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

2015 AMERICAN COMMUNITY SURVEY RESEARCH AND EVALUATION REPORT
MEMORANDUM SERIES # ACS14-RER-28-R1

MEMORANDUM FOR ACS Research and Evaluation Advisory Group

From: James B. Treat **(signed 01/26/2015)**
Chief, American Community Survey Office

Prepared by: Todd R. Hughes
American Community Survey Office

Subject: Reducing Respondent Burden in the American Community Survey's Computer Assisted Personal Visit Interviewing Operation – Phase 3 Results

Attached is the final American Community Survey Research and Evaluation report on reducing burden on respondents in the Computer Assisted Personal Visit Interviewing (CAPI) data collection mode. This project includes three phases. Phase 1 summarizes the current workloads, costs, burden, and quality associated with CAPI. Phase 2 identifies a set of potential stopping rules, and Phase 3 simulates the impacts of these stopping rules. This report has been corrected from the previously released version to correct errors associated with the calculation of the burden scores. Specifically, it includes recalculated numbers in the Tables and Figures and revised text referring to those exhibits. Additionally, Appendix A has been revised, the layout of Table 9 has been corrected, and a few rounding errors from the last version have also been corrected.

If you have any questions about this report, please contact Todd Hughes at (301) 763-6686.

Attachment

cc:
ACS Research and Evaluation Working Group
E. Slud (CSRM)
C. Erdman
M. Zelenak (DSSD)
P. Murphy
D. Daily (ACSO)
D. Nelson (CSM)
G. Walejko
T. Olson (FLD)
J. Magruder
F. Bradshaw

Reducing Respondent Burden in the American Community Survey's Computer Assisted Personal Visit Interviewing Operation – Phase 3 Results

FINAL REPORT

Deborah H. Griffin, Eric V. Slud & Chandra Erdman
American Community Survey Office &
Center for Statistical Research and Methodology

United States™ Census Bureau

Intentionally blank

Table of Contents

Tables

Figures

EXECUTIVE SUMMARY

The goal of this research was to identify a set of possible stopping rules to use in the American Community Survey (ACS) Computer Assisted Personal Interviewing (CAPI) operation to reduce the burden that sample households may feel due to the level of effort that we currently make to obtain high response rates. We analyzed data from the 2012 ACS to understand:

- how frequently we contact sample households,
- the proportion of households that express reluctance and how we address that reluctance,
- the cumulative burden that sample households might feel due to the sequential mixed mode design of the ACS,
- the quality of CAPI data collection in terms of response rates, item-level completeness of responses, and total number of completed sample interviews, and
- the number of hours spent during CAPI to convert noninterviews.

We proposed a set of 27 stopping rules that rely on paradata from the Contact History Instrument (CHI) and frame data such as geographic location. We estimated the costs, quality, and burden associated with these stopping rules, comparing each to the "full 2012 CAPI". This report summarizes the estimated reductions in burden that we might realize if we implemented these 27 stopping rules. It also estimates the associated losses in interviews. We estimate the potential cost savings in terms of CAPI hours saved.

We identified five rules that we believe hold promise to reduce respondent burden (based on several alternative measures) while maintaining high national levels of CAPI response. They include stopping CAPI data collection:

- after the second expression in CAPI of any form of reluctance (rule 6),
- after the first expression in CAPI of firm reluctance (rule 8),
- after four attempts if the case was a CATI refusal, after five attempts if it was another CATI noninterview, and after seven attempts for all other cases (rule 13),
- when the case reaches a cumulative burden score of 40 (rule 18), and
- when the propensity score falls below 20 percent (rule 26).

We estimate that four of these rules (6, 8, 18, and 26) would reduce the CAPI response rate from 95 percent to around 91 - 92 percent; thus reducing the number of CAPI completed interviews by 3 to 4 percent. One rule (13) would decrease the response rate to about 88 percent, reducing CAPI completed interviews by about 7 percent. We estimate reductions in total CAPI hours of 4 to 10 percent and 5 to 9 percent reductions in the number of total contacts made with sample households. All of these rules curtail followup for at least a third of the highest burdened households and about 10 to 14 percent of the CAPI visits that result in a household expressing reluctance.

Based on these results, staff will meet with Field Division to discuss the challenges associated with operationalizing each rule to determine which rule we could implement later this year or in calendar year 2015. We recommend additional analysis at subnational levels to understand the effects on data for specific areas and populations.

1. BACKGROUND

In Congressional testimony about the mandatory nature of the American Community Survey (ACS), it became clear that Congressional staff were advocating on behalf of constituents who felt "harassed" due to repeated efforts by the Census Bureau to obtain interviews (Poe 2011). This burden on sample households is a consequence of existing methods that include multiple mailings, repeated telephone call attempts using Computer Assisted Telephone Interviewing (CATI) methods, and potential personal visits using Computer Assisted Personal Interviewing (CAPI) methods. The ACS, like other household surveys, strives to maximize response to achieve the highest levels of quality by reducing the potential for nonresponse bias. In some instances, households can certainly perceive these multiple contacts and multiple modes as overly intrusive.

Unlike response rates, survey managers do not regularly measure respondent burden and response burden. It is not obvious to most managers which metrics best capture respondents' perception of burden in a survey setting. Respondents may perceive a survey to be burdensome based on the amount of time required to complete it or on the complexity of the questions. Some households could perceive a survey's repeated contact attempts as overly intrusive or burdensome.

The Census Bureau initiated research to understand the degree to which the ACS contacts sample households multiple times and the effectiveness of those follow up attempts (Zelenak & Davis 2013). Griffin & Hughes (2013) studied the CATI operation in detail to identify possible changes to the call parameters that would reduce the number of call attempts without a significant loss in completed interviews or increase in survey costs. The Census Bureau instituted changes in the ACS CATI call parameters in April 2013, reducing the maximum number of call attempts.

2. RESPONSE AND RESPONDENT BURDEN

The concepts of response burden and respondent burden are not well understood or consistently defined. The Office of Management and Budget considers *response burden* to be the time, effort, or financial resources the public expends to provide information. Traditional measures of response burden account for the frequency of a survey, the time required to complete it, and the number of survey respondents (Sears 2011, McCarthy 2011). Response burden measures attempt to quantify the effort expended by the public in responding to a survey request. Giesen et al (2011) suggest that aspects of response burden include actual burden (compliance costs in time and money), cognitive burden (difficulty), and irritation burden (perceived usefulness of the request). Others note that response burden measures should include a broader set of metrics to assess usability motivation, available time, respondent stress, and general perceptions of burden (e.g., Gravem et al 2011, Bradburn 1978, Jones 2012).

Frankel and Sharp (1981) define *respondent burden* as negative feelings (frustration, anger, annoyance, boredom) experienced by persons who participate in household surveys. They note the value of measuring respondent burden to "protect the public" using measures similar to the response burden measures noted previously, survey length and time burden imposed on a respondent. They also cited the value of measuring respondent burden to "protect survey methods", acknowledging the link between response rates and quality. They hypothesized that respondents with negative attitudes about survey participation may provide poor quality data. They conducted research to assess perceptions by survey respondents of how burdensome a survey request actually was. Fisher & Kydoniefs (2001) consider

respondent burden as the attributes of a respondent that affect the survey, including their attitudes towards the survey.

In this research, we do not use either of these traditional measures or definitions of response or respondent burden. Given the concerns expressed by Poe (2011), we wanted to assess and minimize the burden that survey respondents might perceive because we repeatedly contact them, despite possible objections to participate. We chose to measure "respondent burden" in terms of the amount of effort that is required to obtain survey responses. In this paper, we consider respondent reactions as well as total number of attempts as ways to assess respondent burden.

3. INTRODUCTION

This research focuses on the CAPI operation. Phase 1 summarized the current state of CAPI in the ACS using paradata from the Contact History Instrument (CHI), case histories, and CAPI outcomes from the 2011 and 2012 ACS. Zelenak (2014) and Griffin & Nelson (2014) document the level of effort, potential respondent burden, costs, and quality associated with the current CAPI methods and management of CAPI work. We consider these reports a baseline from which we looked for opportunities to reduce respondent burden without appreciable cuts in quality or increases in costs.

Concurrent with that analysis, an interdivisional group that included Field division brainstormed ideas of possible ways to stop data collection based on rules that use available paradata. In this Phase 2 activity, we identified potential changes to CAPI data collection rules that might reduce respondent burden. We created a list of possible interventions and reviewed each intervention with a focus on the complexity of its implementation, the potential for the intervention to reduce costs and burden on respondents, and the likely effect on quality. We identified 27 unique stopping rules that we hypothesized would reduce respondent burden in CAPI. Each rule involves some form of truncation of the data collection effort that we had in place in the 2012 ACS. Griffin (2014) summarizes the process we followed to identify these stopping rules (Phase 2).

Our goal is to determine the likely effect of implementing each of those rules in the ACS CAPI operation. In the third phase (the focus of this report), we use 2012 production data to estimate the effects of the proposed alternative rules on respondent burden, costs, and quality. Based on the combination of the cost/benefit analysis and the feasibility of implementation, we recommend possible changes to existing CAPI methods and procedures.

This report answers the following research questions:

1. What is the estimated effect on CAPI workloads associated with implementing each stopping rule?
2. What is the estimated effect on CAPI costs associated with implementing each stopping rule?
3. What is the estimated effect on CAPI burden associated with implementing each stopping rule?
4. What is the estimated effect on the quality of survey estimates associated with implementing each stopping rule?
5. Which rules meet our goal of reducing respondent burden without major losses in quality?
6. What are the trade-offs in workloads, costs, burden, and quality for the rules identified in research question #5?

4. METHODOLOGY AND LIMITATIONS

4.1 Data Sources

This research uses data from the 2012 ACS, specifically the January through December 2012 sample panels that include the March 2012 through February 2013 CAPI operations. We restricted the analysis to only those sample cases ultimately determined to be eligible for CAPI interviewing (i.e., we eliminated all CAPI cases coded as a Type C noninterview[1]). Griffin & Nelson (2014) found that in 2012 about 6.6 percent of the CAPI workload was ineligible for the survey. All workloads and costs in this report therefore understate the total workloads and costs associated with CAPI.

We base most of the summaries in this report on paradata from the CHI; specifically, contact attempt information and associated outcomes. We also use paradata from the CATI operation when describing contact histories and Cost and Response Management Network (CARMN) data to estimate time spent and costs.

4.2 Stopping Rules

As noted earlier, we identified a set of interventions based on available survey paradata that we thought we could use to define specific CAPI stopping rules. The rules fall into six basic categories; each involves some form of truncation of CAPI data collection, allowing us to estimate the possible effects using our existing full CAPI data collection dataset.

- Stop CAPI after a maximum number of total CAPI contact attempts
- Stop CAPI after certain types and outcomes of CAPI contact attempts (e.g., reluctance status)
- Stop CAPI after a maximum number of CAPI contact attempts based on CATI status
- Stop CAPI when reach maximum cumulative burden estimate
- Stop CAPI after a maximum number of CAPI contact attempts based on where the address is located
- Stop CAPI based on propensity scores[2]

We converted the general idea behind each stopping rule into a set of specific rules that use available paradata and/or frame data. We based our choice of thresholds in these rules on Phase 1 data, when available. Table 1 summarizes the 27 stopping rules that we studied. Refer to Griffin (2014) for detailed descriptions of each of these stopping rules.

[1] Type C noninterviews include addresses that do not meet the requirements for an ACS housing unit, including addresses that are associated with a unit that is under construction, demolished, condemned, or nonexistent.

[2] These propensity scores use contact history and frame characteristics to estimate the likelihood of the next contact attempt resulting in an interview.

Table 1. CAPI Stopping Rules

Rule	Type of Rule
	Maximum number of total CAPI contact attempts
1	when a case reaches 6 contact attempts (contact attempts includes all attempts, regardless of outcome)
2	when a case reaches 7 contact attempts (contact attempts includes all attempts, regardless of outcome)
3	when a case reaches 8 contact attempts (contact attempts includes all attempts, regardless of outcome)
4	when a case reaches 9 contact attempts (contact attempts includes all attempts, regardless of outcome)
5	when a case reaches 10 contact attempts (contact attempts includes all attempts, regardless of outcome)
	Maximum number of CAPI contact attempts with certain outcomes
6	after 2 contact attempts with any form of reluctance
7	after 3 contact attempts with any form of reluctance
8	after 1 contact attempt with a firm reluctance
9	after 2 contact attempts with a firm reluctance
10	after 2 contacts (with a household member) regardless of reluctance status
11	after 3 contacts (with a household member) regardless of reluctance status
	Maximum number of CAPI contact attempts, considering CATI status
12	when a case reaches a total of 3 contact attempts if the case was a CATI refusal, 4 if the case was another type of CATI noninterview, and 6 for all other cases
13	when a case reaches a total of 4 contact attempts if the case was a CATI refusal, 5 if the case was another type of CATI noninterview, and 7 for all other cases
14	when a case reaches a total of 5 contact attempts if the case was a CATI refusal, 6 if the case was another type of CATI noninterview, and 8 for all other cases
15	when a case reaches a total of 6 contact attempts if the case was a CATI refusal, 7 if the case was another type of CATI noninterview, and 9 for all other cases
16	when a case reaches a total of 7 contact attempts if the case was a CATI refusal, 8 if the case was another type of CATI noninterview, and 10 for all other cases
	Maximum cumulative burden score
17	when case reaches a cumulative burden score of 30 or more
18	when case reaches a cumulative burden score of 40 or more
	Maximum number of CAPI contact attempts, considering location
19	based on total contact attempts, vary based on FPS (Max of 6 if FPS1, Max of 8 if FPS2, Max of 10 if FPS3)[3]
20	based on total contact attempts, vary based on FPS (Max of 4 if FPS1, Max of 6 if FPS2, Max of 8 if FPS3)
21	based on total contact attempts, vary based on CSS (Max of 6 if CSS1, Max of 7 if CSS2, Max of 8 if CSS3, Max of 10 if CSS4 or CSS5)[4]
22	based on total contact attempts, vary based on CSS (Max of 4 if CSS1, Max of 5 if CSS2, Max of 6 if CSS3, Max of 8 if CSS4 or CSS5)
23	based on total contact attempts, vary based on historical survey response rates (Max of 6 if HRA, Max of 8 if not)[5]
24	based on total contact attempts, vary based on historical survey response rates (Max of 4 if HRA, Max of 6 if not)
	Minimum propensity score
25	if the propensity score is less than 15%
26	if the propensity score is less than 20%
27	if the propensity score is less than 25%

FPS = Field Performance Strata
CSS = CAPI Subsampling Strata
HRA = High Response Area

[3]Rules 19 and 20 use the three Field Performance Strata (FPS). This stratification partitions all addresses in a block group into one of three strata based on the expected ease of CAPI data collection in that block group. FPS1 includes the addresses that we expect to be the easiest to interview; FPS2 includes addresses that we expect to be harder, but not the hardest to interview; addresses in FPS3 should be the most challenging to contact and interview.

[4] Rules 21 and 22 use the CAPI Subsampling Strata (CSS). The ACS currently employs five different CAPI subsampling rates. We sample addresses in selected American Indian and Alaska Native areas at 100 percent. We sample addresses that we determined were unmailable, and therefore ineligible for previous data collection modes, at a rate of 2-in-3. We sample mailable addresses at 1-in-2, 2-in-5, and 1-in-3 based on the expected level of response in the self-response and CATI modes of their tract.

[5] Rules 23 and 24 define areas with historically high survey response rates and restricts the maximum contacts in these areas. In these rules, we identify areas with historical survey response rates that exceed 95 percent as high response areas (HRA).

4.3 Tabulations

We used the case histories and survey paradata from the 2012 ACS sample panels to tabulate a set of workloads that include tallies of contact attempts by mode and outcome. These workloads became the building blocks for the set of evaluation measures. All workload, burden, and cost tabulations are unweighted[6]. We produced two sets of measures: one for all eligible sample housing units and one for all eligible occupied sample housing units. The measures for occupied units are the best measures of burden that we are placing on our sample household respondents.

We organized the measures into the following groupings: interviews, contact attempts, reluctance, and cumulative burden. We tabulated each metric for each stopping rule and for our current (control) methodology. Refer to Griffin & Nelson (2014) for details of how we used CHI to define each of these measures.

Interviews - Total number of interviews

Contact attempts

- Total number of contact attempts
  - Total telephone attempts
  - Total personal visit attempts

- Total telephone attempts with contact made
  - Total interviews obtained in telephone attempts with contact made
  - Total noninterviews associated with telephone attempts with contact made

- Total personal visit attempts with contact made
  - Total interviews obtained in personal visit attempts with contact made
  - Total noninterviews associated with personal visit attempts with contact made

- Total attempts without contact (telephone or in-person) either with a sample household member or any other person
  - Total telephone attempts with no sample household contact
  - Total telephone attempts with no contact (sample household or other)

  - Total personal visit attempts with no sample household contact
  - Total personal visit attempts with no contact (sample household or other)

Reluctance

- Total contact attempts resulting in expressed reluctance[7] (of any form)
  - Total contact attempts resulting in expressed FIRM reluctance[8]

- Total contact attempts after expressed reluctance (of any form)
  - Total contact attempts after expressed FIRM reluctance

[6] We weighted the response data to account for the probabilities of selection when we estimated nonresponse bias (see section 4.7.2).

7 Expressed reluctance includes time, privacy, and anti-government concerns, among others.

[8] For this research project, we defined "firm" reluctance as attempts with specific strong reluctance outcomes that FRs recorded in CHI. These outcomes include hang-ups, door slamming, and hostile or threatening outcomes.

Cumulative burden

- Total cumulative burden[9]
- Total cases exceeding current 95th percentile of burden (score of 45.5)
- Total cases exceeding current 90th percentile of burden (score of 37.5)

4.4 Measuring Workloads

We used the tabulations described in section 4.3 to calculate the following workloads. For each metric, we define change as the difference between the control and each alternative – they are all reductions. We calculate percent reduction relative to the control. The first three measures include contact attempts associated with both vacant and occupied housing units; the last three are restricted to occupied units only.

- ***CAPI telephone contact attempts*** - Total telephone attempts with contact made (interviews and noninterviews) + Total telephone attempts with no contact (sample household or other)

- ***CAPI personal visit contact attempts*** - Total personal visit attempts with contact made (interviews and noninterviews) + Total personal visit attempts with no contact (sample household or other)

- ***Total CAPI contact attempts*** - CAPI telephone contact attempts + CAPI personal visit contact attempts

- ***Contacts with sample household (occupied interviews)*** - Total interviews obtained in personal visit attempts with contact made (occupied only) + Total interviews obtained in telephone attempts with contact made (occupied only)

- ***Contacts with sample household (occupied noninterviews)*** - Total noninterviews associated with personal visit attempts with contact made (occupied only) + Total noninterviews associated with telephone attempts with contact made (occupied only)

- ***No contact with sample household (occupied units)*** - Total personal visit attempts with no sample household contact (occupied only) + Total telephone attempts with no sample household contact (occupied only)

4.5 Measuring Costs

In the ACS, as in other surveys, Field Representatives (FRs) report their data collection costs in terms of hours spent and mileage for each day. They do not associate specific mileage or hours with each contact attempt. In order to assess the effects of a truncated interviewing cycle, we needed a measure of the effort expended for each interview attempt. Specifically, we needed an estimate of the average hours spent on contact attempts with the following six outcomes:

[9] See Appendix A for details on the methodology used to calculate cumulative burden.

- Telephone interviews – vacant addresses
- Telephone interviews – occupied addresses
- Telephone noncontacts or noninterviews
- Personal visit interviews – vacant addresses
- Personal visit interviews – occupied addresses
- Personal visit noncontacts or noninterviews

Staff in the Survey Analytics area of the Center for Survey Measurement within the Census Bureau developed a methodology that used an ordinary least squares regression model to produce estimates of hours spent for each CAPI contact attempt (Lawrence et al., forthcoming). This model is used to measure the *relative* amount of time spent on different types of contact attempts to inform decisions on how to allocate limited FR resources. They used CHI paradata and CARMN data on interviewing hours and mileage from the 2012 and the 2013 ACS, including Remote Alaska. For each interviewer-day, they modeled hours worked as a linear function of miles driven, personal visits resulting in an interview (occupied versus vacant), personal visits without a completed interview, and similarly for telephone attempts. These models were fit for each Regional Office. These data include interviews and noninterviews, occupied and vacant units, and cases ultimately determined to be ineligible for the survey.

The parameter estimates (βs) give the relative amount of time for each type of attempt. For example, a β of 0.21 for "pv_attempts_none" and β of 0.38 for a "pv_attempts_interview" indicate that an attempt resulting in an interview takes almost twice as long as an unsuccessful interview. We use these estimates to derive the proportion of the day's hours related to personal visits and hours that are telephone attempts. Mileage hours are included in the personal visit hours. Since there are no mileage data on the attempt-level, we made the simplifying assumption that, if there are X miles driven that day for N personal visits, then the miles driven for each personal visit attempt is X/N. We assumed that a telephone attempt has 0 miles associated with it.

Using the data from the model, this methodology assigned a number of hours worked to each type of attempt and therefore, to each attempt within a day. As a final step, the estimation process ratio-adjusted the resulting predicted hours to control totals for the total number of hours worked. Survey Analytics staff summarized these per contact attempt estimates for the 2013 ACS CAPI workload to estimate the total attempts by outcome and the total hours associated with those attempts. The authors of this model recognize that there are other approaches to estimating the relative time spent in a day by contact attempt type and outcome, but these Survey Analytics cost estimates are currently the best available for deriving preliminary cost implications of the alternative CAPI policies studied in this report. We used these estimates to approximate the average hours charged for each of the six bulleted outcome groupings above. Note that unlike the 2012 production data utilized in this report, the cost data included interviewing in Remote Alaska.

This methodology is one of several ways to approximate per case costs. Unlike other estimates in this report, these cost estimates are based on model assumptions. We provide these estimates as a measure of the relative cost implications of the alternative stopping rules.

4.6 Measuring Respondent Burden

We chose three different approaches to measure respondent burden. For each of these measures, we restricted our tabulations to occupied housing units.

The first approach defines burden by the total number of CAPI contacts that an FR makes with a sample household member. Given that a contact without an interview means a respondent was reluctant to participate, multiple contacts without an interview could indicate situations where the respondent feels harassed. To assess reductions in burden, we estimated the reduction in the number of total contacts with a sample household member each year and the percent difference relative to the current methodology.

- ***Contacts with sample household*** - Total telephone contacts with a sample household member (interviews) + Total telephone contacts with a sample household member (noninterviews) + Total personal visit contacts with a sample household member (interviews) + Total personal visit contacts with a sample household member (noninterviews)

We also looked at several measures that use paradata on respondent reluctance. Specifically, we measured the change and percent reduction in the number of contact attempts resulting in expressed reluctance (and firm reluctance) and the total number of contact attempts made after a respondent expresses reluctance (and firm reluctance).

- ***Contact attempts expressing reluctance –*** Total attempts with an indication that the respondent expressed some form of reluctance

- ***Contact attempts expressing firm reluctance*** – Total attempts with an indication that the respondent expressed firm reluctance

- ***Post-reluctance contact attempts*** - Total contact attempts after expressed reluctance

- ***Post-firm reluctance contact attempts*** - Total contact attempts after expressed firm reluctance

We based the third burden measure on a score that results from assigning different burden values to contact attempts, based on their outcomes. The basic idea behind this method is to tally each contact attempt (in any mode) as a separate increment of burden. We assigned a score based on the relative burden of the various contact attempts. We established a set of "incoming burden scores" based on the CATI/mail status. We acknowledged that cases that were mailable had four or five mail contact attempts. Cases that entered CATI with a good telephone number have additional contact attempts and possibly, additional contact outcomes that suggest levels of burden. The rule then increments the burden score for every CAPI contact attempt based on the type of attempt and its outcome.

We calculated three metrics related to this cumulative burden score. The cumulative burden score alone is hard to interpret. We consider the relative change as the key measure of reduction in cumulative burden. We also identified cases with the highest burden scores as the cases that, under our current methodology, had scores that exceeded the 90^{th} percentile and the 95^{th} percentile. We estimated the reduction in cases with these high burden scores under the new stopping rules.

- ***Mean cumulative burden score*** – Ratio of total cumulative burden to workload
- ***Cases with very high cumulative burden*** – Total cases with a cumulative burden score of 45.5 or greater (exceeding current 95th percentile)
- ***Cases with high cumulative burden*** – Total cases with a cumulative burden score of 37.5 or greater (exceeding current 90th percentile)

4.7 Measuring Quality

4.7.1 Measuring Response Rates, Completed Interviews, and Completeness Scores

To assess the effect of changes in stopping rules on quality, we considered three measures of quality: levels of nonresponse, final interviewed sample, and completeness scores. To measure the increase in unit nonresponse, we estimated the CAPI response rate under the current contact rules and again for each stopping rule. We then produced measures of the change and the percent reduction associated with each rule. We also calculated annual estimates of the reduction and the percent reduction in the estimated number of completed CAPI interviews.

- ***CAPI response rate*** – Ratio of total completed CAPI interviews to total eligible CAPI sample housing units

- ***CAPI interviews*** – Complete (occupied, temporarily occupied, and vacant) and sufficient partial CAPI interviews (occupied)

In the ACS, we calculate an aggregate completeness score for every completed and sufficient partial interview. It is a simple ratio of the number of items with nonblank responses to the total number of items requiring a response. The algorithm used to create these scores classifies entries such as "R" and "DK" for "refused" and "don't know" as blank responses. We summarize these completeness scores as percentages; so, a value of 100 means that the respondent answered every item that they should have answered. In our Phase 1 research, we determined that the completeness scores did not vary by number of contact attempts and other paradata (Griffin & Nelson 2014). Given this outcome, we chose not to use these scores as part of our assessment of quality in Phase 3.

4.7.2 Measuring Respondent Characteristics

To improve our understanding of the potential effects of increases in nonresponse on bias in survey estimates, we compared the characteristics of completed interviews under each stopping rule. We identified 24 ACS characteristics for our analysis (3 demographic, 11 social, 4 economic, and 6 housing). These 24 items came from two sources. We reviewed the key estimates used in the ACS Statistical Quality Standards (King 2012) and considered additional items that research identified as being more likely to vary across modes (Joshipura 2008). Table 2 summarizes the characteristics we selected.

We limited our research to assessments of the nonresponse bias associated with "lost interviews." We do not analyze the nonresponse bias that currently exists in the ACS CAPI (noninterviews under the current

methodology that we assume remain noninterviews). We also do not consider any potential reductions in bias resulting from noninterview adjustments.

To determine whether the specific reductions in CAPI follow up efforts associated with each rule resulted in an increase in nonresponse bias, we produced weighted estimates of these 24 characteristics for the completed interviews under each stopping rule. We only weighted for probabilities of selection. We did not make any adjustments for nonresponse or coverage. We used the replicate weight files to estimate the sampling errors associated with each of these estimates under each of the proposed rules. We calculated the difference between the completed interviews under each stopping rule and the completed interviews under our current methods.

We also calculated mean absolute percent error estimates for selected rules by averaging the absolute values of the percentage differences across all 24 characteristics. We used this as a way to compare the possible differences that each rule might introduce. See Appendix B for details.

Table 2. Characteristics for Analysis of Changes under New Policies

Characteristic
DEMOGRAPHIC
Percent of the total population reporting a race of Black or African-American alone
Percent of the total population reporting as Hispanic
Percent of the total population under 3 years
SOCIAL
Percent of the total population with a relationship of spouse
Percent of households that are living alone
Percent of the female population 15 years and over that are divorced
Percent of the population 25 years and over enrolled in college or graduate school
Percent of the population 25 years and over that have a graduate or professional degree
Percent of the population 18 years and over that are civilian veteran
Percent of the population 1 year and over with a residence 1 year ago of a different house in the U.S.
Percent of the total population that are non-citizens
Percent of the total population that are foreign born
Percent of the population 5 years and over that speak a language other than English at home
Percent of the total population with health insurance
ECONOMIC
Percent of the civilian labor force that is unemployed
Percent of workers 16 years and over commuting to work by car, truck, or van
Percent of total households (occupied housing units) with income and benefits of $15,000 - $25,000
Percent of total households (occupied housing units) with food stamp/SNAP benefits in the past 12 months
HOUSING
Percent of occupied housing units that are renter-occupied
Percent of total housing units built before 1940
Percent of total housing units with 1 bedroom
Percent of occupied housing units with house heating fuel from electricity
Percent of owner-occupied housing units with a property value of less than $50,000
Percent of occupied housing units that are in multi-unit structures

4.8 Reduction to a Set of Representative Rules

As noted earlier, the goal of this research was to identify a set of stopping rules that held promise to reduce respondent burden without a major loss in quality. Before pursuing analysis of cost, burden, and quality trade-offs, we wanted to eliminate any rules that did not meet the basic requirement of reducing burden. We also wanted to eliminate any rules that we felt introduced too great a risk in reducing quality.

The thresholds that we chose for quality were conservative but allowed us to winnow the large list to a more manageable number of rules to study in detail.

We looked across the set of respondent burden measures and identified the rules that consistently had the lowest reductions in burden. In this report, we dropped them from our later analyses. We used the following thresholds to determine the rules that met the burden reduction goal. We required that a stopping rule had to:

1. reduce total contacts with occupied sample households by 4 percent or more,
2. reduce total attempts in occupied cases expressing firm reluctance by 8 percent or more,
3. reduce post reluctance attempts in occupied cases by 15 percent or more, and
4. reduce "very high burden" occupied cases by 25 percent or more.

When considering how to define a "major" loss in quality, we identified rules with a final national-level CAPI response rate that fell to 85 percent or less. (These rules resulted in a 10 percent or greater reduction in completed interviews over all ACS response modes.) In addition, we looked to see if any rule had noteworthy differences in ACS characteristics among the completed interviews when compared to the completed interviews under the full CAPI.

4.9 Measuring Cost, Burden, and Quality Trade-offs

We chose to calculate ratios of selected cost, quality, and burden metrics as a way to assess cost/benefit trade-offs. Specifically, we calculated the ratio of cost savings associated with each lost interview and the following ratios of reductions in burden associated with each lost interview:

- eliminated contact attempts per lost interview
- eliminated contacts with a sample household member per lost interview (occupied only)
- eliminated post-reluctance contact attempts per lost interview (occupied only)
- eliminated "very high burden" cases per lost interview (occupied only)

In addition, we plotted several graphs summarizing the trade-off between measures of burden versus non-response. The idea in comparing rules using graphs like Figures 1 and 2 in Section 5.6.2 is that the better rules are those with burden and nonresponse pairs lowest and furthest to the left. Rules falling closest to the lower-left boundary (i.e., low burden and low nonresponse) are those that are most successful in trading lost interviews for reduced burden.

Ferguson (1967) clarifies why decision rules producing multiple loss outcomes at the lower boundary of the convex feasible-outcome region are decision-theoretically optimal. This idea of lower-boundary efficiency is analogous to H. Markowitz's Nobel prize-winning notion of the Efficient Frontier in Portfolio selection (the set of all portfolios that will give the highest expected return for each given level of risk). Line segments between plotted points associated with decisions A and B can be interpreted as the set of average outcomes obtained by mixture-policies choosing A with probability p and B with probability 1-p, as p ranges from 0 to 1. The plotted lower boundary shows the best set of outcomes including all such mixtures, where an outcome is "best" only if it cannot be improved in both burden

(measured by mean cumulative burden score and mean number of contacts with sample household) and nonresponse rate by the outcome from some probability mixture of available decisions.

4.10 Limitations

Several important limitations might lead to different results in production than those found in this evaluation. We expect that our Phase 3 estimates are an imperfect approximation of what might happen when we implement these revised rules in production. In CAPI production, unlike in CATI production, we cannot "truncate" data collection. The behavior of FRs and survey managers during CAPI will likely change when we impose new stopping rules. Operationalizing these rules is also not straightforward. We will need to assess the best ways to use existing technology and training to implement each of these rules. For these reasons, we believe that it is critical to monitor results under full implementation.

In addition, any intervention that reduces the number of completed interviews in a way that affects FR-specific response rates requires research into modifying ACS FR performance standards prior to implementation in order to manage performance under these new procedures.

The cost estimates included in this research involve models and multiple assumptions of per case costs that we may not fully realize in production. Differences could be due to associated staffing and scheduling implications, and limitations in the methodology used to estimate costs per case, per interview, and per contact attempt. We acknowledge that implementing any changes in production may yield cost savings different from those estimated in this analysis.

We limited our review of nonresponse bias to studying changes in the characteristics of the CAPI responses under each stopping rule. We did not look at total CAPI nonresponse or the potential mitigation of bias possible with noninterview adjustments or calibration to population controls.

A major shortcoming of this report is the reliance on national-level findings to select the best rules. We strongly recommend that the ACS Research and Evaluation Working Group pursue sub-national analyses before implementing any changes in CAPI production.

5. RESULTS

5.1 CAPI Workloads

What is the estimated effect on CAPI workloads associated with implementing each rule?

Table 3 includes projected annual reductions for 2012 in total CAPI contact attempts broken out by mode. The CAPI procedures encourage FRs to use the telephone to set up appointments or to obtain completed interviews. Table 3 displays estimates of change and percent reduction relative to current methods. We estimate that in the 2012 ACS FRs made 1,909,591 total contact attempts – 429,235 by phone and 1,480,356 in person. The stopping rules indicate opportunities for large reductions in these workloads.

Reductions in total contact attempts range from insignificant changes (less than 1 percent for rule 9) to major changes (over 10 percent for rules 1, 12, 13, 17, 20, 23, 24, and 27). It is interesting to note that the stopping rules appear to have the greatest effects on the telephone attempts. We find that many rules

result in more than a 10 percent reduction in CAPI telephone attempts (16 of the 27) with six rules reducing CAPI telephone attempts by more than 20 percent. The reductions in personal visit attempts range from under 1 percent under rule 9, to over 16 percent under rule 24.

Table 3. Change in 2012 CAPI Workloads with All households – Contact Attempts by Mode

	Total CAPI Contact Attempts		Telephone Contact Attempts		Personal Visit Contact Attempts	
Full CAPI	**1,909,591**		**429,235**		**1,480,356**	
Rule	**Change**	**Percent Reduction**	**Change**	**Percent Reduction**	**Change**	**Percent Reduction**
1	**-209,884**	**11.0**	**-82,718**	**19.3**	**-127,166**	**8.6**
2	-154,055	8.1	-61,744	14.4	-92,311	6.2
3	-114,087	6.0	-46,377	10.8	-67,710	4.6
4	-85,152	4.5	-35,227	8.2	-49,925	3.4
5	-64,023	3.4	-26,868	6.3	-37,155	2.5
6	-75,405	3.9	-34,811	8.1	-40,594	2.7
7	-22,964	1.2	-11,016	2.6	-11,948	0.8
8	-75,235	3.9	-27,007	6.3	-48,228	3.3
9	-14,661	0.8	-5,565	1.3	-9,096	0.6
10	-122,627	6.4	-59,386	13.8	-63,241	4.3
11	-39,040	2.0	-19,401	4.5	-19,639	1.3
12	**-308,191**	**16.1**	**-119,837**	**27.9**	**-188,354**	**12.7**
13	**-223,208**	**11.7**	**-88,613**	**20.6**	**-134,595**	**9.1**
14	-163,570	8.6	-66,021	15.4	-97,549	6.6
15	-120,871	6.3	-49,743	11.6	-71,128	4.8
16	-89,936	4.7	-37,544	8.7	-52,392	3.5
17	**-247,876**	**13.0**	**-98,248**	**22.9**	**-149,628**	**10.1**
18	-98,764	5.2	-40,286	9.4	-58,478	4.0
19	-154,738	8.1	-64,954	15.1	-89,784	6.1
20	**-298,984**	**15.7**	**-121,376**	**28.3**	**-177,608**	**12.0**
21	-77,406	4.1	-32,029	7.5	-45,377	3.1
22	-142,219	7.4	-56,938	13.3	-85,281	5.8
23	**-202,016**	**10.6**	**-80,466**	**18.7**	**-121,550**	**8.2**
24	**-387,853**	**20.3**	**-149,125**	**34.7**	**-238,728**	**16.1**
25	-56,683	3.0	-20,787	4.8	-35,896	2.4
26	-145,209	7.6	-50,904	11.9	-94,305	6.4
27	**-321,405**	**16.8**	**-106,398**	**24.8**	**-215,007**	**14.5**

The bolded rules have at least a 10 percent reduction in total contact attempts.
Source: January – December 2012 American Community Survey Sample Panels (March 2012 through February 2013 Computer Assisted Personal Interviewing)

Table 4 looks at the workload associated with occupied housing units only. It summarizes the change and percent reduction in contacts with sample households (by interview status) and the change and percent reduction in contact attempts that resulted in no contact with a sample household member. The attempts include both telephone and in-person efforts.

Several rules show promise to reduce unproductive efforts. Specifically, seven rules reduce the contacts with sample households that result in noninterviews by 10 percent or more (bolded). Four rules reduce the number of attempts resulting in no contact with the sample household by 20 percent (rules 12, 20, 24, and 27).

Table 4. Change in 2012 CAPI Workloads with Occupied Households - Contact Attempts by Outcome

Full CAPI	Contacts with Sample Household (Completed Interviews) 413,063		Contacts with Sample Household (Noninterviews) 222,776		No contact with Sample Household (Noninterviews) 874,122	
Rule	**Change**	**Percent Reduction**	**Change**	**Percent Reduction**	**Change**	**Percent Reduction**
1	-33,200	8.0	-19,928	8.9	-136,804	15.7
2	-22,921	5.5	-13,998	6.3	-103,089	11.8
3	-15,960	3.9	-9,959	4.5	-78,087	8.9
4	-11,245	2.7	-7,163	3.2	-59,417	6.8
5	-8,056	2.0	-5,157	2.3	-45,408	5.2
6	-17,682	4.3	-15,846	7.1	-40,145	4.6
7	-4,977	1.2	-5,363	2.4	-12,241	1.4
8	-14,083	3.4	-15,759	7.1	-42,467	4.9
9	-2,345	0.6	-3,730	1.7	-8,275	0.9
10	**-32,202**	**7.8**	**-24,423**	**11.0**	**-63,361**	**7.2**
11	-9,554	2.3	-8,795	3.9	-20,037	2.3
12	**-53,748**	**13.0**	**-32,481**	**14.6**	**-192,161**	**22.0**
13	**-36,275**	**8.8**	**-22,484**	**10.1**	**-144,116**	**16.5**
14	-25,007	6.1	-15,782	7.1	-108,558	12.4
15	-17,438	4.2	-11,173	5.0	-82,093	9.4
16	-12,252	3.0	-7,961	3.6	-62,328	7.1
17	**-51,530**	**12.5**	**-37,210**	**16.7**	**-147,360**	**16.9**
18	-18,129	4.4	-16,149	7.2	-61,085	7.0
19	-24,106	5.8	-13,870	6.2	-101,675	11.6
20	**-53,121**	**12.9**	**-29,182**	**13.1**	**-183,275**	**21.0**
21	-10,506	2.5	-6,483	2.9	-53,482	6.1
22	-21,877	5.3	-12,918	5.8	-93,724	10.7
23	-32,040	7.8	-19,010	8.5	-131,670	15.1
24	**-70,264**	**17.0**	**-39,940**	**17.9**	**-235,490**	**26.9**
25	-6,801	1.6	-5,435	2.4	-40,264	4.6
26	-21,639	5.2	-15,962	7.2	-96,337	11.0
27	**-56,268**	**13.6**	**-38,464**	**17.3**	**-198,906**	**22.8**

The bolded rules have at least a 10 percent reduction in contacts with sample households that result in a noninterview.
Source: January – December 2012 American Community Survey Sample Panels (March 2012 through February 2013 Computer Assisted Personal Interviewing)

5.2 CAPI Costs

What is the estimated effect on CAPI costs associated with implementing each rule?

The 2012 CAPI operation involved a total of nearly 1.5 million hours. Using the model described in Lawrence et al (forthcoming), we estimate that FRs spent an average of 1.17 hours for each personal visit contact attempt resulting in an interview of an occupied housing unit and 1.06 hours for each personal visit contact attempt resulting in a vacant interview[10]. Personal visit contact attempts resulting in noninterviews and noncontacts averaged 0.72 hours. Telephone contact attempts resulting in an interview averaged 0.62 hours if vacant and 0.72 hours if occupied. Telephone contact attempts resulting in a noncontact or noninterview required, on average, 0.28 hours. We used these per attempt costs to estimate the time and therefore, cost savings that might be possible under each of the stopping rules.

[10] For vacant units, selected housing information is collected from neighbors or an alternative source such as a rental office, or determined by observation.

Table 5 includes estimates of the change and the percent reduction in total interviewing hours by rule for all households. The bolded rules (12, 13, 17, 20, 24, and 27) have expected reductions of 10 percent or more. These results are consistent with the results in Tables 3 and 4.

Table 5. Change in 2012 Estimated Data Collection Costs for All Households

Full CAPI	Total CAPI interviewing hours	
Rule	**Change**	**Percent Reduction**
1	-130,166	9.4
2	-94,396	6.8
3	-69,141	5.0
4	-51,027	3.7
5	-38,008	2.7
6	-46,943	3.4
7	-13,922	1.0
8	-48,724	3.5
9	-9,169	0.7
10	-76,665	5.5
11	-23,862	1.7
12	**-194,175**	**14.0**
13	**-138,544**	**10.0**
14	-100,301	7.2
15	-73,215	5.3
16	-53,905	3.9
17	**-158,652**	**11.4**
18	-61,579	4.4
19	-94,079	6.8
20	**-186,863**	**13.5**
21	-46,515	3.4
22	-87,539	6.3
23	-124,951	9.0
24	**-246,566**	**17.8**
25	-34,802	2.5
26	-92,111	6.6
27	**-210,571**	**15.2**

The bolded rules have at least a 10 percent reduction in total CAPI interviewing hours.
Source: January – December 2012 American Community Survey Sample Panels (March 2012 through February 2013 Computer Assisted Personal Interviewing

5.3 Respondent Burden

What is the estimated effect on CAPI burden associated with implementing each rule?

Tables 6 through 9 summarize respondent burden measures by rule. All of these tables are restricted to occupied housing units.

5.3.1 Contacts with Sample Household Members

In 2012, we made 635,839 personal visit or telephone contacts with sample household members in the pursuit of interviews at occupied housing units. “Contacts” means an actual interaction with a sample household member. If we implemented new stopping rules, we might be able to eliminate as many as 110,000 of these CAPI contacts each year; a 17 percent reduction. Table 6 summarizes the estimated change and percent reduction in annual CAPI contacts with sample households (interviews and noninterviews) that are possible under each rule. Five rules suggest a 10 percent or greater reduction

(rules 12, 17, 20, 24, and 27 - bolded) while eight rules have little effect (less than four percent) on reducing this form of burden (rules 4, 5, 7, 9, 11, 16, 21, and 25).

Table 6. Change in 2012 Estimated Respondent Burden for OccupiedHouseholds – Contacts with Sample Household

	Contacts with Sample Household	
Full CAPI	**635,839**	
Rule	**Change**	**Percent Reduction**
1	-53,128	8.4
2	-36,919	5.8
3	-25,919	4.1
4	-18,408	2.9
5	-13,213	2.1
6	-33,528	5.3
7	-10,340	1.6
8	-29,842	4.7
9	-6,075	1.0
10	-56,625	8.9
11	-18,349	2.9
12	**-86,229**	**13.6**
13	-58,759	9.2
14	-40,789	6.4
15	-28,611	4.5
16	-20,213	3.2
17	**-88,740**	**14.0**
18	-34,278	5.4
19	-37,976	6.0
20	**-82,303**	**12.9**
21	-16,989	2.7
22	-34,795	5.5
23	-51,050	8.0
24	**-110,204**	**17.3**
25	-12,236	1.9
26	-37,601	5.9
27	**-94,732**	**14.9**

The bolded rules have at least a 10 percent reduction in contacts with sample households.
Source: January – December 2012 American Community Survey Sample Panels (March 2012 through February 2013 Computer Assisted Personal Interviewing)

5.3.2 Respondent Reluctance

In 2012, 236,802 of our CAPI contact attempts at occupied housing units resulted in the respondent expressing some form of reluctance to participate. FRs recorded 56,796 of these expressions as "firm." Expressing reluctance is a possible measure of burden because it indicates how often we asked a respondent to do something that they did not want to do. It is however, subjective on the part of the FR and can range from meaning, "not now" to "not ever." We can reduce the total number of reluctant outcomes (and firm reluctant outcomes) under our stopping rules. Table 7 includes estimates of the change and the percentage reduction in the number of CAPI attempts that result in reluctance (and firm reluctance).

As expected, the stopping rules performed consistently on the two measures. Rules 12, 17, 20, 24, and 27 (bolded) had estimated reductions in reluctant outcomes of 18 percent or greater (25 percent or greater for

firm reluctance). Several rules had only a minor effect on reducing reluctant or firm reluctant outcomes by less than four percent (rules 5, 7, 9, and 25).

Table 7. Change in 2012 Estimated Respondent Burden for Occupied Households – Reluctant Outcomes

	Attempts with reluctance Expressed		Attempts with firm reluctance Expressed	
Full CAPI	**236,802**		**56,796**	
Rule	**Change**	**Percent Reduction**	**Change**	**Percent Reduction**
1	-30,172	12.7	-10,761	18.9
2	-21,711	9.2	-7,992	14.1
3	-15,689	6.6	-5,937	10.5
4	-11,419	4.8	-4,491	7.9
5	-8,341	3.5	-3,417	6.0
6	-23,517	9.9	-8,902	15.7
7	-7,944	3.4	-3,232	5.7
8	-23,274	9.8	-13,526	23.8
9	-5,378	2.3	-3,710	6.5
10	-31,617	13.4	-10,796	19.0
11	-11,437	4.8	-4,194	7.4
12	**-46,581**	**19.7**	**-15,921**	**28.0**
13	-33,280	14.1	-11,829	20.8
14	-23,967	10.1	-8,856	15.6
15	-17,349	7.3	-6,625	11.7
16	-12,585	5.3	-4,965	8.7
17	**-52,391**	**22.1**	**-19,162**	**33.7**
18	-23,674	10.0	-10,230	18.0
19	-21,693	9.2	-7,948	14.0
20	**-42,873**	**18.1**	**-14,366**	**25.3**
21	-10,273	4.3	-4,063	7.2
22	-19,714	8.3	-7,209	12.7
23	-28,996	12.2	-10,328	18.2
24	**-56,901**	**24.0**	**-18,423**	**32.4**
25	-8,419	3.6	-3,895	6.9
26	-23,040	9.7	-9,137	16.1
27	**-51,089**	**21.6**	**-17,276**	**30.4**

The bolded rules have at least a 25 percent reduction in attempts expressing firm reluctance.
Source: January – December 2012 American Community Survey Sample Panels (March 2012 through February 2013 Computer Assisted Personal Interviewing)

We can also look at burden in terms of the additional efforts that FRs make to convert reluctant households. Specifically, we chose to study the number of contact attempts made after a household member expresses reluctance (or firm reluctance) to participate. Respondents may perceive these types of visits or calls as burdensome given that they chose not to participate during the previous contact.

We made 274,547 additional contact attempts in 2012 to convert reluctant cases (for firm reluctance, we made 74,187 additional attempts). We can reduce the burden that is likely associated with those attempts under many of these stopping rules. Table 8 shows that the most effective rules to reduce the attempts after firm reluctance is expressed include rules 8, 12, 17, 18, 24, and 27 (bolded).

Table 8. Change in 2012 Estimated Respondent Burden for Occupied Households - Post-Reluctance Attempts

	Post-reluctance attempts		Post-firm reluctance attempts	
Full CAPI	**274,547**		**74,187**	
Rule	**Change**	**Percent Reduction**	**Change**	**Percent Reduction**
1	-83,006	30.2	-28,711	38.7
2	-63,227	23.0	-22,484	30.3
3	-48,321	17.6	-17,566	23.7
4	-37,118	13.5	-13,758	18.5
5	-28,477	10.4	-10,766	14.5
6	-73,673	26.8	-32,116	43.3
7	-22,581	8.2	-12,151	16.4
8	**-72,309**	**26.3**	**-72,309**	**97.5**
9	-14,350	5.2	-14,350	19.3
10	-99,541	36.3	-35,268	47.5
11	-34,803	12.7	-15,109	20.4
12	**-115,809**	**42.2**	**-38,660**	**52.1**
13	-88,697	32.3	-30,849	41.6
14	-67,840	24.7	-24,367	32.8
15	-51,878	18.9	-19,082	25.7
16	-39,620	14.4	-14,879	20.1
17	**-154,907**	**56.4**	**-57,023**	**76.9**
18	**-77,297**	**28.2**	**-37,269**	**50.2**
19	-61,607	22.4	-21,224	28.6
20	-107,065	39.0	-34,631	46.7
21	-33,541	12.2	-12,532	16.9
22	-57,595	21.0	-20,552	27.7
23	-79,949	29.1	-27,610	37.2
24	**-136,814**	**49.8**	**-43,617**	**58.8**
25	-26,209	9.5	-14,873	20.0
26	-59,231	21.6	-28,491	38.4
27	**-107,791**	**39.3**	**-42,898**	**57.8**

The bolded rules have at least a 50 percent reduction in the number of CAPI attempts after firm reluctance is expressed.
Source: January – December 2012 American Community Survey Sample Panels (March 2012 through February 2013 Computer Assisted Personal Interviewing)

5.3.3 Cumulative Burden

As described earlier, we developed an algorithm to compute a cumulative burden score (described in Appendix A) that accounts for the combination of contact attempts and outcomes in mail, CATI, and CAPI. The score itself is hard to interpret, but we believe it is useful to define the potential change in the cumulative burden of our multimode contact efforts and also in identifying the reduction in highly burdened cases. We identified the cases that, under our current methodology, have scores exceeding the 90^{th} and 95^{th} percentiles. We refer to the 90^{th} percentile cases as, "high burden" cases and the 95^{th} percentile cases as, "very high burden" cases. Given the thresholds that defined these cases under our current methods, we estimated the reduction in cases exceeding these "high" and "very high" thresholds under each stopping rule.

Table 9 summarizes each rule in terms of its success in reducing cumulative burden based on these scores. Many rules are very successful in reducing the number of "high burden" and "very high burden" households. Rules 12, 13, 17, 18, 20, 24, and 27 all have expected reductions in "very high burden" cases of 50 percent or more (bolded). Note that rule 18 shows no reduction in the number of cases with high

burden. That is because rule 18 stops CAPI data collection when a case reaches the 90th percentile burden threshold and so it has no effect in changing the set of cases reaching it.

Table 9. Change in 2012 Estimated Respondent Burden for Occupied Households - Cumulative Burden

	Cumulative Burden		Cases with very high cumulative burden		Cases with high cumulative burden	
Full CAPI	**11,573,336**		**31,940**		**64,833**	
Rule	**Percent Reduction**	**Change**	**Percent Reduction**	**Change**	**Percent Reduction**
1	5.9	-14,581	45.6	-18,943	29.2
2	4.2	-11,305	35.4	-13,334	20.6
3	3.0	-8565	26.8	-9,242	14.3
4	2.2	-6,585	20.6	-6,239	9.6
5	1.6	-4,814	15.1	-4,326	6.7
6	3.2	-10,394	32.5	-8,795	13.6
7	1.0	-1,892	5.9	-896	1.4
8	3.3	-9841	30.8	-9,528	14.7
9	0.7	-1,331	4.2	-831	1.3
10	4.9	-14,527	45.5	-15,750	24.2
11	1.6	-4,217	13.2	-2,942	4.5
12	**9.0**	**-22,228**	**69.6**	**-33,141**	**51.1**
13	**6.4**	**-17,711**	**55.4**	**-23,819**	**36.7**
14	4.6	-13,761	43.1	-16,339	25.2
15	3.3	-10,429	32.6	-10,912	16.8
16	2.4	-7,691	24.1	-7,093	10.9
17	**8.8**	**-31,010**	**97.1**	**-50,121**	**77.3**
18	**3.7**	**-19,939**	**62.4**	**0**	**0.0**
19	4.2	-10,741	33.6	-13,349	20.6
20	**8.4**	**-17,440**	**54.6**	**-25,503**	**39.3**
21	2.0	-5,607	17.6	-5,533	8.5
22	3.9	-9,848	30.8	-11,472	17.7
23	5.6	-14,022	43.9	-18,150	28.0
24	**11.2**	**-21,758**	**68.1**	**-33,749**	**52.1**
25	1.6	-4,313	13.5	-4019	6.2
26	4.4	-10,653	33.4	-13,320	20.6
27	**10.0**	**-17,620**	**55.2**	**-26,348**	**40.6**

The bolded rules have at least a 50 percent reduction in cases with very high cumulative burden.
Source: January – December 2012 American Community Survey Sample Panels (March 2012 through February 2013 Computer Assisted Personal Interviewing)

5.4 Quality

What is the estimated effect on the quality of survey estimates associated with implementing each stopping rule?

5.4.1 Response Rates, Completeness Scores, and Completed CAPI Interviews

To assess the change in quality, we looked at two nonresponse metrics for all rules. First, we summarized the reduction in survey response by estimating unweighted CAPI response rates. Table 10 includes the estimated CAPI response rates, the percentage point change in CAPI response rate, the loss in CAPI interviews and the percent reduction in CAPI interviews. The percentage reduction in response rate is equal to the percentage reduction in total completed interviews. This loss in completed interviews is most likely to affect the quality (reliability) of survey estimates.

The 2012 ACS CAPI response rate (unweighted) was 95.0 percent resulting in 616,718 total completed interviews (occupied and vacant housing units). When we look at Tables 9 and 10 we see that some of the rules that did quite well in reducing burden did so at the expense of interviews and high response rates. Rules 12, 17, 20, 24, and 27 (bolded) all show at least a 9 percent reduction in completed CAPI interviews and therefore, drops in CAPI response rates of about 9 percent or more.

Table 10. Change in 2012 Quality for All Households – Response Rates and Completed Interviews

	CAPI Response Rate		Total CAPI Interviews	
Full CAPI	**95.0**		**616,718**	
Rule	**Rate**	**Change**	**Change**	**Percent Reduction**
1	88.7	-6.3	-40,758	6.6
2	90.7	-4.3	-28,082	4.6
3	92.0	-3.0	-19,577	3.2
4	92.9	-2.1	-13,782	2.2
5	93.5	-1.5	-9,873	1.6
6	92.1	-2.9	-18,888	3.1
7	94.2	-0.8	-5,291	0.9
8	92.6	-2.4	-15,823	2.6
9	94.6	-0.4	-2,567	0.4
10	89.7	-5.3	-34,317	5.6
11	93.5	-1.6	-10,116	1.6
12	**84.9**	**-10.1**	**-65,811**	**10.7**
13	88.2	-6.8	-44,206	7.2
14	90.3	-4.7	-30,336	4.9
15	91.8	-3.3	-21,115	3.4
16	92.7	-2.3	-14,860	2.4
17	**86.2**	**-8.8**	**-57,313**	**9.3**
18	92.0	-3.1	-19,862	3.2
19	90.4	-4.6	-29,820	4.8
20	**84.7**	**-10.3**	**-66,870**	**10.8**
21	93.0	-2.0	-12,906	2.1
22	90.9	-4.2	-26,975	4.4
23	89.0	-6.1	-39,333	6.4
24	**81.5**	**-13.5**	**-87,648**	**14.2**
25	93.8	-1.3	-8,195	1.3
26	91.1	-4.0	-25,738	4.2
27	**84.7**	**-10.3**	**-67,022**	**10.9**

The bolded rules have at least a 9 percent reduction in total CAPI interviews.
Source: January – December 2012 American Community Survey Sample Panels (March 2012 through February 2013 Computer Assisted Personal Interviewing)

5.4.2 Changes in Respondent Characteristics

In addition to analyzing response rates and lost interviews, we studied the effect that the increase in nonresponse might have on nonresponse bias. To assess nonresponse bias, we chose to compare the characteristics of the completed CAPI interviews under each stopping rule with the characteristics of the completed CAPI interviews with the current, full CAPI level of effort. Tables 11 - 14 summarize the differences in the characteristics of the CAPI respondent populations if we implemented these stopping rules rather than continued follow up efforts. Each table also displays the mean standard error for the differences. We see very little variation in the standard errors across rules.

Table 11 summarizes national-level results for six demographic characteristics. In the 2012 CAPI, 16.3 percent of the completed CAPI interviews reported a race of Black or African-American alone. This table shows that under rule 1, for example, the percent of the CAPI completed interviews reported as Black or African American alone dropped by 0.4 percentage points (to 15.9 percent). The standard errors for these differences were approximately 0.1 percentage points. If we look across rules, we see a range of differences from less than 0.05 percentage points to 0.6 percentage points. For percent Hispanic, the mean difference was 0.2 percentage points and the maximum difference was 0.6 percentage points. Most rules had differences in the population under age 3 that fell below 0.05 percentage points. Percent spouse and percent of females divorced also had small differences, a mean difference of 0.1 percentage point and maximum differences of 0.3 percentage points. The differences in the percent of the population living alone ranged from less than 0.05 to greater than 1 percent.

Table 11. Differences in the 2012 Characteristics of the CAPI Respondent Populations - Demographic Characteristics

	Difference from Full CAPI					
Rule	Percent Black	Percent Hispanic	Percent Under 3 Years	Percent Spouse	Percent Living Alone	Percent of Females Divorced
Full CAPI	**16.3**	**25.1**	**4.9**	**14.6**	**26.9**	**11.8**
1	-0.4	0.1	0.0	0.2	-0.7	-0.1
2	-0.3	0.1	0.0	0.2	-0.5	-0.1
3	-0.3	0.1	0.0	0.1	-0.4	-0.1
4	-0.2	0.0	0.0	0.1	-0.3	-0.1
5	-0.2	0.0	0.0	0.1	-0.2	0.0
6	-0.2	-0.2	0.0	0.0	0.3	0.1
7	-0.1	-0.1	0.0	0.0	0.1	0.0
8	-0.1	0.1	0.0	0.0	0.0	0.0
9	0.0	0.0	0.0	0.0	0.0	0.0
10	-0.3	-0.5	0.0	0.0	0.9	0.2
11	-0.2	-0.2	0.0	0.0	0.3	0.1
12	-0.6	0.2	0.1	0.2	-0.7	-0.1
13	-0.5	0.2	0.1	0.2	-0.6	-0.1
14	-0.4	0.1	0.0	0.1	-0.4	-0.1
15	-0.3	0.1	0.0	0.1	-0.3	-0.1
16	-0.2	0.1	0.0	0.1	-0.2	-0.1
17	-0.6	-0.1	0.1	0.0	0.4	0.1
18	-0.3	-0.1	0.0	0.0	0.2	0.0
19	-0.3	0.3	0.0	0.1	-0.5	-0.2
20	-0.4	0.6	0.1	0.2	-0.8	-0.3
21	-0.2	0.0	0.0	0.1	-0.2	0.0
22	-0.3	0.0	0.0	0.1	-0.4	-0.1
23	-0.4	0.1	0.0	0.2	-0.7	-0.1
24	-0.6	0.1	0.1	0.3	-1.2	-0.2
25	-0.1	0.0	0.0	0.1	-0.2	0.0
26	-0.3	-0.1	0.0	0.1	-0.4	0.0
27	-0.5	-0.5	0.1	0.3	-0.7	0.0
Mean SE	**0.1**	**0.2**	**0.0**	**0.1**	**0.1**	**0.1**

* 0.0 indicates a value of less than 0.05
SE = standard error of the difference
Source: January – December 2012 American Community Survey Sample Panels (March 2012 through February 2013 Computer Assisted Personal Interviewing)

Table 12 summarizes the differences between the completed interviews under the current full CAPI and the completed CAPI interviews by stopping rule for eight social characteristics. For the first two

education characteristics and the percent civilian veterans, percent with different residence 1 year ago, and percent non-citizens, the mean differences were 0.1 with little variability across rules. For the percent foreign-born and the percent speaking a language other than English at home, the mean difference was also 0.1 percentage point with a wider range of differences. The percent with health insurance was very similar across rules and had a mean difference of less than 0.05 percentage points.

Table 12. Differences in the 2012 Characteristics of the CAPI Respondent Populations - Social Characteristics

	Difference from Full CAPI							
Rule	Percent in College or Graduate School	Percent with Graduate or Prof. Degree	Percent Civilian Veterans	Percent with Different Residence 1 Year Ago	Percent non-citizens	Percent Foreign-born	Percent Speaking language other than English at home	Percent with Health Insurance
Full CAPI	**23.9**	**7.8**	**7.0**	**18.9**	**10.1**	**16.1**	**72.7**	**79.7**
1	-0.1	-0.1	0.1	-0.2	0.0	0.1	-0.1	0.0
2	-0.1	-0.1	0.1	-0.2	0.0	0.1	-0.1	0.0
3	-0.1	-0.1	0.1	-0.1	0.0	0.1	-0.1	0.0
4	0.0	0.0	0.0	-0.1	0.0	0.0	-0.1	0.0
5	0.0	0.0	0.0	-0.1	0.0	0.0	-0.1	0.0
6	0.2	0.0	0.1	0.1	-0.1	-0.2	0.3	0.1
7	0.1	0.0	0.0	0.0	0.0	-0.1	0.1	0.0
8	0.1	0.0	0.0	0.1	0.1	0.0	0.0	0.0
9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
10	0.5	0.0	0.2	0.1	-0.2	-0.4	0.6	0.1
11	0.2	0.0	0.1	0.0	-0.1	-0.1	0.3	0.1
12	0.0	-0.2	0.2	0.0	0.1	0.1	-0.2	-0.1
13	0.0	-0.1	0.1	-0.1	0.1	0.1	-0.1	0.0
14	0.0	-0.1	0.1	-0.1	0.1	0.1	-0.1	0.0
15	0.0	-0.1	0.1	-0.1	0.0	0.1	-0.1	0.0
16	0.0	-0.1	0.0	0.0	0.0	0.1	-0.1	0.0
17	0.5	-0.1	0.2	0.4	0.0	-0.2	0.3	-0.1
18	0.2	0.0	0.1	0.1	0.0	-0.1	0.2	0.0
19	0.0	-0.1	0.0	-0.1	0.1	0.2	-0.4	0.0
20	0.0	-0.2	0.1	-0.2	0.4	0.5	-0.8	-0.1
21	0.0	0.0	0.0	-0.1	0.0	0.0	-0.1	0.0
22	-0.1	-0.1	0.1	-0.1	0.0	0.0	-0.1	0.0
23	-0.1	-0.1	0.1	-0.2	0.1	0.1	-0.1	0.0
24	-0.1	-0.3	0.2	-0.3	0.1	0.1	-0.2	0.0
25	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
26	-0.1	-0.1	0.1	0.0	-0.1	-0.2	0.3	0.0
27	-0.3	-0.2	0.3	-0.2	-0.4	-0.6	0.8	0.0
Mean SE	**0.2**	**0.1**	**0.1**	**0.1**	**0.1**	**0.1**	**0.2**	**0.1**

* 0.0 indicates a value of less than 0.05
SE = standard error of the difference
Source: January – December 2012 American Community Survey Sample Panels (March 2012 through February 2013 Computer Assisted Personal Interviewing)

Table 13 includes four economic characteristics. In this table, we see fewer potential differences than observed in Table 11. The mean difference in the percent of the civilian labor force unemployed under full versus truncated CAPI is less than 0.05 percentage points. The mean differences in the percent of workers commuting by car, truck, or van and the population with income and benefits of $15K to $25K are 0.1 percentage points. That mean is slightly higher for the percent with Food Stamp/SNAP benefits.

Table 13. Differences in the 2012 Characteristics of the CAPI Respondent Populations – Economic Characteristics

	Difference from Full CAPI			
Rule	Percent of civilian labor force unemployed	Percent workers commuting by car, truck, van	Percent with income and benefits of $15K - $25K	Percent with Food Stamp/SNAP benefits
Full CAPI	**7.0**	**85.8**	**12.6**	**18.7**
1	0.0	0.1	0.1	0.3
2	0.0	0.1	0.0	0.2
3	0.0	0.1	0.0	0.1
4	0.0	0.1	0.0	0.1
5	0.0	0.0	0.0	0.1
6	0.0	0.1	0.0	0.0
7	0.0	0.0	0.0	0.0
8	0.0	0.1	0.0	0.1
9	0.0	0.0	0.0	0.0
10	-0.1	0.2	0.1	-0.2
11	0.0	0.1	0.0	-0.1
12	0.1	0.1	0.2	0.4
13	0.0	0.1	0.1	0.3
14	0.0	0.1	0.1	0.2
15	0.0	0.0	0.0	0.1
16	0.0	0.0	0.0	0.1
17	0.0	0.1	0.2	0.2
18	0.0	0.1	0.1	0.0
19	0.0	-0.2	0.1	0.2
20	0.1	-0.4	0.2	0.5
21	0.0	0.0	0.0	0.0
22	0.0	0.1	0.0	0.1
23	0.0	0.0	0.1	0.3
24	0.1	0.0	0.2	0.6
25	0.0	0.1	0.0	0.0
26	0.0	0.3	0.0	0.1
27	0.0	0.7	0.1	0.3
Mean SE	**0.1**	**0.1**	**0.1**	**0.1**

* 0.0 indicates a value of less than 0.05
SE = standard error of the difference
Source: January – December 2012 American Community Survey Sample Panels (March 2012 through February 2013 Computer Assisted Personal Interviewing)

Table 14 summarizes results for six housing characteristics. Some differences emerge as noteworthy under some rules. For example, the differences in the percent of renter-occupied units ranges from less than 0.05 percentage points to 1.2 percentage points. The mean difference is 0.3 percentage points. The percent of housing units built before 1940 and the percent with one bedroom have mean differences of 0.1 percentage points. The percent with heating fuel from electricity and the percent with property values less than $50K have mean differences of 0.2 percentage points. Similar to renter-occupied units, the percent multi-unit structures shows greater variability across rules with two rules of about 1 percentage point or higher.

It is important to remember that these results are national-level estimates and that the absence of major differences does not imply that implementing these stopping rules would not lead to differences of concern at subnational levels or for certain population groups. It is reassuring that we do not see indications of substantive national-level differences.

Table 14 Differences in the 2012 Characteristics of the CAPI Respondent Populations – Housing Characteristics

	Difference from Full CAPI					
Rule	Percent Renter-occupied	Percent built before 1940	Percent with 1 bedroom	Percent with heating fuel from electricity	Percent with property value less than $50K	Percent multi-unit structures
Full CAPI	**50.5**	**14.7**	**13.3**	**42.9**	**11.6**	**34.2**
1	-0.6	-0.2	-0.2	0.1	0.2	-0.7
2	-0.5	-0.1	-0.2	0.1	0.2	-0.5
3	-0.3	-0.1	-0.1	0.0	0.1	-0.4
4	-0.3	-0.1	-0.1	0.0	0.1	-0.3
5	-0.2	-0.1	-0.1	0.0	0.1	-0.2
6	-0.1	0.0	0.1	0.2	0.2	0.0
7	0.0	0.0	0.0	0.1	0.0	0.0
8	0.1	0.0	0.0	0.1	0.1	-0.1
9	0.0	0.0	0.0	0.0	0.0	0.0
10	-0.1	0.0	0.2	0.3	0.3	0.0
11	-0.1	0.0	0.1	0.1	0.1	0.0
12	-0.3	-0.3	-0.1	0.4	0.5	-0.5
13	-0.3	-0.2	-0.1	0.2	0.3	-0.4
14	-0.2	-0.2	-0.1	0.2	0.2	-0.3
15	-0.2	-0.1	-0.1	0.1	0.2	-0.3
16	-0.2	-0.1	-0.1	0.1	0.1	-0.2
17	0.5	-0.1	0.3	0.8	0.7	0.3
18	0.1	-0.1	0.1	0.3	0.2	0.0
19	-0.1	0.0	-0.1	-0.1	0.2	-0.1
20	0.0	0.2	0.0	-0.3	0.3	0.0
21	-0.2	-0.1	-0.1	0.0	0.0	-0.3
22	-0.4	-0.1	-0.2	0.0	0.1	-0.5
23	-0.5	-0.1	-0.2	0.1	0.2	-0.6
24	-0.9	-0.2	-0.4	0.0	0.6	-1.1
25	-0.2	-0.1	-0.1	0.1	0.1	-0.3
26	-0.5	-0.3	-0.2	0.3	0.2	-0.7
27	-1.2	-0.6	-0.5	0.7	0.5	-1.5
Mean SE	**0.2**	**0.1**	**0.1**	**0.1**	**0.1**	**0.1**

* 0.0 indicates a value of less than 0.05
SE = standard error of the difference
Source: January – December 2012 American Community Survey Sample Panels (March 2012 through February 2013 Computer Assisted Personal Interviewing)

5.5 Representative Rules

Which rules meet our goal of reducing respondent burden without major losses in quality?

When we look at the results in Tables 6 – 9, we draw similar conclusions about which stopping rules are likely to reduce respondent burden and which rules appear to have limited value in reducing burden. Table 6 identifies eight rules with less than a 4 percent reduction in contacts with a sample household member (4, 5, 7, 9, 11, 16, 21, and 25). These same eight rules have less than an 8 percent reduction in firm reluctant outcomes and less than a 15 percent reduction in post-reluctance attempts. They also emerge as the eight rules with less than a 25 percent reduction in "very high burden" cases. Based on these results, we do not recommend implementing rules 4, 5, 7, 9, 11, 16, 21 or 25.

From Table 10 we find that rules 12, 17, 20, 24, and 27 all show CAPI response rates of 86.2 percent or less and at least a 9 percent reduction in completed CAPI interviews. The remaining representative rules have nonresponse rates of less than 11.8 percent. We conclude that rules 12, 17, 20, 24, and 27 have an

extreme effect on levels of nonresponse and do not recommend that they be considered for implementation at this time.

5.6 Cost, Burden, and Quality Trade-offs

What are the trade-offs in workloads, costs, burden, and quality for the rules identified in research question #5?

5.6.1 Cost/Benefit Ratios

Table 15 summarizes several cost/benefit ratios for the 14 rules that show noteworthy reductions in respondent burden while maintaining CAPI response rates of 85 percent or more. We display two measures that assess cost versus quality. The first is the reduction in the number of total CAPI hours for each lost interview. This is a measure of cost effectiveness of reductions in contact attempts. The second measures the number of eliminated contact attempts for each lost interview. Like the first measure, this is primarily a workload and cost measure and, for that reason, it includes both occupied and vacant housing units.

The other three measures are burden-focused and therefore only include occupied housing units. Each show the ratio of a specific burden measure (eliminated very high burden cases, eliminated contacts with household members, and eliminated post-reluctance attempts) relative to lost interviews.

Table 15. 2012 Cost-Benefit Measures

	Cost vs. Quality		**Burden vs. Quality**		
Rule	**Reduction in total CAPI hours per lost interview (occupied & vacant)**	**Eliminated contact attempts per lost interview (occupied & vacant)**	**Eliminated very high burden cases per lost interview (occupied only)**	**Eliminated contacts with HH member per lost interview (occupied only)**	**Eliminated post-reluctance attempts per lost interview (occupied only)**
1	3.2	5.1	0.4	1.5	2.4
2	3.4	5.5	0.5	1.5	2.6
3	**3.5**	**5.8**	0.5	1.5	2.9
6	2.5	4.0	**0.6**	**1.8**	**4.0**
8	3.1	4.8	**0.7**	**2.0**	**4.9**
10	2.2	3.6	0.4	**1.7**	3.0
13	3.1	5.0	0.5	1.5	2.3
14	3.3	5.4	0.5	1.5	2.6
15	**3.5**	**5.7**	0.6	1.6	2.8
18	3.1	5.0	**1.0**	**1.8**	**4.1**
19	3.2	5.2	0.4	1.5	2.4
22	3.2	5.3	0.4	1.5	2.5
23	3.2	5.1	0.4	1.5	2.4
26	**3.6**	**5.6**	0.5	**1.7**	2.6

The bolded rules identify some of the best performing rules under each of these cost/benefit measures.
Source: January – December 2012 American Community Survey Sample Panels (March 2012 through February 2013 Computer Assisted Personal Interviewing)

All of the stopping rules, by definition, stop data collection efforts earlier than our current methods. For this reason, they all result in fewer completed interviews. It is these "lost interviews" that we want to study. The ratios in Table 15 summarize which stopping rules provide the greatest reductions in burden and reductions in CAPI interviewing hours for every lost interview. We see that rule 26 (which uses propensity scores) does best in eliminating total CAPI hours (this is likely due to its success in eliminating attempts that do not lead to interviews). Rules 3 and 15 also have high cost reduction ratios. It makes sense that rules 3, 15, and 26 also perform the best in eliminating total contact attempts relative to lost interviews. Each eliminated 5.6 or more attempts for every lost interview. In contrast, rules 6 and 10 appear to be less effective in reducing CAPI hours and total contact attempts for each lost interview.

Rules 6, 8, 15 and 18 eliminated the greatest number of high burden cases for every lost interview (0.6 or more). When we look at the potential reduction in contacts with household members, we see that rules 6, 8, 10, 18, and 26 eliminate 1.7 contacts or more per lost interview. If we measure burden in terms of contact attempts made after a household member expresses reluctance, we find that rules 6, 8, and 18 eliminate four or more post-reluctance attempts for every lost interview. All of the rules summarized in this table do a good job in reducing burden, regardless of the measure used.

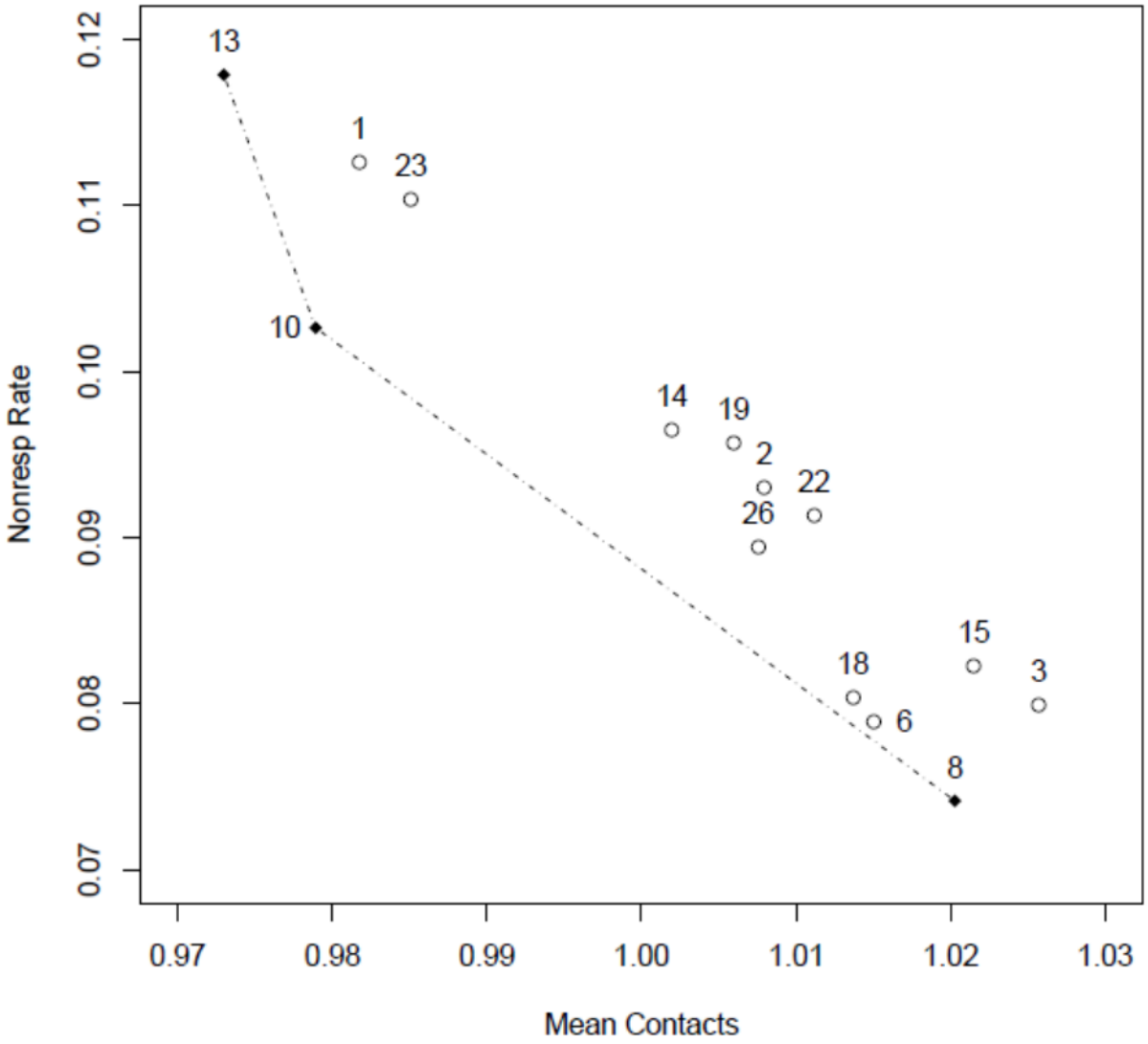
Based on these ratios, we find that rules 3, 6, 8, 10, 15, 18 and 26 appear to provide the best benefits in cost savings or burden reduction for the cost of a lost interview. The relative importance of these two benefits might suggest one rule over another.

5.6.2 Visual Summaries

In addition to these ratios, we analyzed quality and burden trade-offs using several graphical summaries. This report includes two of the figures that we considered. Each plot displays the estimated final nonresponse rate against a single measure of burden for the 14 rules that we determined met the basic requirements of reducing burden while maintaining quality.

Figure 1 displays the estimated final nonresponse rate and the mean number of contacts with sample households for the reduced set of representative rules. Rules 13, 10 and 8 define the lower boundary – with rules 18 and 6 extremely close to it. This suggests that these rules, and probability weighted mixtures of them, achieve the highest response rates for each specific level of burden reduction. For example, rule 26 reduces mean contacts to about the same level as rule 2, but rule 18 has the lower nonresponse rate. Similarly, rules 8 and 15 share the same estimated mean contacts but rule 8 has a lower associated level of nonresponse. Figure 1 allows us to see, for a specific level of nonresponse, which rule requires the least burden (as measured by contacts). For example, rules 6 and 3 have similar nonresponse rates with rule 6 showing lower burden (lower mean contacts).

Figure 1. Comparison by Stopping Rule of Nonresponse Rate and Mean Contacts with Sample Household

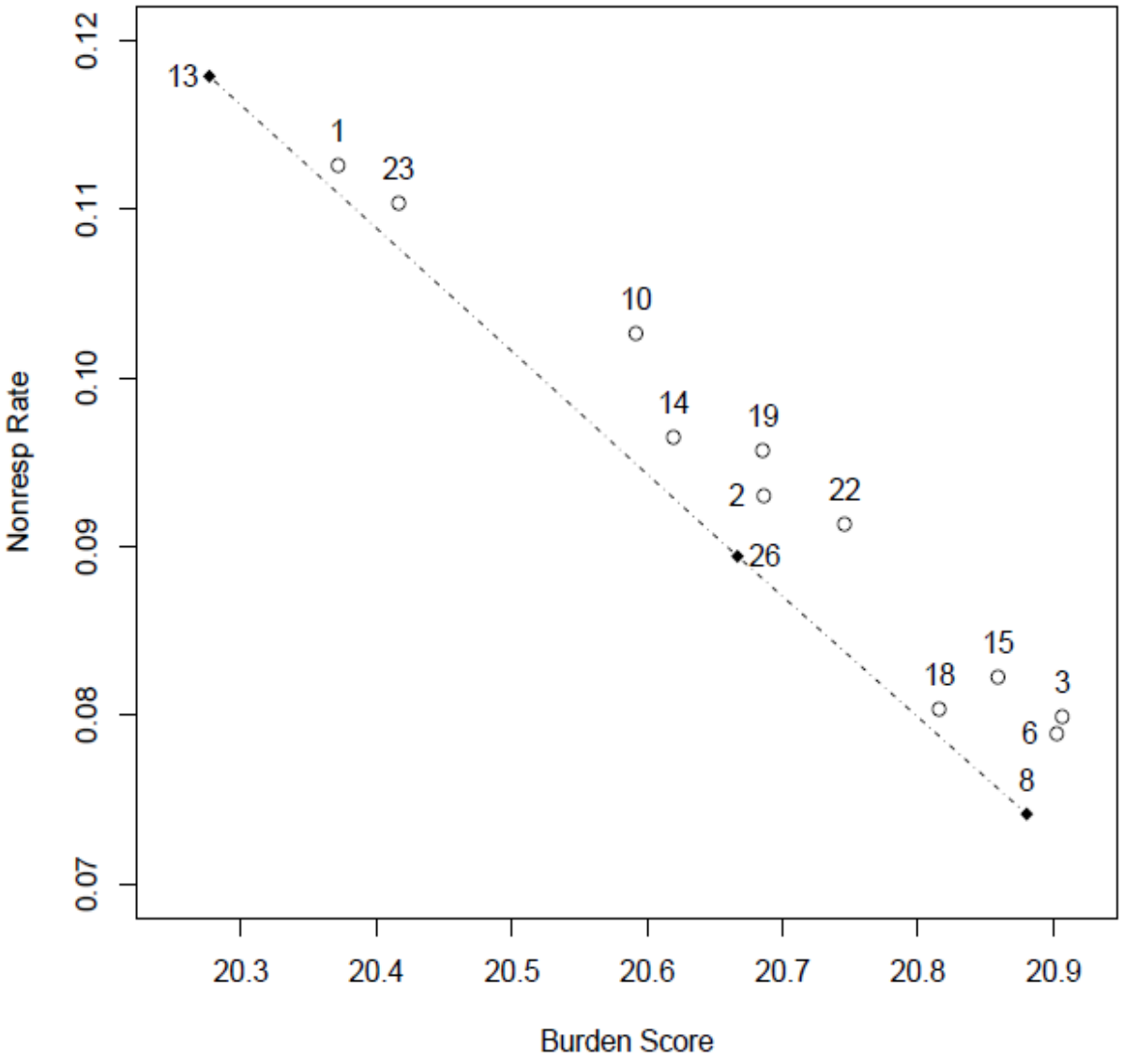


Source: January – December 2012 American Community Survey Sample Panels (March 2012 through February 2013 Computer Assisted Personal Interviewing)

Figure 2 graphs the nonresponse rate and the mean cumulative burden score by stopping rule. Rules 13, 26, and 8 define the lower boundary, which is essentially a straight line, describing the trade-off under these "best" rules as an additional 1 percent of nonresponse being equivalent to a reduction of 0.138 in per-case cumulative burden score. As in Figure 1, for a given level of cumulative burden, these lower-boundary rules have the lowest nonresponse rates. For example, rules 2, 19, and 26 have similar expected levels of cumulative burden with rule 26 having the lowest nonresponse rate. Rules 18 and 3 also have similar levels of nonresponse with 18 having the lower level of cumulative burden.

These two visual summaries suggest that, from a burden perspective (mean contacts and mean cumulative burden), rules 6, 8, 13, 18, and 26 perform relatively better than the other rules.

Figure 2. Comparison by Stopping Rule of Nonresponse Rate and Cumulative Burden Score



Source: January – December 2012 American Community Survey Sample Panels (March 2012 through February 2013 Computer Assisted Personal Interviewing)

6. CONCLUSIONS

Our analysis of respondent burden metrics for 27 CAPI stopping rules identified 14 rules that we believe reduce respondent burden in a meaningful manner without major increases in nonresponse. When we consider which of these stopping rules optimize reductions in burden while maintaining the highest levels of CAPI survey quality (Figures 1 and 2), we find rules 6, 8, 13, 18 and 26 worthy of additional analysis, with 10 also a possibility from Figure 1. If we look at the cost effectiveness of these six rules (Table 15), we note that rule 10 shows relatively low reductions in hours and total attempts per lost interview.

We selected rules 6, 8, 13, 18, and 26 as the rules that best balance burden reductions and cost savings with losses in total interviews.

- Rule 6 stops after the second contact attempt with any form of reluctance expressed by the sample household.
- Rule 8 stops after the first expression of firm reluctance.
- Rule 13 stops after four attempts if the case is a CATI refusal, after five attempts if it is another CATI noninterview, and after seven attempts for all other cases.
- Rule 18 uses the cumulative burden score, stopping when the case reaches a score of 40.
- Rule 26 stops when the propensity score falls below 20 percent.

Table 16 summarizes several estimates of burden and cost reduction and quality losses associated with these five rules. Table 16 also includes one aggregate measure of nonresponse bias – the mean absolute percent error. We calculated this as the sum of the absolute values of the percent differences between the full CAPI and the truncated CAPI across all 24 characteristics. We interpret it as a measure of the average percent difference in the characteristics of the completed interviews under these stopping rules. Appendix B provides detail on this measure.

Four of these rules result in about a 3 to 4 percent drop in the number of completed interviews and a final CAPI response rate of about 91 to 92 percent. Rule 13 brings the response rate to about 88 percent, reducing the CAPI completed interviews by about 7 percent. The mean absolute percent error measures are low across all rules. As expected, the rules with the greatest reductions in response show the highest mean absolute percent error measures.

Each of these rules holds promise to reduce burden in a meaningful way. From Table 16 you can see that because we defined some rules using certain burden metrics, they perform very well when we measure burden reduction by that metric. Rule 26 is very effective in reducing non-contacts given the use of propensity scores to identify cases for additional follow up. Rule 13 is similar in some ways as it does well in reducing total CAPI hours and total CAPI contact attempts. Rule 18 is very effective in eliminating very high burden cases because it uses the burden score to determine when to discontinue contact efforts. Rule 8 essentially eliminates all post-firm reluctance attempts due to its definition of stopping after any expression of firm reluctance.

Table 16. Summary of Results

Cost, Burden and Quality Measures	Rule 6	Rule 8	Rule 13	Rule 18	Rule 26
Percent reduction in completed interviews	3.1	2.6	7.2	3.2	4.2
Final CAPI response rate	92.1	92.6	88.2	92.0	91.1
Pct. reduction, very high burden scores in occupied cases	32.5	30.8	55.5	62.4	33.4
Pct. reduction, occupied post-firm reluctance attempts	43.3	97.5	41.6	50.2	38.4
Pct. reduction, occupied attempts with reluctance	9.9	9.8	14.1	10.0	9.7
Pct. reduction, contacts with occupied sample households	5.3	4.7	9.2	5.4	5.9
Percent reduction in total CAPI attempts	3.9	3.9	11.7	5.2	7.6
Percent reduction in non-contacts	4.6	4.9	16.5	7.0	11.0
Percent reduction in total CAPI hours	3.4	3.5	10.0	4.4	6.6
Mean absolute percent error	0.6	0.3	1.0	0.5	1.0

Source: January – December 2012 American Community Survey Sample Panels (March 2012 through February 2013 Computer Assisted Personal Interviewing)

7. NEXT STEPS

Staff needs to discuss the implementation issues associated with each of these stopping rules and determine the additional analysis that we should undertake before implementation. We recommend subnational analysis of results to identify if any of these rules have differential effects on specific areas or populations. In particular, we suggest the production of summaries, similar to those in Table 16, at regional office, state, and field supervisory levels. Tabulating expected results by field performance strata, American Indian and Alaska Native Areas, and 2010 census segmentation groups would provide additional insights. It is also useful for us to consider minor variations on these rules that may warrant further analysis; for example, rules with alternative thresholds. We see these five rules as representative examples of types of rules that may well benefit from refinements.

The lost interviews associated with implementing these stopping rules will reduce the reliability of sample estimates. Additional research should determine if we should consider any changes in the CAPI sampling rates to offset these losses.

8. ACKNOWLEDGMENTS

The authors acknowledge the huge contributions to this multi-phase project of the other members of the interdivisional "CAPI Analysis" team – Cavan Capps, Todd Hughes, Pad Murphy, Dawn Nelson, Barbara O'Hare, and Mary Frances Zelenak. In addition, we received important comments from the critical reviewers assigned to this project, specifically, Tony Tersine, Gianna Dusch, and Beth Tyszka.

9. REFERENCES

Bradburn, N. 1978. Respondent Burden. *Health Survey Research Methods*, DHEW Publication No. 79-3207, 49-53.

Ferguson, T. 1967. Mathematical Statistics: A Decision Theoretic Approach. Academic Press: San Francisco.

Fisher, J. & Kydoniefs, L. 2001. Using a Theoretical Model of Respondent Burden to Identify Sources of Burden in Surveys. Paper presented at the 12th International Workshop on Household Survey Nonresponse, Oslo.

Frankel, J. & Sharp, L. 1981. Measurement of Respondent Burden. Statistical Reporter. January 1981.

Giesen, D., Bavdaz, M. & Haraldsen, G. 2011. Response Burden Measurement: Current Diversity and Proposal for Moving Towards Standardization. BLUE-ETS Conference on Burden and Motivation in Official Business Surveys.

Gravem, D., Haralsen, G., & Lofgren, T. 2011. Response Burden Trends and Consequences. BLUE-ETS Conference on Burden and Motivation in Official Business Surveys.

Griffin, D. & Hughes, T. 2013. Analysis of Alternative Call Parameters in the American Community Survey's Computer Assisted Telephone Interviewing. ACS13-RER-11. http://www.census.gov/acs/www/Downloads/library/2013/2013_Griffin_02.pdf

Griffin, D. 2014. Reducing Respondent Burden in the American Community Survey's Computer Assisted Personal Visit Interviewing Operation – Phase 2 Results. 2014 American Community Survey Research And Evaluation Report Memorandum Series #ACS14-RER-07. http://www.census.gov/acs/www/Downloads/library/2014/2014_Griffin_01.pdf

Griffin, D. & Nelson, D. 2014. Reducing Respondent Burden in the American Community Survey's Computer Assisted Personal Visit Interviewing Operation – Phase 1 Results (Part 2). ACS14-RER-22. http://www.census.gov/acs/www/Downloads/library/2014/2014_Griffin_02.pdf

Jones, J. 2012. Response Burden: Introductory Overview Lecture. Presented at the 2012 Joint Statistical meetings.

Joshipura, M. 2008. 2005 ACS Respondent Characteristics Evaluation. DSSD ACS Research and Evaluation Memorandum Series Chapter #ACS-RE-2http://www.census.gov/acs/www/Downloads/library/2008/2008_Joshipura_01.pdf

King, K. 2012. Meeting the Statistical Quality Standards for the 2011 American Community Survey Data Products. DSSD 2011 American Community Survey Memorandum Series - #ACS11-QM-07

Lawrence, J., Adams, T. & O'Hare, B. (forthcoming). CAPI Case Cost Estimation Model. Office of Survey Analytics Memorandum Series #2014-03.

McCarthy, J. 2011. Using Predictive Models with Measures of Burden Reporting History, and External Data to Identify Likely Establishment Survey Nonrespondents. BLUE-ETS Conference on Burden and Motivation in Official Business Surveys.

Poe, T. 2011. "The Pros and Cons of Making the American Community Survey Voluntary." Prepared Testimony of Ted Poe (TX-02) Before the House Oversight and Government Reform Committee's Subcommittee on Health Care, District of Columbia, Census, and National Archives. March 6, 2011.

Sears, J. 2011. Response Burden Measurement and Motivation at Statistics Canada. 2011. BLUE-ETS Conference on Burden and Motivation in Official Business Surveys.

Zelenak, M.F. & 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. http://www.census.gov/acs/www/Downloads/library/2013/2013_Zelenak_01.pdf

Zelenak, M.F. (2014). Reducing Respondent Burden in the American Community Survey's Computer Assisted Personal Visit Interviewing Operation – Phase 1 Results (Part 1). 2014 American Community Survey Research And Evaluation Report Memorandum Series #ACS14-RER-06. http://www.census.gov/acs/www/Downloads/library/2014/2014_Zelenak_01.pdf

APPENDIX A

CUMULATIVE BURDEN SCORING

The basic idea behind this method is to tally each contact attempt (in any mode) as a separate increment of burden. We assign a score based on the relative burden of the various contact attempts. The stopping rule involves a threshold of cumulative burden and when a case reaches or passes that threshold, we stop CAPI contact attempts. This option starts by establishing a set of "incoming burden scores" based on the mail/CATI status as follows.

For mail attempts, a case receives four burden points for being sent the initial mailing, one additional burden point for being sent the second mailing, and an additional point for being sent the third mailing. In addition to any mail burden points, cases that went to CATI receive twenty points for a final CATI status of refusal, and twelve points for reaching the call maximum or resulting in a non-interview in CATI for some other reason.

The incoming burden score for each case is equal to the sum of its CATI and mail burden. For example, a case that received all three mailings and reached the call maximum in CATI has an incoming burden score of 18. The rule then increments the burden for every CAPI contact attempt as follows.

* A CAPI contact attempt by phone or in-person that makes no contact and is likely invisible to the respondent (e.g., a drive-by or a ring-no-answer) has an added score of 1.

* A CAPI contact attempt that makes no contact and is visible to the respondent (e.g., a message left on a machine, a letter sent, a note left on a door) has an added score of 2 if by phone or 3 if in-person.

* A CAPI contact attempt that makes contact with no reluctance expressed has an incremental score of 4 if by phone or 6 if in-person.

* A CAPI contact attempt that expresses reluctance (but not firm) has an added score of 8 if contacted by phone or 10 if contacted in-person.

* A CAPI contact attempt that expresses firm reluctance has an added score of 12 if the attempt was by phone or 15 if it was in-person.

We could establish a stopping rule based on several different cumulative burden values. For example, if our stopping rule was set at 30, we could have the following outcomes. A case that received all three mailings and reached the call maximum in CATI has an incoming burden score of 18. If the next contact attempt for this case was in-person and resulted in reluctance (but not strong), the new score would be 28 and we would continue. If the next contact attempt for this case was in-person and resulted in a strong refusal, the new score would reach 33 and we would stop.

APPENDIX B

CALCULATION OF MEAN ABSOLUTE PERCENT ERRORS

To compare the nonresponse bias that we might introduce in our national-level estimates, we chose to calculate a Mean Absolute Percent Error (MAPE) for the five stopping rules that we thought most viable to pursue. To calculate each MAPE we used the results in Tables 11 through 14 to estimate the percent difference between the full CAPI and the truncated CAPI for each of the 24 characteristics. Table 17 summarizes these percent differences by rule and characteristic. The MAPE is the sum of the absolute value of these differences divided by 24.

Table 17. Percent Differences and MAPEs by Stopping Rule

Characteristic	Rule 6	Rule 8	Rule 13	Rule 18	Rule 26
DEMOGRAPHIC					
Percent of the total population reporting a race of Black or African-American alone	-1.3	-0.8	-2.9	-1.9	-2.1
Percent of the total population reporting as Hispanic	-0.9	0.5	0.6	-0.2	-0.5
Percent of the total population under 3 years	0.1	0.4	1.1	0.6	0.5
SOCIAL					
Percent of the total population with a relationship of spouse	0.1	0.2	1.2	0.1	0.9
Percent of households that are living alone	1.3	0.1	-2.1	0.6	-1.6
Percent of the female population 15 years and over that are divorced	0.5	0.0	-1.0	0.4	0.2
Percent of the population 25 years and over enrolled in college or graduate school	0.8	0.2	0.1	0.8	-0.6
Percent of the population 25 years and over that have a graduate or professional degree	0.3	0.2	-1.8	-0.1	-1.2
Percent of the population 18 years and over that are civilian veteran	1.0	0.3	1.6	1.1	1.6
Percent of the population 1 year and over with a residence 1 year ago of a different house in the U.S.	0.4	0.5	-0.3	0.7	0.2
Percent of the total population that are non-citizens	-0.8	0.7	0.9	-0.1	-1.1
Percent of the total population that are foreign born	-1.1	0.2	0.5	-0.6	-1.4
Percent of the population 5 years and over that speak a language other than English at home	0.4	0.1	-0.2	0.2	0.4
Percent of the total population with health insurance	0.1	0.0	0.1	0.0	0.0
ECONOMIC					
Percent of the civilian labor force that is unemployed	0.6	0.3	0.5	-0.2	0.0
Percent of workers 16 years and over commuting to work by car, truck, or van	0.2	0.1	0.1	0.1	0.4
Percent of total households (occupied housing units) with income and benefits of $15K - $25K	0.4	0.2	0.8	0.5	0.2
Percent of total households (occupied housing units) with food stamp/SNAP benefits in the past 12 months	0.2	0.3	1.7	0.3	0.8
HOUSING					
Percent of occupied housing units that are renter-occupied	-0.1	0.1	-0.6	0.2	-1.0
Percent of total housing units built before 1940	0.2	0.2	-1.2	-0.5	-2.0
Percent of total housing units with 1 bedroom	0.7	0.0	-0.9	0.7	-1.8
Percent of occupied housing units with house heating fuel from electricity	0.4	0.3	0.6	0.6	0.7
Percent of owner-occupied housing units with a property value of less than $50K	1.6	1.1	2.9	1.9	2.0
Percent of occupied housing units that are in multi-unit structures	0.1	-0.2	-1.2	0.1	-2.0
MAPE	**0.6**	**0.3**	**1.0**	**0.6**	**1.0**

Source: January – December 2012 ACS Panels (March 2012 through February 2013 CAPI)