone during the next month.[1] A sample of cases that have not responded by the end of the telephone period is selected to be interviewed via personal visit.

At the beginning of 2016, the ACS implemented a design change for the Computer-Assisted Telephone Interview (CATI) data collection mode. Prior to 2016, the ACS had been acquiring telephone number data for sampled addresses from third-party vendors. In a cost-saving effort, the ACS decided to acquire telephone numbers from the Census Bureau's Center for Administrative Records Research and Applications (CARRA), who purchases the numbers from outside vendors. CARRA uses a model to determine which telephone numbers are most likely to reach the correct household, and they provide the model quality score and rank to ACS along with the telephone numbers. This design change was fully implemented with the January 2016 panel for February 2016 CATI calling.

The general goal of CARRA's modeling strategy is to be able to sift the many telephone numbers that are matched to a particular address to arrive at the telephone numbers with the highest likelihood of resulting in a successful contact with the correct address. After ACS provides CARRA with a list of addresses in a panel for which a telephone number is required, CARRA looks for matches in their various sources of telephone number data and gives a quality score based on the quantity of sources in which a particular telephone number was found, as well as the year the telephone number was confirmed. In general, a telephone number that has more matches across the vendor lists and more recent matches will have a higher quality score, although the model also takes into account that sources vary considerably in their quality. For example, some sources are known to be of such poor quality that matched numbers were categorically excluded from being sent to CATI. Each address and telephone number match has its own quality score, even where there exists two or more telephone numbers for a single address. This is an important point, as it allows us to select the "best" telephone number for each address when CARRA has several telephone numbers for a single address. Telephone number quality is especially critical because it directly influences the cost-effectiveness of the CATI operation.

---

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

The purpose of this research project is to analyze the results of this change to the ACS CATI program in terms of interview responses and contact center (CC) labor productivity, and to examine the relationship between CARRA model scores and contactability. However, it is important to note that while CARRA's modeling strategy and CC workload staffing and methodology potentially affect CATI outcomes and are worthy research enterprises, they are outside the scope of this report. Rather, this project will provide a more specific analysis of the success of telephone numbers scoring strategy in terms of interview responses, model score and contactability relationship, and CC labor productivity.

2. RESEARCH QUESTIONS AND ANALYSIS METHODOLOGY

- Research Question 1. How many complete interviews did ACS receive in the CATI operation? What is the completion rate for CATI? How does this compare with previous months?

The most basic question this research seeks to answer is how the overall number of complete CATI interviews changes with the introduction of CARRA telephone numbers. This analysis will include the number of complete interviews as well as completion rates to account for the change in CATI workload size variation. For the purpose of this report, complete interviews include completes and sufficient partial interviews. For this analysis, we use what is called the CATI transaction file, which has the outcome of all calls made, to calculate the number of complete interviews and the completion rate for January, February, and March 2016 CATI months, as well as the preceding 12 CATI months in 2015. In this way, we will be able to compare February and March 2016 with the same months from 2015, while also accounting for the overall decline in CATI response rates that has been occurring in recent years. Completions and completion rates will be calculated at the case level, as opposed to analyses in later sections that are calculated at the telephone number-level.

Note that the completion rates are unweighted, which differs from the weighted response rates used in many reports. We use unweighted data because our interest in this report is to find out how this change affects the operations, and we are not interested in making inferences to the entire universe of possible CATI cases. We use the following formula to calculate the completion rate (C):

Equation 1.1 $\quad C = (a + b)/(x - l)$

where a and b are completes and sufficient partials, respectively. In the denominator, x is the total number of cases, and l is the number of Late Mail Returns (LMRs) that also did not have a complete or sufficient partial interview. Indeed, ACS occasionally received completed interviews for a single case in more than one mode, in which case a separate algorithm is used to decide

2

which interview is kept. For the purpose of calculating the completion rate, we include cases in the universe when both an LMR and complete or sufficient partial interview was received.

- Research Question 2. What is the distribution of contactability in CATI after the switch to CARRA telephone numbers? Nominally, how does this compare to previous months?

For this project, we conceptualize contactability as the level of contact established with a particular telephone number that was called in ACS CATI. This is operationalized using a combination of the "outcome" and "outcome subtype" variables from the CATI transaction file, which contains these outcome codes for each phone call made during the CATI month. To obtain each phone number's level of contactability, outcomes from each call on the transaction file are sorted into a hierarchical contactability index.[2] Considering the cumulative contactability history of each phone number over the course of the entire month, we take the highest contactability score that occurred for each phone number. For example, if contact was made and address verified in the first call attempt, but contact was not made on any subsequent attempts, the phone number is still categorized as contacted and verified. Contactability has three main categories, in descending order of outcome quality:

1. Contact
   a. Complete or sufficient interview
   b. Verified telephone number (but not a sufficient interview)
   c. Refusal
2. Not enough information
3. Bad telephone number

Taken together, the umbrella category of "Contact" includes phone numbers that had outcomes as "complete," "verified telephone number," or "refusal," because all three indicate contact was made with the respondent signifying that the phone number from CARRA was of high contactability. A phone number is registered as "complete" if at some point ACS obtained a complete or sufficient partial interview from the phone number. A "verified" telephone number indicates that an interviewer obtained contact at the correct address, but did not complete an interview. An example of this is if the interviewer verifies that the phone number is for the sampled address, but the respondent claims to have already mailed in the paper questionnaire. A phone number is registered as a "refusal" if contact was made, but the respondent refuses to participate. Telephone numbers that register as "not enough information" are those that did not receive enough information to decipher their quality, such as a ring with no answer or an answering machine. Generally, "bad" telephone numbers are those that are not in service, are not able to be completed as dialed, or result in contact with ineligible or out of scope households.

---

[2] The contactability index was developed by the American Community Survey Office's Survey Analytics and Methods Branch and the Housing Unit Nonresponse Branch.

As with the first research question, given the seasonal and long-term fluctuation in CATI operation success, we will compare the February and March 2016 panel outcomes with outcomes from January 2016 as well as 2015 panels to see if any trends become elucidated while controlling for recent and seasonal trends. Analyses of the level of contact will be determined at the telephone number level rather than at the case level as in Research Question 1.

- Research Question 3. How do the telephone number scores provided by CARRA relate to the levels of contact shown in the CATI transaction file? When ACS sent second phone numbers for some households in March, how did they perform? And lastly, how do the 10,000 least promising phone numbers, or numbers that are typically on the cusp of being sent to CATI, perform? [3]

As mentioned above, CARRA provides a quality score for each telephone number. Higher scores are considered better and more likely to be correctly matched to the address. With this information, along with call outcomes from the CATI transaction file, we can statistically and visually verify if higher scores are indeed related to higher contact rates. Statistically, this can be achieved by performing a statistical test for correlation between telephone number score and contactability using the Spearman's Rho test, which is a monotone, nonparametric test and appropriate for both continuous and discrete variables. Since contactability is measured on an ordinal scale and telephone number score is a ratio variable, the Spearman's Rho test is appropriate for measuring correlation (Raveh, 1986). These data will be visualized by breaking down the distribution of scores into relatively small and evenly distributed score groups and fitting to each score group a graphic display of the distribution of outcome scores. With this data visualization, we will be able to see, for example, if the proportion of cases with contacts increases across score groups.

In March, ACS decided to send CATI two telephone numbers for selected households that had particularly highly rated telephone number match scores for two numbers. Having two phone numbers for a case adds almost no additional cost to the ACS program in the case that neither number results in an interview. This is because calling parameters occur at the case-level and thus results in minimal additional phone calls or time spent calling by CATI interviewers.[4] However, since the second-best number is only dialed after several unproductive attempts with the first-best number, it is likely that interviews completed with second-best telephone numbers would not have been otherwise obtained.

---

[3] In actuality, that number may vary somewhat. The research includes an analysis of the lowest ranked numbers, and 10,000 was chosen as a reasonable guess of what will be in the group of "worst numbers."

[4] Second-best matched phone numbers are not called until the first phone number is found to be bad or after seven unproductive calls, often leaving the second telephone number with a maximum of five call attempts.

Finally, we want to know how telephone numbers with a CARRA quality score that is typically on the cusp of being sent or not sent to CATI perform during the CATI month. For March, this included performing a separate analysis for the lowest scored 10,000 cases that were sent to CATI, about ten percent of the cases, which includes scores 0.1679595014 through 0.438421627. For February CATI, due to the modeling errors that occurred,[5] the analysis will look at cases with scores that corresponded to the lowest scored 10,000 cases in March. However, the February modeling errors also present an opportunity to learn more about numbers that typically would not have been sent to CATI, thus we will also separately analyze cases with a score greater than zero and less than 0.1679595014. Similar to the first part of Research Question 3, these cases will be split into smaller groups to visualize the distribution of contactability across improving scores.

- Research Question 4. How are CATI labor hours and labor efficiency affected by the CARRA telephone number changes?

Lastly, our data show that in recent years the cost-effectiveness of the CATI operation has changed to such an extent that the cost of a response in CATI is approaching or surpassing the cost of a response in the Computer-Assisted Personal Interview (CAPI) operation. One goal of this research is to evaluate the effects of CARRA modeling and targeted telephone number selection on cost in terms of labor hours and labor efficiency. These metrics will include: 1) total login hours and, 2) completes per login hour. As with Research Questions 1 and 2, these data will be nominally compared with previous panels in 2016 as well as the 2015 CATI panels in order to distinguish broad trends. These data will come from the ACS Daily Receipts Report. These measures are superior to actual financial cost numbers, as costs from the Financial Management Reports (FMRs) are influenced by many factors outside of the direct effort by CC staff to complete the cases. For this reason, this report will focus on labor efficiency and productivity.

3. LIMITATIONS

The main limitation of this study is that its scope is considerably narrow. The purpose of this project is explicitly to analyze the results of the change to CARRA-provided phone numbers in the ACS CATI program in terms of interview responses and Contact Center labor productivity, and to examine the relationship between CARRA model scores and contactability. Admittedly, many things affect household contactability and CATI completion rates in general that were not in the scope of this study, such as Contact Center calling or staffing procedures.

An additional limitation is that while this project directly examines the relationship between CARRA telephone number scores and outcomes, the CARRA modeling strategy itself is also

---

[5] These errors are further described in the limitations.

outside the scope of this project. It should be noted that beginning with the June CATI calling operation (May 2016 ACS panel), CARRA altered its modeling strategy based on results from February calling, and further research into CATI outcomes and CARRA's telephone number vendors will continue through other venues. However, these topics are not directly examined in this project.

We also want to note that although the ACS switch from third party to CARRA-provided telephone numbers was implemented in time for the February CATI month, the bulk of attention will be placed on March calling. This is due to some unexpected modeling errors that occurred with the CARRA telephone numbers for February calling, which resulted in ACS sending the CCs significantly lower quality telephone numbers than had been planned. This error essentially rendered the February 2016 calling month unrepresentative in terms of overall completion and contactability rates. However, February calling data are still included in this report because many of the numbers sent were of average or above average quality and analyzing the corrected telephone number scores with contactability will still shed light on the model's utility.

4. RESULTS

In the following section, we analyze the data to answer the research questions posed in Section 2, in the same order they were posed. 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.

4.1. How many complete interviews did ACS receive in the CATI operation? What is the completion rate for CATI? How does this compare with previous months?

Table 1 shows the number of complete CATI interviews and completion rates by CATI month from January 2015 to March 2016. Additionally shown are the averages for the first six months, the second six months, and for all twelve months of 2015. The figures in Table 1 take into account LMRs in a given month, which are excluded from the analysis if a completed CATI interview also did not occur. The number of monthly CATI interviews fell over the course of 2015. For the first six months of the year, CATI averaged over 11,000 interviews per month, but averaged only 9,681 in the second six months. The completion rate also fell throughout the year, starting with over 15.1 percent in January 2015 to 12.8 percent in December, and an average of 13.4 percent for the first six months of 2015 to 12.4 percent in the second six months. When CATI began using phone numbers from CARRA in February 2016, programming errors caused an initial steep decline in the number of interviews for the month. March 2016 CATI month finished with 8,455 interviews, which is a sizable drop from the nominal figures from 2015. Additionally, the nominal interview completion rate dropped by more than two percentage points from January to March 2016.

6

Table 1. Case-level Completes and Completion Percentages by CATI
Month, January 2015-March 2016

CATI Month	Completes	Non-LMR Cases	Completion Rate[1]
January 2015	12,829	85,031	15.1
February 2015	10,605	79,373	13.4
March 2015	11,021	82,703	13.3
April 2015	10,641	80,874	13.2
May 2015	10,324	79,810	12.9
June 2015	10,621	84,330	12.6
1st half of 2015	*11,007*	*82,020*	*13.4*
July 2015	9,977	78,957	12.6
August 2015	9,935	79,354	12.5
September 2015	9,441	76,851	12.3
October 2015	9,149	78,882	11.6
November 2015	9,692	78,088	12.4
December 2015	9,890	77,468	12.8
2nd half of 2015	*9,681*	*78,267*	*12.4*
Average 2015	*10,344*	*80,143*	*12.9*
January 2016	9,938	77,301	12.9
February 2016	6,592	101,667	6.5
March 2016	8,455	79,485	10.6

Source: U.S. Census Bureau, 2015 and 2016 ACS CATI Paradata
[1] Values are in percent

Thus, on the one hand, CATI responses and rates of completed interviews seem to have fallen after the implementation of using CARRA-provided telephone numbers. On the other hand, it is somewhat difficult to fully separate this decrease from the larger downward trends that had already been occurring for some time in the CATI operation, and indeed has been a reality that is nearly universally shared across phone surveys (Groves 2011, Keeter et al. 2006). The nominal year-over-year change from March 2015 to March 2016 was about 2,500 interviews.

4.2.   What is the distribution of contactability in CATI after the switch to CARRA telephone numbers? How does this compare to previous months?

The second research question asks about the distribution of outcomes at the telephone-number level. The results in Table 2 show the distribution of outcomes across CATI months from January 2015 to March 2016. The universe for the data in Table 2 is all phone numbers that were called at least once during the CATI month.

Table 2. Telephone Number-level Percent Distribution of Outcomes by CATI Month, January 2015-March 2016

CATI Month	1. Complete or sufficient interview	2. Good number, but not complete	3. Refusal	*Contact: 1, 2, or 3*	4. Not enough information	5. Bad telephone number	*Non-Contact: 4 or 5*
January 2015	12.6	13.0	4.1	*29.6*	19.4	51.0	*70.4*
February 2015	11.5	12.2	4.0	*27.7*	20.2	52.1	*72.3*
March 2015	10.5	15.1	4.5	*30.2*	21.0	48.8	*69.8*
April 2015	11.0	13.3	4.5	*28.8*	20.4	50.8	*71.2*
May 2015	11.0	12.6	4.2	*27.8*	20.2	51.9	*72.2*
June 2015	10.3	13.5	4.7	*28.6*	21.1	50.4	*71.4*
1st half of 2015	*11.1*	*13.3*	*4.3*	*28.8*	*20.4*	*50.8*	*71.2*
July 2015	10.7	12.5	4.8	*28.0*	20.9	51.1	*72.0*
August 2015	10.7	12.3	4.8	*27.8*	20.8	51.4	*72.2*
September 2015	10.3	13.0	4.7	*28.1*	20.7	51.3	*71.9*
October 2015	9.9	12.0	5.1	*27.0*	21.8	51.2	*73.0*
November 2015	10.4	13.1	4.9	*28.5*	20.7	50.8	*71.5*
December 2015	10.7	13.6	4.8	*29.1*	20.4	50.5	*70.9*
2nd half of 2015	*10.4*	*12.8*	*4.9*	*28.1*	*20.9*	*51.1*	*71.9*
Average 2015	*10.8*	*13.0*	*4.6*	*28.5*	*20.6*	*50.9*	*71.5*
January 2016	10.9	12.7	4.6	*28.2*	20.2	51.6	*71.8*
February 2016	6.2	7.7	3.0	*16.9*	14.6	68.5	*83.1*
March 2016	9.1	12.5	4.6	*26.2*	21.8	52.0	*73.8*

Source: U.S. Census Bureau, 2015 and 2016 ACS CATI Paradata

8

Similar to Table 1, the data in Table 2 seem to demonstrate a slight decrease in contactability over time. The nominal proportion of phone numbers that made contact with the sample household changed from 28.8 percent to 28.1 from the first half to the second half of 2015. This included a decrease in both complete interviews and verified phone numbers that did not complete an interview, as well as an increase in refusals. Data from after the switch to CARRA telephone numbers also shows a downward trend. In March 2016 only 26.2 percent of numbers resulted in a contact. This is a decrease of about two percentage points from January 2016 and four percentage points from March 2015. Bad numbers in March 2016 comprised 52.0 percent of telephone numbers called, which is a nominal increase of about 1.1 percentage points from the 2015 average, and a 0.4 percent increase from January 2016. In sum, CARRA telephone numbers appear to have moderately lower contactability than numbers used previously, but it is unclear the direct effect of the change to CARRA compared to the overall downward trend observed throughout 2015.

4.3. **How do the telephone number scores provided by CARRA relate to the levels of contact shown in the CATI transaction file? In addition, specifically, how do the 10,000 least promising phone numbers, or numbers that are typically on the cusp of being sent to CATI, perform?**

The third research question asks about the relationship between the phone number contactability scores provided by CARRA and CATI outcomes. As described above, CARRA produced a scoring algorithm based on the presence of phone numbers matched to housing units in various vendors' data across several years. [6] The various levels of contact are consistently represented across Figures 1-5 for ease of interpretation. Also in each figure, the black line indicates the relative proportion of cases in the respective sample or subsample that belong in the associated score group, labeled on the right y-axis. For example, in Figure 1, the score group "0.0-0.4" has the most cases, with between 24 and 27 percent as indicated on the right y-axis. The figures show percentage distributions of CATI contactability across CARRA score groups for February and March. We constructed score groups in the figures by finding relatively consistent score cutoffs in terms of number of cases in each group, combined with sizable score gaps in the score distribution. In Figures 1 and 2, the result was 11 score groups that each contain roughly seven to ten percent of the caseload based on March CARRA scores. February has an additional "missing" category that contains numbers that did not receive a score after the model was corrected due to poor quality. We used March scores to construct the score groups because, as described earlier, February scores were not representative due to programing errors. The same score group cutoffs are used in both figures so that the relationship between scores and outcomes is emphasized in the visualization, rather than the lower distribution of scores alone.

---

[6] Note that the actual numeric scores do not have a meaning other than higher scores are assumed to indicate a higher probability of contactability.

We use this same method when analyzing the March cases that were sent a second phone number in Figure 3, as well as the lowest 10,000 scored telephone numbers in Figures 4 and 5.

Figure 1. Contactability for February 2016 CATI, by CARRA Score Group

Sources: 2016 ACS CATI Paradata and CARRA telephone number matches

Figure 2. Contactability for March 2016 CATI, by CARRA Score Group

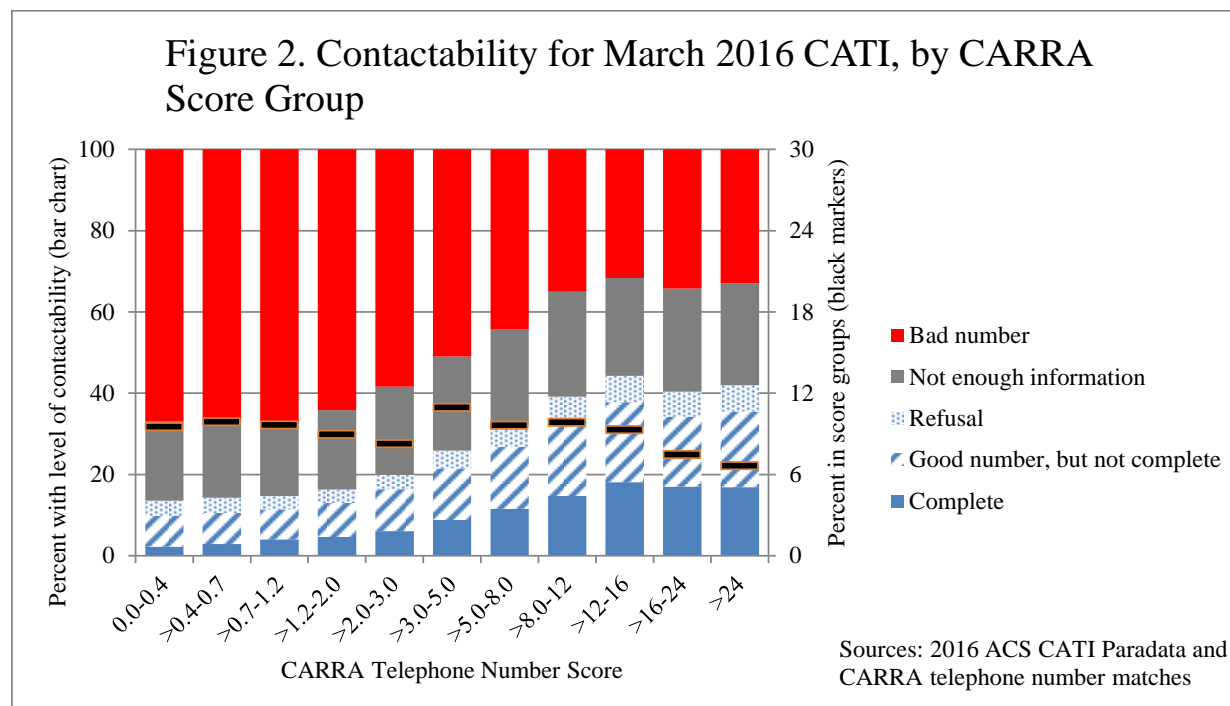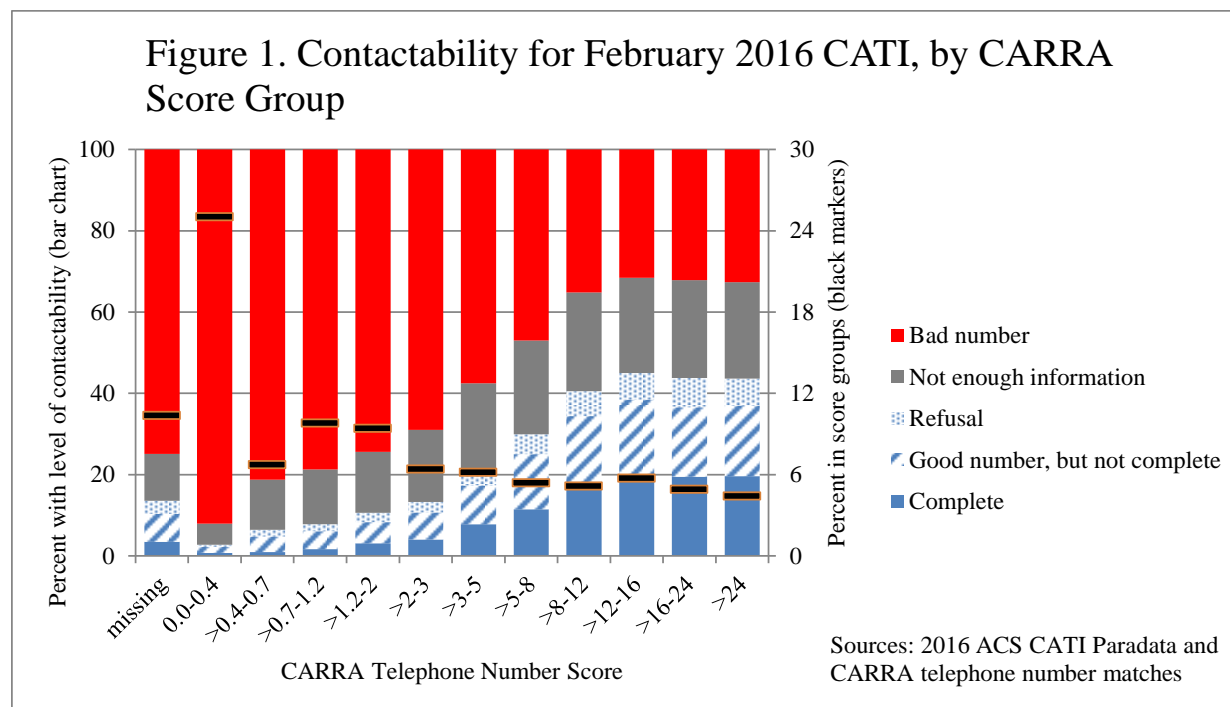Sources: 2016 ACS CATI Paradata and CARRA telephone number matches

Figure 1 shows that about one-fourth of the February phone numbers were in the lowest score group, which is a score between 0.0 and 0.4, which performed very poorly. Almost 90 percent of phone numbers in this score group were bad phone numbers, as indicated by the red area. However, although the February telephone numbers had a sizable oversampling of lower CARRA scores, there is still a visible relationship between CARRA score and contactability. Whereas the lowest three score groups had lower than a 10 percent contact rate (complete, verified, and refusals combined), the highest three groups had contactability rates of well over 40 percent, and completion rates of about 20 percent. Likewise, the proportion of bad phone numbers in the lowest three score groups was around 70 percent, but closer to 30 percent for the highest score groups. Figure 2 shows the results of March 2016 CATI. The relationship between CARRA scores and connectivity looks similar to February, with the lowest score groups containing about 70 percent bad telephone numbers and the highest groups between 30 and 40 percent bad telephone numbers. In addition, completion rates top out around 15-20 percent in the top score groups in both months. While there very well may be room for improvement in CARRA's modeling strategy, Figures 1 and 2 appear to suggest a considerable level of success in the modeling of contactability.

Figure 3. Second Number Contactability for March 2016 CATI, by CARRA Score Group
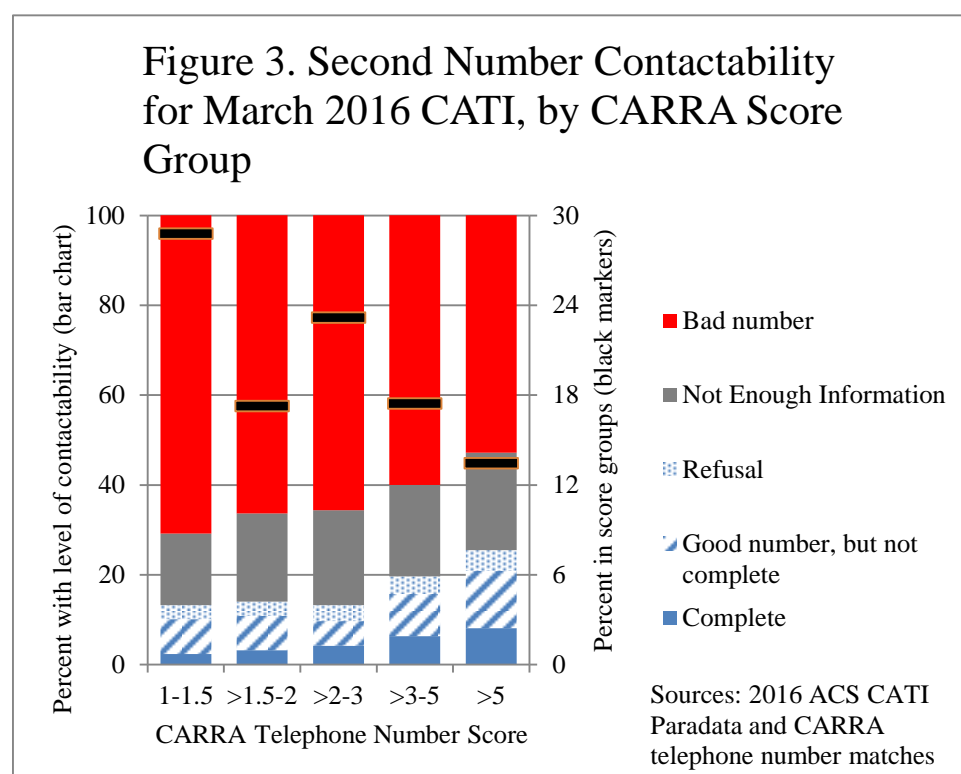
In March, ACS sent CATI interviewers a second number for cases where CARRA had a second telephone number with a score of 1.0 or higher. Figure 3 shows a separate analysis including only these second telephone numbers. ACS chose 1.0 as the cutoff because of the results from February and the desired percentage of cases to have a second number (from a workload

standpoint.) Slightly fewer than 5,500 such cases were sent to CATI. The figure shows that while the associated scores had a right-skewed distribution, the relationship between score and contactability was largely maintained. The highest scored second numbers had over a 20 percent contact rate and nearly a 10 percent completion rate.

We separately analyzed the telephone numbers that have the lowest 10,000 scores to gain insight into the predicted effect of increasing or decreasing the total number of cases sent to CATI. Again, due to February programming errors, these 10,000 lowest scores are based on the score distribution of March 2016, as were the score group ranges. For the February data in Figure 4, due to sending a high number of cases with very low scores, we are additionally able to analyze telephone numbers with scores below the lowest 10,000 scores for March. For both months, the score quality within the lowest scored telephone numbers seems to mean little in terms of CATI outcomes. For February, scores in the groups "0.213-0.269" and "0.269-0.328" have higher levels of contact, but higher scored groups "0.328-0.399" and "0.399-0.441" have higher completion rates. In March contact and completion rates show little variation across score groups, making contact in about 14 percent of cases, and having around a two percent completion rate.

Finally, for the score-outcome relationships in Figures 1-5, we calculated an associated completion percentage and Spearman's Rho correlation coefficient to gauge the relationship between CARRA's match model scores and CATI contactability. In the correlation test, we used the three-category measure of contactability described in Section 2, which is comprised of "Contact" (Complete, Verified, and Refusal), "Not Enough Information," and "Bad Telephone Number." Results are shown in Table 3, along with the associated numbers of completes and score ranges for each group being analyzed. It is important to keep in mind that the data used in the correlation analysis is unweighted and thus the correlation coefficients are meant only to reflect general relationships within the CATI operation.[7] When considering all called telephone numbers within each month, there exists a positive relationship between score and contact ($r = 0.292$ in February, $r = 0.481$ in March). The scores associated with the group of second phone numbers sent in March also show a positive relationship with contactability ($r = 0.106$). More importantly, the second-best phone numbers had a completion percentage of 4.44 and 244 completed interviews, which is worth noting as this was accomplished on fewer calls than a best-match only case, and likely would have been incomplete cases without the addition of second-best telephone number matches. The lowest 10,000 scores show a weaker relationship with contactability in either month, and had a completion percentage of less than one percent in February and 2.5 percent in March.

---

[7] Again, these correlation tests were also performed on weighted ACS data, and the relationships were consistent. We report the results on unweighted data here, as we are interested in the operational effects on CATI.

Figure 4. Contactability for Lowest 10,000 and Lower Scores for February 2016 CATI, by Score Group
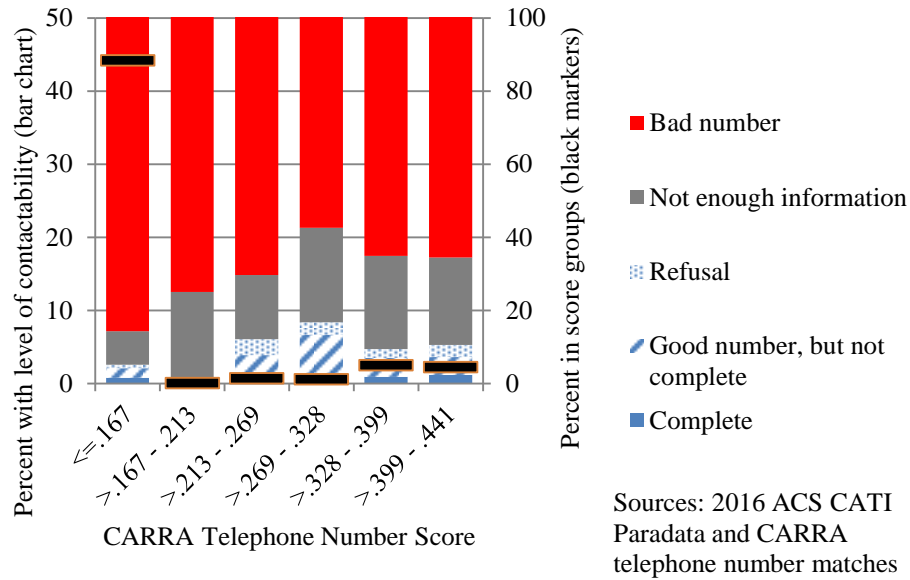
Sources: 2016 ACS CATI Paradata and CARRA telephone number matches

Figure 5. Cotactability for the Lowest 10,000 Scores for March 2016 CATI, by Score Group

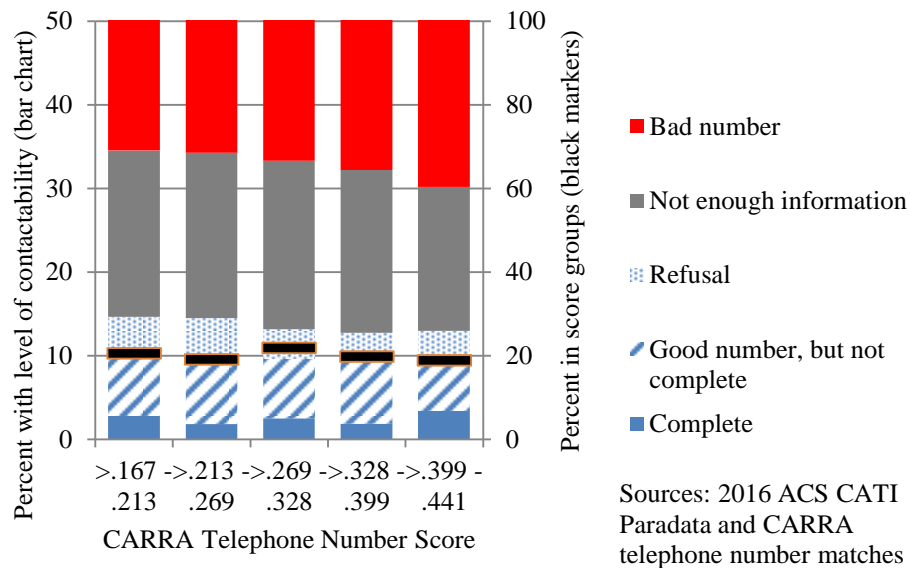Sources: 2016 ACS CATI Paradata and CARRA telephone number matches

Table 3. Percent Complete and Bi-variate Correlations

Month and Group	Total	Complete or Sufficient Interviews	Percent Complete	Score Range	Correlation Coefficient
February					
All Scored	94,891	6,188	6.5	0 - 146	0.481 ***
Bottom 10,000[1]	3,255	27	0.8	0.168 - 0.438	0.026
Below the Bottom 10,000	24,631	179	0.7	0.0 - 0.167	0.016 *
March					
All Scored	86,739	8,132	9.4	0 - 146	0.292 ***
2nd Numbers	5,497	244	4.4	1 - 33.781	0.106 ***
Bottom 10,000[1]	9,965	249	2.5	0.168 - 0.438	-0.017

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

Note: Analysis uses Spearman's Rho Correlation

*** $p < .001$

* $p < .05$

[1] "Bottom 10,000" refers to the lowest-scored 10,000 telephone numbers for March CATI

4.4.   How are CATI labor hours and labor efficiency affected by the CARRA telephone number changes?

The fourth research question asks about the effects of the change to CARRA-provided telephone numbers on interviewer efficiency. Efficiency is tracked through a "completes per login hour" measure provided by the National Processing Center (NPC) in the ACS Daily Receipts Report. Table 4 shows that efficiency fell throughout 2015, starting with a year-high of 0.66 completes per login hour in January 2015, to 0.55 in August and September and 0.56 in the final three months of 2015. Overall, the average completes per login hour was 0.59 in the first six months of 2015 and 0.56 in the second six months. Efficiency dropped precipitously in February 2016 to 0.45 after the CARRA change, which was expected to some extent due to the coding errors that sent lower quality phone numbers. However, March showed little improvement in efficiency with 0.48 completes per login hour. Thus, after the change to CARRA, efficiency is nominally lower than the prior downward slope in efficiency, but it is unclear how much this is directly due to the change to CARRA telephone numbers. More research into this area is needed to decipher why efficiency has decreased, and how this decrease relates to measures such as number eligibility, calls per case by outcome type, and interviewer login hours.

Table 4. CATI Center Efficiency by CATI Month, January 2015-March 2016

Month and Year	Login Hours	Completes/Login Hour
January 2015	19,465	0.66
February 2015	17,466	0.61
March 2015	18,810	0.58
April 2015	18,569	0.57
May 2015	17,640	0.58
June 2015	18,908	0.56
Average 1st half 2015	*18,476*	*0.59*
July 2015	17,545	0.56
August 2015	17,866	0.55
September 2015	17,096	0.55
October 2015	16,059	0.56
November 2015	17,126	0.56
December 2015	17,719	0.56
Average 2nd half 2015	*17,235*	*0.56*
Average 2015	*17,856*	*0.58*
January 2016	17,174	0.58
February 2016	14,727	0.45
March 2016	17,225	0.48

Source: U.S. Census Bureau, ACS Daily Receipts Reports

5. CONCLUSION

The purpose of this paper is to evaluate the ACS CATI program efficiency after the change to CARRA for phone number matching. Program efficiency in CATI has been decreasing for some time, and ACS staff made the decision to change telephone number sources to CARRA, who proposed using a modeling approach to predict cases that have the most likely address-telephone number match, and specifically target those cases in CATI calling.

We are somewhat limited in the definitive conclusions we can draw about the direct effect changing to CARRA for telephone numbers had on data collection, and this research does not delve into the modeling strategy or other covariates that might affect CATI responses. However, the analyses in this paper suggest that the change to CARRA had a slightly negative impact on nominal completion rates and contactability during the CATI month. These nominal effects do not appear to be large at this point in time, and it is difficult to separate the direct effects of the change to CARRA from the overall downward trend in the CATI operation. On the other hand, the CARRA model for predicting correct address-telephone number matches seems to function as designed, as telephone numbers with higher model scores had higher contactability rates,

which resulted in higher rates of obtaining a sufficient interview. It is conceivable that improvements to the matching algorithm could improve results further.

Finally, calling efficiency in terms of interviews per login hour has also been decreasing since last year, from 0.66 to 0.56 from January to December 2015. In March 2016, interviewers obtained 0.48 interviews per login hour, which appears to be a sizable drop in efficiency. More research is needed into this area to know how this might have been affected by the change to CARRA.

6. REFERENCES

Groves, R. M., (2011), "Three Eras of Survey Research," *Public Opinion Quarterly*, Vol. 75(5): 861-871.

Keeter, S., Kennedy, C., Dimock, M., Best, J., and Craighill, P., (2006), "Gauging the Impact of Growing Nonresponse on Estimates from a National RDD Telephone Survey," *Public Opinion Quarterly*, Vol. 70(5): 759-779.

Raveh, A., (1986), "On Measures of Monotone Association," *The American Statistician*, Vol. 40(2): 117-123.