Changing the Way We Look at Survey Nonresponse

Deborah H. Griffin and Dawn V. Nelson

U.S. Census Bureau
Washington, DC 20233

Proceedings of the 2013 Federal Committee on Statistical Methodology (FCSM) Research Conference

Abstract

For decades, survey sponsors, data users, and other stakeholders have judged the success of a survey by its response rate. The Office of Management and Budget and the Department of Commerce cite this measure of nonresponse as a key performance metric. The Census Bureau's Statistical Quality Standards require that a survey achieve certain levels of response to demonstrate acceptable quality. Because of the value placed on response rates, they have become a major measure used to evaluate data collection operations. We evaluate the performance of telephone interviewers, field representatives, and their managers based largely on the response rates they achieve. This focus on response rates has many consequences. One is a high burden placed on respondents when a survey repeatedly contacts sample households. Another is the pressure on interviewing staff to spend time and resources to increase their response rates and not to accept a noninterview. The reliance on response rates has overshadowed the value of alternative measures that may be equally, or even more, important from a survey management perspective. In the past few years, survey methodologists have questioned the utility of the response rate as a primary measure of survey quality. There is also growing concern about the costs associated with obtaining high response rates.

This paper summarizes historical measures of survey nonresponse in the American Community Survey (ACS). It shows that the attention given to response rates in the ACS has resulted in annual weighted survey response rates between 97 and 98 percent with consistently high response rates achieved across all areas of the country. If controlling for survey nonresponse (as measured by consistently high subnational response rates) is the most important metric, the ACS is very successful. This paper looks more closely at ACS nonresponse, specifically at nonresponse associated with the final ACS data collection mode, Computer Assisted Personal Interviewing (CAPI). Here too we find exceptionally high response rates of about 95 percent with little subnational variation. The Census Bureau employs a contact history instrument during CAPI data collection. Paradata from this instrument allow us to look below the surface at what is required, in terms of both respondent burden and costs, to obtain these high response rates. We use these paradata to propose and demonstrate the utility of a set of alternative metrics to assess quality and productivity. We also suggest new measures of efficiency and quality that hold promise to tell us more than response rates ever could.

Key words: quality, response rates, respondent burden, efficiency

1 Introduction

The survey world is dealing with many realities. Reductions in respondent cooperation and increases in the costs of doing business means that maintaining high response rates requires increasing levels of effort and greater resources. In addition, public complaints about survey requests from the government have risen. These criticisms come to the Census Bureau in the form of respondent letters and calls about the intrusive and burdensome nature of the survey questions. Congressional testimonies have criticized the American Community Survey for its pursuit of a response through multiple contacts (Poe 2011).

With flat data collection budgets and depressed levels of respondent cooperation, some have argued that our current methods are not sustainable (Groves 2011). This paper offers suggestions for changes in the way that managers and sponsors judge survey success by taking into account a broader set of metrics than survey nonresponse.

2 Background

2.1 Response Rates

For decades, survey sponsors, data users, and other stakeholders have judged the success of a survey by its response rate. The Office of Management and Budget and the Department of Commerce have historically cited this measure of nonresponse as a key performance metric (OMB 2006; U.S. Census Bureau 2013a). The Census Bureau's statistical quality standards require that a survey conduct nonresponse bias analyses when unit response rates fall below 80 percent and define "serious data quality issues" when unit response rates are below 60 percent (U.S. Census Bureau, 2013b). In turn, response rates have become the major measure used to evaluate data collection operations. We assess the performance of telephone interviewers, field representatives and their managers based largely on the response rates they achieve.

This focus on response rates has many consequences. One is an increased burden placed on survey respondents when a survey repeatedly contacts them. In a sequential mixed mode survey, this can mean receipt of multiple mail materials, numerous contact attempts by phone and in-person visits. In telephone surveys, double-digit call attempts are common in an effort to increase response rates. Relentless pursuit of a respondent through repeated personal visit attempts is not unusual in some surveys.

Another consequence is the pressure on interviewing staff to spend time and money to increase their response rates and not to accept a noninterview. Performance reports track response rates and "worst case" response rates at multiple levels, identifying regional offices as well as individuals that fail to achieve a desirable level of response. Managers reward interviewers and field supervisors for high response rates. This pressure to obtain high response rates can result in interviewer error and even falsification. It also encourages interviewer behavior that a household respondent might perceive as stalking (Poe 2011).

The difference in the characteristics of respondents and nonrespondents must accompany response rates to understand nonresponse bias and the impact of a given response rate on the accuracy of survey estimates. Survey methodologists question the utility of the response rate as a measure of nonresponse bias (Groves 2006; Groves & Peytcheva 2008). We know that low levels of nonresponse cannot ensure high quality. High response rates can even mask measurement error, coverage error, and item nonresponse problems. Biemer (2003) notes that "there has been much focus on response rates in surveys, and some surveys with low response rates have been discontinued on the basis of the response rate without evaluating the bias."

In addition, response rates are not a measure of efficiency. They do not account for the costs associated with these responses. In many surveys the goal is to obtain "as high as possible" response rates. Data collection managers who think they have done a good job by simply obtaining a high response rate may be wrong. Response rates do not tell the full story. Managers need to understand the costs and benefits associated with the response rates they achieve. What level of effort and respondent burden are associated with achieving those rates and do these extra efforts improve the quality of survey estimates? More specifically, managers need to know if reducing nonresponse was worth the cost.

2.2 Paradata

Falling response rates in the early to mid-1990's resulted in statistical agency heads creating the Inter-Agency Household Survey Nonresponse Group (IHSNG) to better understand nonresponse in major U.S. government surveys. The IHSNG included staff from a broad set of federal statistical agencies. In 2002, the IHSNG and the Census Bureau sponsored a Response Rate Summit to provide a forum for discussion among experts in the field about how to address concerns related to the decreasing response rate trend in household surveys. The top recommendation was to collect and analyze detailed contact/call record data and to use these paradata for analytic purposes and as a feedback mechanism for regional offices and field representatives (Salvucci, S. et al 2002). This type of information could help managers understand reasons for nonresponse and where certain interventions might be appropriate. The participants proposed an automated system to capture case histories including information such as number of contacts, interim outcomes, reasons for refusals, date and time of contact attempt, and demographics of refusal and contact households.

In response to this recommendation, the U.S. Census Bureau developed the Contact History Instrument (CHI). Bates (2003) describes the CHI as a single survey instrument designed for use in all surveys that contract with the Census Bureau for data collection. Field representatives record information into the CHI each time they make a contact attempt and the CHI automatically records the date and time of the entry, the case ID, and the duration that the CHI application was open. The CHI prompts field representatives to answer a series of questions describing the contact experience and outcome.

The first field applications of the CHI took place in 2004 with the National Health Interview Survey and in 2005 with the Consumer Expenditures Interview Survey. Since then, most Census Bureau demographic surveys incorporated the CHI into their CAPI instruments. Tan (2011) provides an excellent overview of the CHI with examples of how the Consumer Expenditures Survey uses CHI data. The CHI is a key source of information that allows survey managers to better understand the successes and challenges associated with personal visit data collection. New management approaches, as well as evaluations, benefit from the availability of CHI data. The American Community Survey (ACS) began collecting CHI data in 2011 and is beginning to take advantage of this information.

2.3 American Community Survey

The ACS is an on-going survey that provides information to communities (large and small) to aid them in planning investments and services. Information from the ACS generates estimates that help to determine how federal and state governments distribute more than $400 billion in funds each year. The data collected in the ACS are a part of the decennial census and the Census Bureau, prior to 2005, collected these data in combination with the decennial head count once every ten years. Given the rapid demographic, social, and economic changes experienced in the past decades and the strong expectation that such changes will continue and accelerate, this once-a-decade detailed data collection was recognized as unacceptable for producing the data required by federal, state, local, and tribal governments. To meet these growing needs, the Census Bureau designed the ACS which, unlike other surveys, must produce statistics for very small communities and population groups.

The ACS surveys the population in the United States and Puerto Rico, collecting demographic, social, and economic data from the population living in housing units (e.g., houses, apartments, mobile homes) and group quarters facilities (e.g., prisons, skilled nursing facilities, college dormitories). In addition, the ACS collects data on the physical and financial characteristics of the places where people live. For the population living in housing units (the focus of this paper), we select an annual sample of about 3.54 million housing unit addresses and distribute them into 12 monthly sample panels for data collection. We use four data collection modes over a 3-month time period to interview each sample panel.

Figure 1 summarizes this sequential design with rows representing the calendar months of each operation and columns defining each 3-month sample panel. Referring to the first column for January 2013, the Internet and mail modes are first, with follow-ups conducted initially by telephone and ultimately (for a subsample) in person. We accept Internet and mail responses throughout the 3-month data collection period. This data collection design means that all modes are active each month across different sample panels. U.S. Census Bureau (2010) provides details on the mixed mode data collection methods used prior to 2013. The ACS added the Internet response option in 2013. See Matthews et al (2012) for information about the development and testing associated with the new Internet mode.

CALENDAR MONTH	SAMPLE PANELS			
	January 2013	February 2013	March 2013	April 2013
January 2013	Internet/Mail			
February 2013	Telephone Internet/Mail (cont'd)	Internet/Mail		
March 2013	Personal Visit Internet/Mail (cont'd)	Telephone Internet/Mail (cont'd)	Internet/Mail	
April 2013		Personal Visit Internet/Mail (cont'd)	Telephone Internet/Mail (cont'd)	Internet/Mail
May 2013			Personal Visit Internet/Mail (cont'd)	Telephone Internet/Mail (cont'd)

Figure 1. 2013 ACS Data Collection

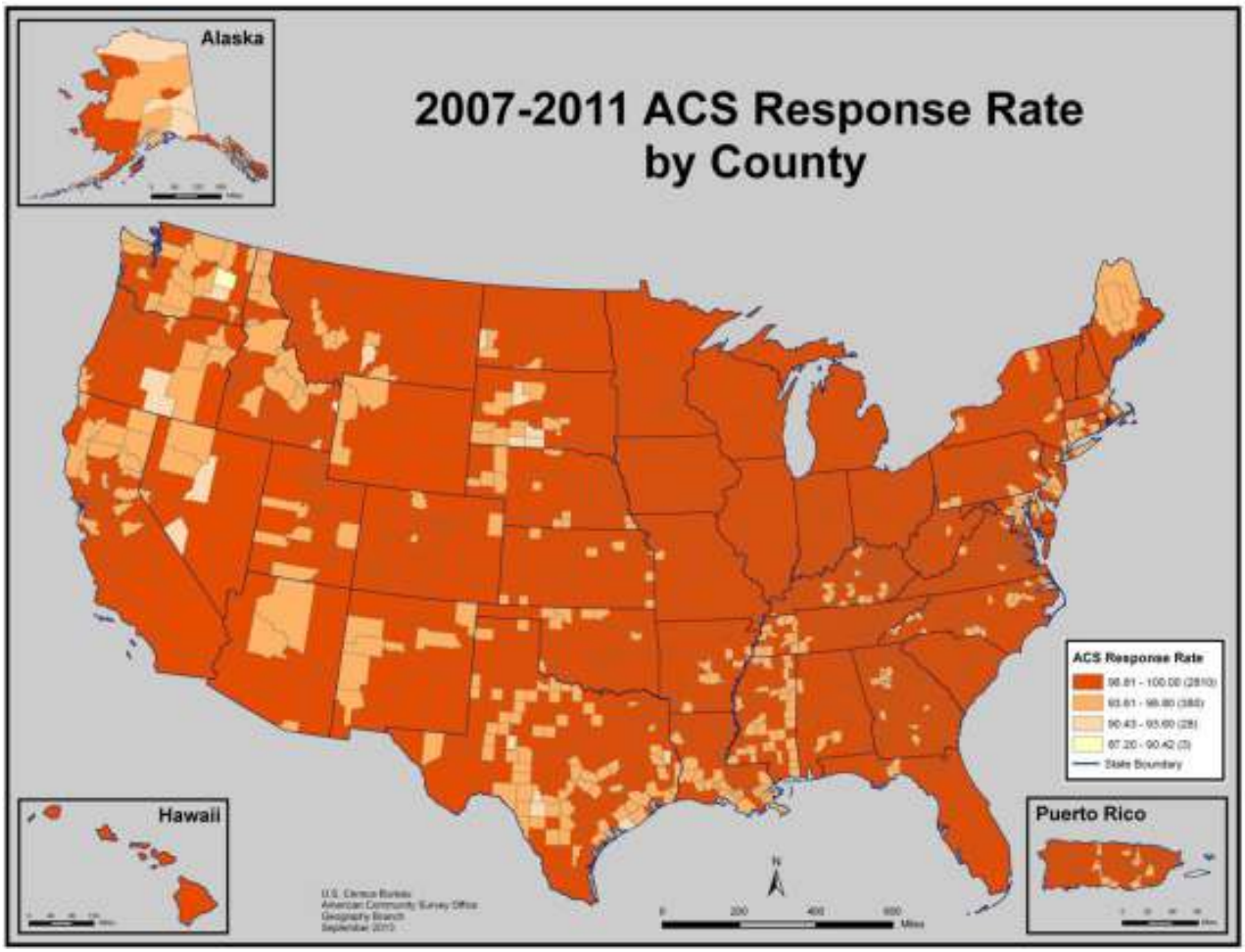
2.4 Monitoring ACS Performance

Each year, we combine sample interviews across months and modes to produce survey estimates. The Census Bureau summarizes ACS data in the form of 1-year, 3-year, and 5-year aggregations, permitting increasing characteristic and geographic detail as we combine more years of sample. Along with those estimates, we release various quality measures (U.S. Census Bureau 2013c). In contrast, performance measures focus almost exclusively on the survey's response rates. The performance benchmark for the Department of Commerce had historically required that the ACS achieve a 92 percent response rate nationally (U.S. Census Bureau 2013a). Since the survey's start in 2005, the weighted survey response rate has been high – over 97 percent (U.S. Census Bureau 2013c).

The Census Bureau's Statistical Quality Standards require that each survey manager assess the quality of key survey estimates based on the major survey deliverables. Given that the goal of the ACS is to produce estimates for small geographic areas and small population groups, ACS managers consider it critical to achieve high response rates below the national level. For this reason, annual ACS statistical quality assessments involve the calculation and review of response rates at the county-level. Our annual statistical quality review also includes assessments of item nonresponse, coverage, and sampling error (King 2012). Figure 2 shows ACS weighted county-level survey response rates. They range from a low of 87.2 percent to a high of 100 percent. The darkest orange highlights the large number of counties with rates of about 97 percent or more. Even the areas with the lowest relative rates have exceptionally high rates – the paler orange shading is still over 90 percent.

In the ACS, we calculate and monitor mode-specific response rates (e.g., Cyffka 2013). Given the sequential nature of ACS data collection, the response rates in the early modes (mail, Internet, and telephone) drive the workloads for subsequent modes. The interviewing success in our final mode, computer-assisted personal visit interviewing (CAPI) defines the survey's nonresponse universe. CAPI data collection is also the most expensive mode. For these reasons, this paper focuses on the response rates, costs, and burden associated with this final data collection mode.

While ACS managers have analyzed the survey's response rates across the country and monitor mode-specific response rates regularly, we have not evaluated what it has cost us to maintain these rates. This paper discusses some additional metrics that can help us understand the burden placed on survey respondents in the final mode of data collection and the costs in terms of hours spent in converting nonrespondents.



Source: 2007-2011 ACS 5-Year Estimates

Figure 2. County-level Survey Response Rates Associated with the 2007-2011 ACS 5-Year Estimates

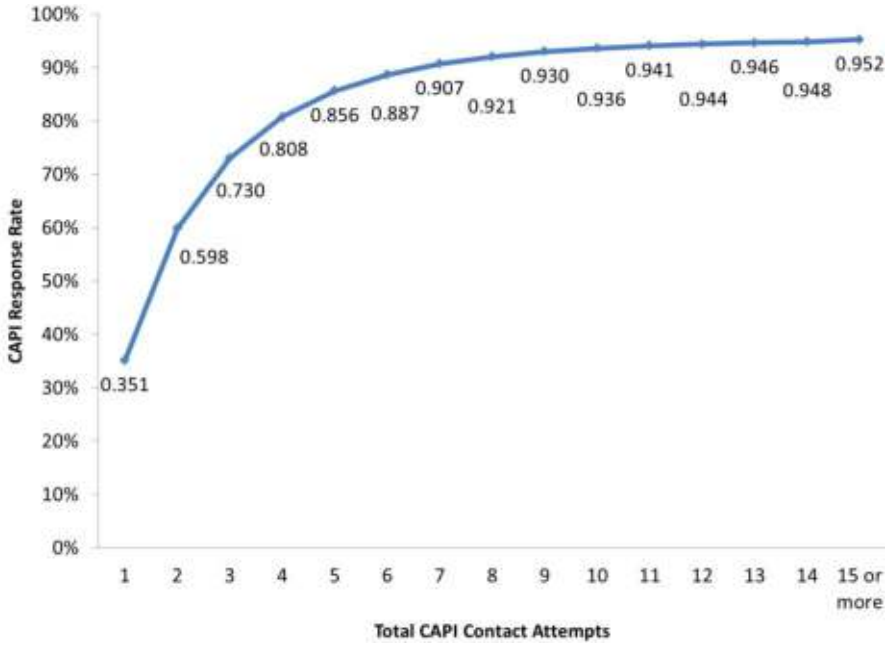
3 Measuring Costs

3.1 Burden

The Contact History Instrument (CHI) provides a resource for survey managers to understand the potential respondent burden associated with the response rates they achieve. Figure 3 shows the relationship between the response rates obtained in the 2012 ACS CAPI operation and the associated number of contact attempts. The final ACS CAPI response rates have consistently been around 95 percent. We estimated the response rates in Figure 3 as ratios of the number of completed interviews (occupied and vacant) after each round of contact attempts relative to the total eligible universe (RR2 in AAPOR 2011). For example, after all seventh contact attempts, we estimate that the response rate was about 90.7 percent. You can see that significant efforts in terms of contact attempts are required to obtain the final 5 percent.

In this figure, "contact attempts" means any form of attempt to contact the sample household. This could include a successful contact leading to an interview, a refusal, or a request for a call back or many forms of noncontact. Noncontacts include some attempts that could be invisible to a sample respondent like a field representative driving

by the address to see if anyone was home. We chose to be generous in defining "potential respondent burden" as any contact attempt. It is hard to know which forms of attempted contacts a respondent would perceive as burdensome.



Source: January – December 2012 ACS CAPI

Figure 3. Potential Burden Associated with CAPI Response Rates – 2012 ACS

Given the large sample size in the ACS, even a low percentage of cases with high numbers of contact attempts can represent a notable universe of respondents that may feel harassed by the survey's repeated contact efforts. To give you a sense of scale, we looked at the universe of housing units in the 2012 CAPI operation that were ultimately determined to be survey-eligible, occupied housing units (about 463,000 cases). This universe includes sample cases that we were able to interview and those that we could not interview. Table 1 shows that interviewers resolved about 54 percent of these occupied housing units after 1 or 2 CAPI contact attempts. While we contacted some portion of these households by mail and phone in previous modes, in this final mode they have the lowest levels of respondent burden. In contrast, we estimate that about 19,000 households (or just over 4 percent) involved 10 or more CAPI contact attempts.

Table 1. CAPI Contact Attempts - 2012 ACS

Total CAPI Contact Attempts	Percent of eligible, occupied sample housing units (in personal visit mode)	Estimated Respondents Each Year
1 or 2	54.2	251,000
3 to 5	30.7	142,000
6 to 9	10.9	50,000
10 or more	4.2	19,000

Source: January – December 2012 ACS CAPI

The Census Bureau is beginning to use CHI paradata to monitor field performance. While dashboards and performance reports focusing on response rates still exist and all levels of management give them great consideration, some managers are starting to see the value of additional performance information. We recommend that managers produce and review statistics such as those listed below to understand respondent burden associated

with final response rates. The CHI paradata provide additional detail about the reasons for noncontacts and noninterviews by contact attempt. Field representatives record the reasons that a respondent cites for not completing the interview at a given attempt. Some of these reasons note degrees of reluctance on the part of the sample household. We should use these reasons to measure the burden in terms of numbers of contact attempts after signs of respondent reluctance. We suggest that managers produce and review each of the metrics listed above after indication of respondent reluctance.

Metrics like these warrant continuous review. The ACS organizes data collection by monthly panels and this is an obvious choice for production and analysis of burden measures. We suggest multiple levels of review and analysis. Headquarters management should review national and regional measures while regional offices should analyze metrics by field supervisor and field representative.

- Mean number of total contact attempts – overall and by final outcome
- Mean numbers of specific types of contact attempts (e.g., those resulting in respondent reluctance)
- Distribution of total number of contact attempts and specific types of contact attempts required
- Workload requiring "high" numbers of contact attempts
- Workloads with repeated reluctance expressed by a respondent
- Response rates by required contact attempts

There are other ways to try to reflect cumulative respondent burden, considering earlier mode's contact attempts. Our future research intends to look more closely at how to both acknowledge this burden in any metric and how to utilize previous burden in developing CAPI stopping rules.

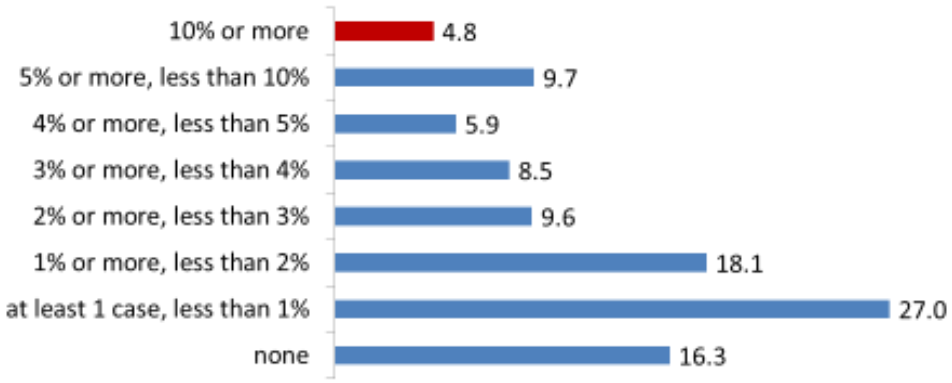
3.2 Efficiency

We recognize that efficiency and burden both involve contact attempts. When we review data on contact attempts to understand the respondents that we may have over-burdened, we may also find that those data tell us about how efficiently our data collection is running. CHI supports the production of data collection efficiency statistics such as:

- Percent of workload completed with fewer than 3 contact attempts
- Percent of workload requiring 10 or more contact attempts
- Percent of contact attempts that resulted in an interview

To understand the level of effort required, efficiency metrics like the percent of the workload that a field representative completed with fewer than 3 contact attempts can identify efficient work that we should acknowledge. Alternatively, field representatives with high proportions of cases requiring 10 or more contact attempts may warrant review of their work. There are many possible explanations for low and high contact cases. Certainly, some areas of the country are easier to gain respondent cooperation, even to gain access to sample households and find them at home. Data collection managers need to understand the reasons for these high contact attempt cases. Calculating the proportion of contact attempts resulting in an interview is another way to acknowledge efficient data collection efforts.

Outlier reports are a good way to identify areas with potential data collection challenges and field representatives that may need additional training. Figure 4 is one example. It summarizes CHI data from 270 ACS field representatives. To ensure that we had sufficiently large workloads, we limited the analysis to the 45 field representatives with the largest ACS workloads in 2012 from each regional office. Using CHI, we calculated the percent of each ACS field representative's workload that required 10 or more contact attempts. This graph is the full distribution. About 16 percent of all field representatives never had a single case warranting 10 or more attempts, another 27 percent had a very small proportion (less than 1 percent). Figure 4 also identifies that close to 5 percent of the field representatives needed to make 10 or more contact attempts to complete about 10 percent of their work. While Figure 4 is a national summary, we recommend regional-level and supervisory-level reports that incorporate this type of information.



Source: January – December 2012 ACS CAPI

Figure 4. Distribution of ACS Field Representatives by the Proportion of Their Workload Requiring 10 or More Contact Attempts – 2012 ACS

Additional efficiency statistics include:

- Percent of interviews completed by telephone
- Distribution of contact attempts by time of day, day of the week
- Percent of cases without an initial contact attempt after the first week of data collection (i.e.,cases that haven't been started)

This type of information could provide a basis for understanding why some field representatives have high numbers of contact attempts. Monitoring cases that a field representative hasn't started is critical to ensuring that each case has the full time period available to reach a household and gain a response.

3.3 Time Spent

The pursuit of high response rates comes at a cost other than burden and efficiency. While the absolute number of cases with high numbers of contact attempts is small, the time spent to convert these noninterviews can be noteworthy. It is important to summarize cost information, including measures such as:

- Mean costs (hours spent, mileage) per contact attempt – this could be by field representative or aggregated to other management levels
- Cumulative costs associated with each case – this would allow managers to determine if additional investments in a case are worth the resources
- Costs per percentage point increase in response rate

In the ACS, as in other surveys, field representatives record their data collection costs (in terms of hours spent and mileage) each day. They do not, however, 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.

Staff in the Survey Analytics area of the Census Bureau developed a methodology that used an ordinary least squares regression model to produce estimates of hours spent for each contact attempt. For each interviewer-day, they calculated the number of personal visit attempts that were successful, the number of personal visit attempts that were unsuccessful, the number of telephone attempts made, and the number of miles driven. They split the personal visit attempts into these two categories because of the differential amount of time required for successful versus unsuccessful efforts. Because most telephone attempts are unsuccessful in obtaining an interview, they chose to collapse the telephone attempts. The model used miles driven as a proxy for time spent traveling. Analysts used these coefficients to predict the hours worked that day at a regional office level. 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. These data include interviews and noninterviews, occupied and vacant units, and cases ultimately determined to be ineligible for the survey.

To approximate the data collection hours required to achieve high response rates we calculated the hours associated with all first contact attempts, second attempts, etc. Table 2 summarizes these results. For example, the row associated with the fifth contact attempt tells us that field representatives spent over 92,000 hours on fifth attempts. The cumulative hours associated with cases having one through five attempts was over 1.3 million hours. The results allow us to estimate the costs (in terms of hours) if we had stopped data collection after reaching some maximum number of attempts. The final two columns estimate the hours saved (and the percent of the total hours saved) if we had not pursued additional contact attempts. Therefore, we estimate that we would have reduced data collection costs by about 12 percent if we had stopped after the fifth attempt.

Because only a small proportion of the total workload for personal visit follow-up required high numbers of contacts, we did not expect the total hours associated with their pursuit to be large. Here we see that we could reduce hours spent on CAPI data collection by about 5 percent if we truncated data collection after the seventh attempt, for example. If we allowed eight contact attempts we would eliminate about 40,000 interviewing hours and reduce hours spent by about 2.6 percent. Stopping rules that allowed up to nine attempts would have a minimal impact on cost reductions.

Table 2. Estimated Hours Worked – 2012 ACS CAPI

Contact Attempt	Estimated Hours	Cumulative Hours	Hours Saved if Stopped After this Attempt	Percent of Total Hours Saved if Stopped After this Attempt
1	576,956	576,956	967,232	62.6
2	347,944	924,899	619,289	40.1
3	207,780	1,132,680	411,508	26.6
4	135,598	1,268,278	275,910	17.9
5	92,180	1,360,457	183,731	11.9
6	64,520	1,424,977	119,211	7.7
7	45,885	1,470,862	73,326	4.7
8	32,881	1,503,742	40,445	2.6
9	23,773	1,527,515	16,673	1.1
10 or more	16,672	1,544,188	0	0.0

Source: January – December 2012 ACS CAPI

These contact-level cost data also allow us to approximate the hours required to pursue noninterviews based on the estimated number of required attempts. Table 3 summarizes those results. We estimate that the cases resolved after a single contact attempt involved an average of 1 hour while the cases requiring 10 or more contact attempts involved an average of 6.8 hours. These data allow us to assess the relative costs (in terms of hours) of obtaining responses for these sample cases. For example, we estimate that a case that required five contact attempts cost about four times as much as a case resolved after a single attempt. Data summaries, such as these, allow managers to assess the investment involved in pursuing higher response rates. Understanding the costs and benefits in terms of hours spent to increase the response rate begins to tell a more complete story.

Table 3. Estimated Total Hours by Required Contact Attempts – 2012 ACS

Total Contact Attempts	Number of Resolved Cases	Total Hours	Hours per Case
1	252,406	246,425	1.0
2	174,221	299,778	1.7
3	92,214	226,059	2.5
4	55,036	175,233	3.2
5	35,213	136,850	3.9
6	23,169	106,804	4.6
7	15,940	85,331	5.4
8	11,317	67,611	6.0
9	7,931	52,929	6.7
10 or more	21,752	147,167	6.8

Source: January – December 2012 ACS CAPI

3.4 Quality

While we usually think of quality as a benefit, another important cost dimension to explore is whether an increase in unit response has unintended negative effects on item nonresponse, coverage, or measurement errors. The pressure on field representatives to achieve high response rates could lead to falsification of data, misclassification of units or to other interviewer errors. Frustrated respondents could also intentionally leave people off the survey form or provide incorrect information in order to "be done with it".

In the ACS we calculate an aggregate completeness score for every completed interview/submitted response. It is a simple ratio of the items with nonblank responses to the items requiring a response. We summarize these scores as percentages so a value of 100 means that the respondent answered every item that they should have answered. When we study the completeness of all CAPI interviews, we find that in 2012, field representatives were able to obtain about 88 percent of all required responses, on average. When we look at the cases that required a single contact attempt, the mean score is about 89 percent. If a case required 10 or more contact attempts, the score drops to 84 percent. The completeness of responses suffers when less cooperative respondents are pushed to respond, but the loss is not that great. Managers should supplement response rates with these types of completeness metrics to be confident that we are not improving unit response at a cost of item-level completeness.

We need to invest more effort into designing methods to produce performance metrics that assess measurement error. The ACS uses reinterview methods to identify instances of falsification of data, misclassification of occupied units as vacant or nonexistent and other interviewer errors. Computer-assisted recorded interviewing (CARI) is another tool that holds promise to identify measurement error. It may be useful to analyze data from our reinterview by level of effort required or to oversample cases with high numbers of contact attempts to be more confident that pressure to reduce nonresponse is not introducing other errors. The Census Bureau recently completed a comprehensive study of response variance in the ACS by conducting a content reinterview survey with sample households. The results estimate simple response variance by item and mode of data collection (Murphy 2014). Studies like this can identify measurement errors warranting question redesign or changes in interviewer training.

The Census Bureau is developing a statistical process control system as a way to try to identify response and interviewer errors during data collection. We could also use this automated system to study the quality of responses occurring in the final week of data collection.

4 Measuring Benefits

The previous section discussed some of the costs associated with high response rates. This section touches on the value or benefit. The reduction of nonresponse bias without increasing variance is the benefit we hope to achieve by obtaining high response rates. We believe that we should be reducing the risk of nonresponse bias when we reduce levels of nonresponse but only a nonresponse bias study can tell us if the characteristics of the households that we convert after these repeated efforts improve the representativeness of the interviewed population. If we find little added value, from a nonresponse bias perspective, we should question if the costs outweigh the benefits of pursuing

these interviews. The ACS is undertaking a nonresponse bias analysis to study the effect of various stopping rules on the resulting survey estimates.

The pursuit of high response rates also leads to a greater number of completed interviews and this can reduce levels of sampling error. Each completed interview has value in this respect and for this reason it is useful to monitor the final numbers of completed interviews. If we determine that we can reduce response rates without adding significant nonresponse bias, we still need to consider the potential increase in sampling error.

5 Understanding Reasons for Nonresponse

Response rates focus on the ratio of completed interviews to eligible cases. Managers should look beyond these rates to review the full distribution of interviewer outcomes including the proportion of cases classified as ineligible. Table 4 provides an example of how CHI data can add richness to the simple response rate statistic. This information, especially at lower levels, can point to possible problems. Distinguishing noncontacts from other noninterviews is also very important when trying to understand the reasons for nonresponse. Noncontacts may be due to poor choices of when interviews are attempted.

In the ACS we collect data from both occupied and vacant units. Usually a household respondent is involved only at occupied units; vacant units require a proxy like a resident manager or a real estate agent. The amount of data collected for vacant housing units (and thus, the burden) is dramatically less than the data collected from occupied housing units. The effort required to make contact with a knowledgeable source can be a harder task for vacant versus occupied units. Therefore, it is important to separate the vacant and occupied universes before looking too closely at contact attempts and costs.

It is easy to derive alternative statistics like contact rates, cooperation rates, and refusal rates from the full breakout of final outcomes. See AAPOR (2011) for detailed descriptions and definitions. Contact rates measure the proportion of all cases for which the survey was successful in reaching some household member. Cooperation rates isolate success in gaining an interview for only the subset of cases contacted. We calculate refusal rates as the proportion of all cases in which a housing unit refuses to do an interview or breaks off out of all potentially eligible cases. All of these metrics tell us much more than the response rate alone.

Table 4. Distribution of Final Outcomes by Regional Office – 2012 ACS CAPI

	Percent of Total CAPI Workload							
	Completed Interviews (occupied)	**Completed Interviews (vacants)**	**Refusal**	**No One Home**	**Other Non-interview**	**In-eligible**	**Total CAPI Workload**	**CAPI Workload**
Atlanta	58.8	28.8	1.5	0.9	1.0	9.0	100.0	116,704
Chicago	56.7	32.0	1.9	0.8	0.9	7.8	100.0	120,206
Denver	63.7	24.2	1.7	0.7	0.9	8.8	100.0	144,851
Los Angeles	70.9	18.8	2.4	1.0	1.6	5.3	100.0	100,383
New York	61.6	25.8	3.2	2.4	1.5	5.5	100.0	106,717
Philadelphia	61.1	27.3	2.6	1.2	1.2	6.6	100.0	100,244
TOTAL	62.0	26.2	2.2	1.1	1.1	7.3	100.0	689,105

Source: January – December 2012 ACS CAPI

6 Tools - Unified Tracking System

The Census Bureau has invested in the development of a data warehouse to organize CHI and other paradata in ways to help survey managers make better decisions. We call this our Unified Tracking System (UTS). The UTS references outcomes and specific paradata to create valuable reports for all Census Bureau surveys. A special feature of the UTS is the integration of survey paradata from our CAPI data collection activities into standard reports that can improve managers understanding of some of the data collection issues confronting their staff. The

goal of the UTS was to create a common enterprise solution for a common problem – tracking of progress, quality, and cost and understanding operational efficiencies.

ACS staff participated in a team, with representatives from all surveys conducted by the Census Bureau, to identify the type of information that survey managers need. The UTS includes a detailed set of reports with capabilities to "drill down" geographically to understand local data collection management and identify potential data collection issues. For example, one report summarizes the number of contact attempts associated with completed interviews at the regional office level and analysts within each regional office can review smaller management areas, down to individual field representatives. Other reports look at outcomes, contact attempt strategies, and costs. The ACS is only beginning to work with these reports to improve our understanding of the level of effort required to maintain high response rates.

7 Conclusions

Survey managers can benefit from increased efforts to understand the full set of costs associated with high response rates. Assessments should look at potential respondent burden, hours spent, and the potential loss in quality due to increases in other forms of error. Managers also need to measure the benefits in terms of reductions in nonresponse bias and reductions in sampling error to determine if the benefits of high response offset these costs. Paradata from CHI and other sources can help us.

To acknowledge the ACS program's commitment to reducing burden on sample households, we are using CHI to measure burden levels and identify areas of concern. We are developing improved tools to identify field representatives with high contact attempts – outlier reports that field managers can use to determine where feedback may be important. Research is underway to assess nonresponse bias reductions associated with achieving high response rates, to understand the possible quality implications of reductions in survey response rates.

ACS staff recently completed a similar analysis of paradata from the ACS telephone follow up operation. Based on that research, we identified possible changes to the call parameters (Griffin & Hughes 2013). The Census Bureau implemented parameter changes in 2012 production that we knew would reduce response rates for this operation. We felt that the reduction in respondent burden and costs offset the loss in response. Subsequent evaluations confirm that these changes had the expected impact on production (Griffin 2013). Moving beyond response rates using paradata will provide us with much greater insights into the true success of the ACS.

8 References

AAPOR. 2011. *Standard definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys.* http://www.aapor.org/Resources

Bates, N. 2003. Contact histories in personal visit surveys: the Survey of Income and Program Participation (SIPP) Methods Panel. AAPOR. http://www.fcsm.gov/committees/ihsng/aapor2003proceedingforpdf.pdf

Biemer, P. and Lyberg, L. 2003. Introduction to Survey Quality. Wiley Series in Survey Methodology.

Cyffka, K. 2013. "Results from the American Community Survey and Puerto Rico Community Survey 2012 Housing Unit Response Rates and Margins of Error by Mode." DSSD 2012 American Community Survey Memorandum Series #ACS12-S-22.

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

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

Groves, R. 2011. "The Future of Producing Social and Economic Statistical Information," Director's Blog, http://directorsblog.blogs.census.gov/2011/09/, U.S. Census Bureau.

Groves, R. 2006. "Nonresponse Rates and Nonresponse Bias in Household Surveys." *Public Opinion Quarterly* 70(5):646-75

Groves, R. & Peytcheva, P. 2008. "The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis." *Public Opinion Quarterly* 72(2):167-89

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

Matthews, B., Davis, M., Tancreto, J., Zelenak, M., & Ruiter, M. 2012. 2011 American Community Survey Internet Tests: Results from Second Test in November 2011. 2012 American Community Survey Research and Evaluation Report Memorandum Series #ACS12-RER-21-R1. http://www.census.gov/acs/www/library.

Murphy, P. 2014. American Community Survey 2012 Content Reinterview Survey. 2013 American Community Survey Research and Evaluation Report Memorandum Series #ACS14-RER-01.
http://www.census.gov/acs/www/library

Office of Management and Budget. 2006. Office of Management and Budget Standards and Guidelines for Statistical Surveys.
http://www.whitehouse.gov/sites/default/files/omb/inforeg/statpolicy/standards_stat_surveys.pdf

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.

Salvucci, S., Wenck, S., Hamsher, S., & Bates, N. (2002). Response Rate Summit: National Health Interview & Consumer Expenditure Quarterly Surveys. Summary Report. Synectics for Management Decisions, Inc., and U.S. Census Bureau. http://www.fcsm.gov/working-papers/summitreportfinal.pdf.

Tan, L. 2011. An Introduction to the Contact History Instrument (CHI) for the Consumer Expenditure Survey. Consumer Expenditure Survey Anthology, 2011.

U.S. Census Bureau. 2010. Design and Methodology – American Community Survey.
http://www.census.gov/acs/www/Downloads/survey_methodology.

U.S. Census Bureau. 2013a. U.S. Census Bureau Fiscal Year 2013 Quarterly Performance Report

U.S. Census Bureau. 2013b. U.S. Census Bureau Quality Standards.
http://www.census.gov/quality/standards/Quality_Standards.pdf

U.S. Census Bureau. 2013c. American Community Survey: Methodology – Sample Sizes and Data Quality.
http://www.census.gov/acs/www/methodology/response_rates_data/index.php.