TRENDS IN U.S. FACE-TO-FACE HOUSEHOLD SURVEY NONRESPONSE AND LEVEL OF EFFORT

DOUGLAS WILLIAMS*
J. MICHAEL BRICK

Research on nonresponse in face-to-face surveys in the United States has shown that nonresponse increases over time for most surveys, but there are also periods of fairly stable rates. Surprisingly, nonresponse for face-to-face surveys has not been as widely studied recently as compared to telephone surveys. The focus on telephone surveys may be due to the dramatic increase in nonresponse for these surveys or perhaps because face-to-face surveys still achieve relatively high levels of response. This paper updates nonresponse trends for face-to-face household surveys conducted in the United States since 2000. The review provides a comprehensive picture of the industry by including surveys conducted by government, private, and academic organizations. The relative role of refusals and noncontacts in the total nonresponse is also presented. We tie the trends in nonresponse with data on the level of effort for some of these surveys. Many researchers have suggested that extra effort has been needed to prevent response rates from falling even more precipitously but lacked the effort data to evaluate this hypothesis for face-to-face surveys. Some data on field effort are becoming available, so this question can be addressed for the first time across more than one survey. To complete the picture, we also look at loss over time for longitudinal face-to-face surveys.

KEYWORDS: Face to face; Level of effort; Re-interview; Response rates.

DOUGLAS WILLIAMS is a Senior Survey Methodologist at Westat, 1600 Research Blvd, Rockville, MD, USA. J. MICHAEL BRICK is a Vice President and an Associate Director at Westat, 1600 Research Blvd, Rockville, MD, USA.

*Address correspondence to Douglas Williams, Westat, 1600 Research Blvd, Rockville, MD, USA; E-mail: DouglasWilliams@westat.com.

1. INTRODUCTION

The trend of increasing nonresponse has been well documented in numerous studies over the last 50 years, but recently most of the interest has been directed to telephone surveys. Examination of face-to-face survey nonresponse is more complicated because changes in procedures (e.g., computerization) and interventions (e.g., increasing incentives) confound the picture. Unlike telephone surveys where data on level of effort are easily available, the lack of comparable data in face-to-face surveys has hampered researching this important aspect of nonresponse analysis.

This paper is a much-needed update to the nonresponse literature for face-to-face surveys. We begin with a historical overview and point out that the increase in nonresponse is not a new concern. We update the trends for face-to-face surveys in the United States beginning in 2000 with data from several studies across a range of survey organizations. This time period is especially important because of cultural and communication changes. These changes clearly had devastating effects on telephone survey response rates, but related changes such as the increase in gated communities and security guards (Dillman, Smyth, and Christian 2014, p. 10) have implications for face-to-face surveys. We also show how the components of nonresponse—refusals and noncontact—have changed over this time period. We link the changes in the rates to data on level of effort, focusing on the number of contacts to complete a case. This is a new and important contribution since this type of data has not been available previously. The data show that more effort is required to elicit response. Our last topic covers incomplete response in longitudinal studies, where some of the initial respondents fail to participate in later waves of the survey.

2. BACKGROUND

2.1 Nonresponse Over Time

The history of investigation into nonresponse trends for face-to-face surveys is long. Steeh (1981) is an early review of nonresponse rates for face-to-face surveys that examined surveys dating back to the 1950s. She found increasing trends in nonresponse and identified refusals as the main reason for the increase. Steeh (1981) noted that the trend was a cause for concern, but the problem was not insurmountable. Making comparisons of response rates across surveys was especially difficult during this time due to the lack of standardized approaches for computing rates. The creation of standards for reporting dispositions of cases and response rates (American Association for Public Opinion Research 1998) facilitated a more careful examination of nonresponse that enhances comparability. Despite these efforts, conducting comparisons between studies remains challenging. This is due to the complexity of defining a final disposition given a case's history of contacts, and that can affect reported rates (Blom 2013).

In his AAPOR presidential address, Bradburn (1992) cited the threat nonresponse posed to the validity of surveys and called for action. This was soon followed by Brehm (1994), who reported that nonresponse was increasing across all types of organizations—"academic, government, business, media." Shettle, Guenther, Kaspryzk, and Gonzalez (1994), however, did not find increases in nonresponse for Federal surveys. Smith (1995) continued the time series of Steeh (1981) and showed that nonresponse rates had plateaued since her review. He also noted that the growing popularity of telephone surveys (at that time) could result in lower response rates for face-to-face surveys due to the need to remain competitive on cost. More recently, Atrostic, Bates, Burt, and Silberstein (2001) examined nonresponse rates for some federal face-to-face surveys conducted by the U.S. Census Bureau between 1990 and 1999. They found increases in nonresponse occurring midway through the decade and showed that the increases were because of increases in both noncontact and refusal.

Brick and Williams (2013) and a similar review (Tourangeau and Plewes 2013) by a panel of the National Research Council (NRC) provided some of the most recent data on nonresponse trends. Brick and Williams (2013) included two telephone and two face-to-face surveys for 1996 to 2007. They showed that nonresponse continued to increase, with telephone surveys having the most dramatic increases, while the change in face-to-face survey nonresponse was less severe. The conclusions of the National Research Council (NRC) panel were similar. Meyer, Mok, and Sullivan (2015) show increasing nonresponse for several large household surveys—although they mainly focus on surveys conducted by the Census Bureau.

We restrict our attention to surveys conducted in the United States because including studies from other countries adds significantly to the complexity of the evaluation. Of course, nonresponse in face-to-face surveys has also been studied in several countries (e.g., Synodinos and Shigeru 1999; de Leeuw and de Heer 2002). While the level of response to surveys varies widely across countries, the general trend of increasing nonresponse, or at least the perception of lower response rates, is a consistent finding internationally.

Our review of the literature on nonresponse shows that nonresponse has been increasing for some time, and many studies found refusals to be the main culprit. The reviews reveal some differences due to the organization collecting the data (e.g., academic and government) and to the specific time periods examined. Furthermore, the pattern appears to be periods of relatively stable nonresponse rates followed by periods of increases.

2.2 Nonresponse in Longitudinal Surveys

Face-to-face surveys are very expensive compared to other modes of data collection, and longitudinal face-to-face surveys that require going to sampled

units over time are among the most expensive designs. The main reasons for conducting longitudinal surveys are the statistical advantages associated with estimating change over time and allowing for estimates of gross change and spells of events (e.g., spells of unemployment). In some longitudinal surveys, the costs are reduced by using other modes such as telephone for later waves, and many of the initial costs of sampling are also reduced by repeating data collection from the same units over time.

One of the disadvantages of longitudinal surveys is attrition. Attrition is when a respondent drops out of the study and does not return (Duncan and Kalton 1987). The effects of attrition nonresponse have been studied because attrition is sometimes correlated with the characteristics being estimated (e.g., Frankel and Hillygus 2014). For example, in a survey measuring medical expenditures, attrition might be due to the respondent suffering a serious medical condition and being unavailable for the interview in subsequent times. While some research shows that bias due to attrition may not be any more severe than bias due to base wave nonresponse (Pyy-Martikainen and Rendtel 2008), this varies from study to study. In any case, attrition is a source of nonresponse bias that needs to be addressed during data collection and analysis (Burton, Laurie, and Lynn 2006).

Despite its importance, trends in attrition rates over time have not been discussed extensively in the literature. Schoeni, Stafford, McGonagle, and Andreski (2013) is the only article we identified that examines changes in attrition rates over time for more than one longitudinal survey. The complexity of longitudinal survey designs makes meaningful comparisons difficult. Attrition rates are greatly affected by many design features such as periodicities (time between collections), following rules (which respondents should be contacted in subsequent rounds), lengths (time between the base wave and final collection), and modes (use of telephone in later waves). All of these features vary from survey to survey.

Schoeni et al. (2013) looked at three U.S. surveys—the Panel Study of Income Dynamics (PSID), the National Longitudinal Survey of Youth 1979 (NLSY79), and the Health and Retirement Survey (HRS)—and three international longitudinal surveys covering the entirety of each survey up to 2009. They present wave-to-wave re-interview rates (proportion of those interviewed in the previous wave who are interviewed in the current wave) to reduce the effect of some of the design feature issues mentioned above, especially the different following rules. The three domestic surveys they studied have all been conducted for many years (the PSID began in 1968, the NLSY79 in 1979, and the HRS in 1992) and may not be very representative of current longitudinal surveys with fixed panel lengths. The conclusion from the work conducted by Schoeni and his colleagues was no evidence of changes in re-interview rates for the time period examined.

2.3 Level of Effort

One of the most successful approaches to improving response rates for a survey is to increase the level of the data collection effort (Bradburn 1992; Fuchs, Bossert, and Stukowski 2013; Heerwegh, Abts, and Loosveldt 2007; Marquis 1977). The interaction between effort (i.e., fieldwork) and response rates has been demonstrated in the European Social Survey (Stoop, Billiet, Koch, and Fitzgerald 2010). The number of contact attempts is the most common way of measuring level of effort. Measures of level of effort over time are important to help interpret trends in nonresponse rates, and also because level of effort and cost are often related to survey climate.

Despite its importance, research associating level of effort with nonresponse rates over time is limited. An early investigation of nonresponse trends by Marquis (1977) noted that the lack of data on level of effort made it difficult to understand changes in nonresponse rates. Smith (1995) lamented the unavailability of these data nearly two decades later. Tourangeau (2004) noted the paucity of available data and stated that most researchers believed it required more effort to contact respondents, but only anecdotal evidence existed to assess this hypothesis. A recent review of a face-to-face survey in the Flanders area of Belgium found stable response rates from 1996 to 2013, but the rule for minimum number of contacts had gradually increased from three in 1996 to five in 2007 (Barbier, Loosveldt, and Carton 2015), suggesting increasing effort to maintain rates.

With the adoption of computerized interviewing methods in the 1990s, access to paradata such as level of effort has become more common. Face-to-face surveys pose more challenges than with other modes because the control of the contact effort and recording are less automated than in telephone and mail surveys, potentially leading to errors in the contact history data record. For example, interviewers can be inconsistent in recording whether driving by and observing a sampled house constitutes a contact, or they may be motivated to keep a case active by not recording a contact (Biemer, Chen, and Wang 2013).

The Census Bureau developed a tool for collecting contact paradata and interviewer observations. The contact history instrument (CHI) was first implemented in the National Health Interview Survey (NHIS) in 2004 and is now used across 11 studies administered by the Census Bureau (Virgile 2016). Although these data are very useful, a review of Census Bureau surveys using their contact history paradata found possible underreporting of noncontacts and inconsistencies in reporting refusals (Bates, Dahlhamer, Phipps, Safir, and Tan 2010). Other errors such as "ghost" contact attempts due to interviewers accessing cases for administrative reasons also add some noise to these data (Virgile 2016).

Some researchers have used data on the number of contact attempts to study the utility of additional contact attempts. These efforts focus on truncating contacts, essentially removing cases, to see the effect on nonresponse bias. Heerwegh, Abts, and Loosveldt (2007) and Fuchs, Bossert, and Stukowski

(2013) found that additional contact attempts resulted in a minimal reduction in nonresponse bias. The increase in response from the additional effort came largely from missed appointments rather than noncontact cases. They also observed increases in refusals with increases in effort, consistent with the negative feedback cycle hypothesized by Brick and Williams (2013), which posits that increases in contact attempts reduce response propensities. This is reliant on respondent awareness of each contact attempt, presuming that a higher number of contacts is perceived as harassing or intrusive.

3. U.S. CROSS-SECTIONAL NONRESPONSE

3.1 Selected Surveys

The nine U.S. surveys chosen for this analysis include studies conducted by government, academic, and private research organizations. A key factor in determining which surveys should be included was that they all had to provide access to the data needed to assess trends over time for our selected time period. We also tried to choose surveys that covered a range of survey topics: health-related (six surveys), economic (one survey), crime (one survey), and social attitudes (one survey). Of the nine surveys, five are cross-sectional, and four are longitudinal surveys. For the longitudinal surveys, we examine the initial or base year interview only; later we look at losses in subsequent waves for these surveys. The time period covered is 2000 to 2014, although some of the surveys do not cover the entirety of this period. This time period coincides with the adoption of computerized field data collection and is a period of rapid technological evolution throughout society. All of the surveys are categorized as face-to-face surveys because this is the predominant and default method of data collection. Some of the surveys use telephone as an alternative mode when it is more convenient, but this is generally a small proportion of interviews. For the longitudinal surveys, the initial interview is predominately face-to-face although follow-up surveys may be done by telephone.

The NHIS is sponsored by the National Center for Health Statistics (NCHS) and conducted by the U.S. Census Bureau. The NHIS is a cross-sectional survey that monitors the health of the noninstitutionalized population of the United States. An adult and a child (if any) are selected from each sampled household, and interviewing is done throughout the year. The last substantial change to the NHIS instrument occurred in 1997, partly in response to declining response rates (National Center for Health Statistics 2000).

The Current Population Survey (CPS) is sponsored by the Bureau of Labor Statistics and conducted by the Census Bureau. The March supplement, referred to as the Annual Demographic Supplement prior to 2003 and the Annual Social and Economic Supplement (ASEC) since 2003, is an additional component that produces the official annual estimate of poverty in the United

States. The sample design for the CPS is a 4-8-4 rotating design. Households are interviewed for four consecutive months, not interviewed for eight months, and then interviewed for the subsequent four months. The first and fifth interviews are conducted face-to-face, where others are primarily conducted by telephone. While the March supplement is administered to the entire March CPS sample, other sample months are included to increase the representation of some groups (see U.S. Census Bureau 2006 for detail on sample groups). The last redesign of the CPS occurred in 1995 with the goal of improving measurement, updating methods, and taking advantage of changes in data collection technology (U.S. Census Bureau 2006).

The National Crime Victimization Survey (NCVS) is a longitudinal survey sponsored by the Bureau of Justice Statistics and conducted by the Census Bureau. It estimates nonfatal criminal victimization in the United States. All household members age 12 years and older are interviewed. The design is a rotating sample conducted throughout the year, with households interviewed every six months for a total of seven interviews. The first interview is generally face to face (20 percent are by telephone), but telephone mode is a much larger proportion for later waves. The NCVS has both a household and personal interview. We focus on the household interview. The last substantial change to the NCVS survey instrument was in 1993.

The General Social Survey (GSS) is conducted by NORC and supported by funding from the National Science Foundation. It monitors social change in the United States. The interview includes a core set of questions along with supplements or topics that vary over time. One adult per household is sampled to complete the interview. Beginning in 1994, the GSS has been conducted in even-numbered years only. The GSS is a cross-sectional survey, but from 2006 through 2012 experimented with a panel design that included two reinterviews corresponding with the every-other-year pattern of the GSS (see Smith, Marsden, and Hout 2015). We include only the initial interview for these years. About 10 to 20 percent of the base wave interviews were conducted by telephone.

The National Survey on Drug Use and Health (NSDUH) is sponsored by the Substance Abuse and Mental Health Services Administration (SAMHSA) and has been conducted by the Research Triangle Institute (RTI) since 1988. It produces estimates of tobacco use, alcohol use, illicit drug use, and mental health. Up to two persons per household age 12 years or older can be sampled from a household based on age and other sampling criteria. The survey instrument is updated annually, but the last substantial change was the conversion to computer-assisted personal interviewing (CAPI) in 1999.

The Medical Expenditure Panel Survey (MEPS) is sponsored by the Agency for Healthcare Research and Quality (AHRQ) and has been conducted by Westat since the current survey design began in 1996. This survey provides data on health utilization of U.S. households, including health status, frequency of health service use and access, health cost, and payment source. The MEPS

includes multiple components, but we only consider the household survey. The sample for the MEPS is drawn from households responding to the NHIS the previous year.[1] The survey is a rotating design where households are interviewed approximately every six months for a total of five interviews.

The MEPS survey has some features of the second wave of a longitudinal survey as the sample includes a subsample of those cooperating in the NHIS. We include it in this analysis because the MEPS is a new survey request from a separate survey organization almost one year after the NHIS. We believe it suits our main purpose of estimating trends well.

The Medicare Current Beneficiary Survey (MCBS) is sponsored by the Centers for Medicare and Medicaid Services (CMS), and Westat collected the data during the time period covered in this review. The MCBS produces estimates of access to health care, cost, and sources for payment of services for U.S. Medicare-eligible households. The MCBS was an early adopter of CAPI interviewing in 1991. The MCBS uses a rotating design where sampled households are interviewed three times a year for a period of four years, yielding a total of 12 interviews. Respondents are sampled from the Medicare enrollment file.

The National Health and Nutrition Examination Survey (NHANES) is sponsored by the NCHS, and Westat is the data collector. It is a program of studies that capture health and nutrition data, including a screening interview to sample eligible household members.[2] Household and individual interviews are conducted for eligible household members age two months and older, and examination data are collected in a Mobile Examination Center. Our focus is on the household interview component. The current design is a continuous cross-sectional survey with two-year cycles that began in 1994.

The ninth survey is the National Survey of Family Growth (NSFG). It is sponsored by the NCHS and the University of Michigan's Institute for Social Research has collected the data since 2002. The survey collects information on family life, general health, and reproductive health. After the NSFG was administered in 2002, the design was changed to collect data annually as part of a four-year continuous survey design that started in 2006. The continuous interviewing design was adopted in response to challenges in maintaining high response rates and managing costs (Groves, Mosher, Lepkowski, and Kirgis 2009). One person per household is selected for the interview, and starting in 2002 females and males age 15 to 44 years could be sampled (starting in 2015, this was extended to age 15 to 49 years). Unlike other surveys, the NSFG is only conducted by female interviewers due to the sensitivity of the interview content.

---

1. Responding households are those completing the NHIS, but also include partial completed interviews that are more difficult to complete in MEPS.

2. Additional criteria, such as race, ethnicity, and household income, are also used for sampling. All interviews are done with adults.

3.2 Historical Trends: 1990–2014

Several of the studies we describe have long histories spanning several decades. Although the focus of this paper is on recent trends across a wide selection of studies, it is useful to provide some context on the changes in nonresponse that go back further in time to support the claim that nonresponse has been increasing for some time. Five of the studies we discuss either report response rates or provide data for determining response back to the early 1990s or the start of the study. Figure 1 shows the long-term trends for these studies.

Figure 1 shows that most surveys had higher response rates during most of the 1990s. The exception is the CPS-ASEC, which has maintained relatively high rates overall, but even this survey does not match the response rates for the earliest years in the series. The early 1990s were just around the time of widespread use of computerized administration and standardized methods for reporting response rates. The expansive view figure 1 provides also demonstrates how small year-to-year changes can have large cumulative effects.

3.3 Response Rate Trends: 2000–2014

The goal of our analysis is to look at trends in response rates from 2000 to 2014 across several surveys with different sponsors, topics, context, and complexity. As a result, the rates we present were selected to support consistent trend estimation, and cross-survey comparisons of response rates from one survey to another are not meaningful. To emphasize this goal, figure 2 gives the average response rates across eight of the nine surveys. We exclude the NSFG from this average because it has only a few data points in this time period and experienced a major change in 2006. For surveys that are not conducted annually like the GSS, we interpolated using data for adjacent years to provide a consistent trend that includes a contribution from each study for each year. The points on the graph are simple averages across the studies for each year and do not take into account sample sizes or other differences. The dotted line is the ordinary least squares regression line. The figure shows that overall nonresponse increased during this time period.

Figure 3 shows the responses rates[3] for all nine surveys in this period. For the longitudinal studies, only the base wave interview rates are shown. For all

---

3. The response rates for each study were computed as follows: NHIS-Household response rate (AAPOR RR2), CPS-ASEC-Base wave overall unconditional household response rate for Basic CPS and March supplement (AAPOR RR2), NSDUH-Overall person response rate (product of unweighted screener and interview response rates for age 12 and older (AAPOR RR6), GSS-sampled adult response rate (AAPOR RR5), NCVS-Base wave household response rate (AAPOR RR2), MEPS-Household response rate for round 1 (AAPOR RR2), MCBS-Base wave response rate (AAPOR RR2), NHANES-Conditional interview response rate (AAPOR RR2), NSFG-Overall (product of screener and main interview) response (AAPOR RR2).
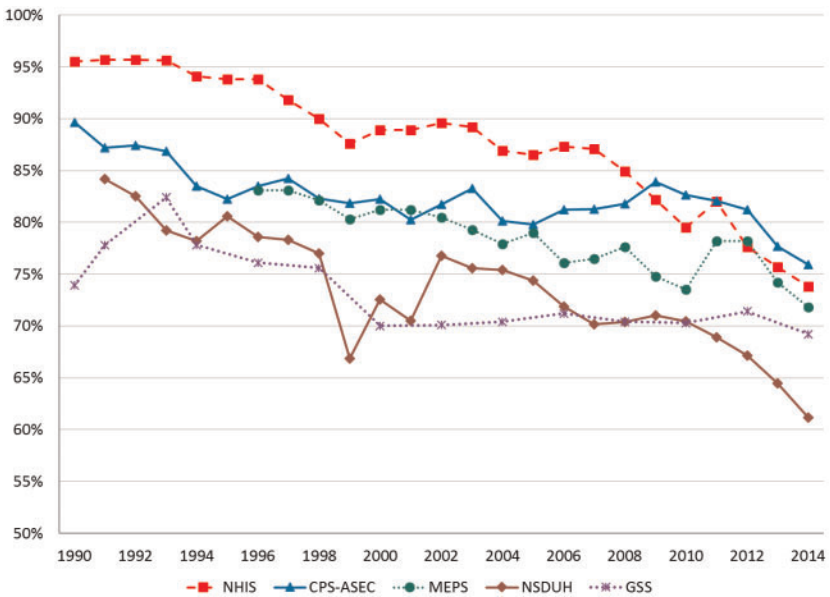
Figure 1. Historical Response Rates back to 1990 through 2014 for Longest-Running In-Person Surveys Selected (Where Data or Information Were Available). *NSDUH 1994:* A questionnaire redesign tested the current (for 1994 and earlier) version and revised version (used starting in 1995), with rates provided for both questionnaire versions. The rate in figure for 1994 reflects the revised version, with a response rate of 78.2 percent compared to 76.5 percent for the earlier version. *NSDUH 1999:* Computer-assisted interviewing (CAI) was tested (previously PAPI), with rates provided for both the CAI and PAPI versions. A CAI response rate of 66.8 percent was used in the figure, compared to 64.2 percent for PAPI. *CPS-ASEC 2014:* A questionnaire redesign was tested for 2014, with rates provided for the existing (current through 2014) version and the revised questionnaire. For consistency, the rate used in the figure is for the existing (for 2014 and earlier) version of 88.0 percent. The response rate for the revised version is 93.7 percent. SOURCES. National Health Interview Survey survey description notes (1997–2014); public use data file readme file (1990–1996); NSDUH data collection final report (1999–2014), public use data file codebook (1991–1998); GSS public use data file codebook, appendix A; CPS-ASEC response was calculated using public use data files (AAPOR RR1); MEPS response was calculated using project data files (AAPOR RR1). CPS-ASEC, Current Population Survey–Annual Social and Economic Supplement; GSS, General Social Survey; MEPS, Medical Expenditure Panel Survey; NSDUH, National Survey on Drug Use and Health.

but two studies, the response rates are unweighted (weighted response rates are given for the GSS starting in 2004 and for the NSFG—the GSS and NSFG use weights primarily to account for subsampling nonrespondents).
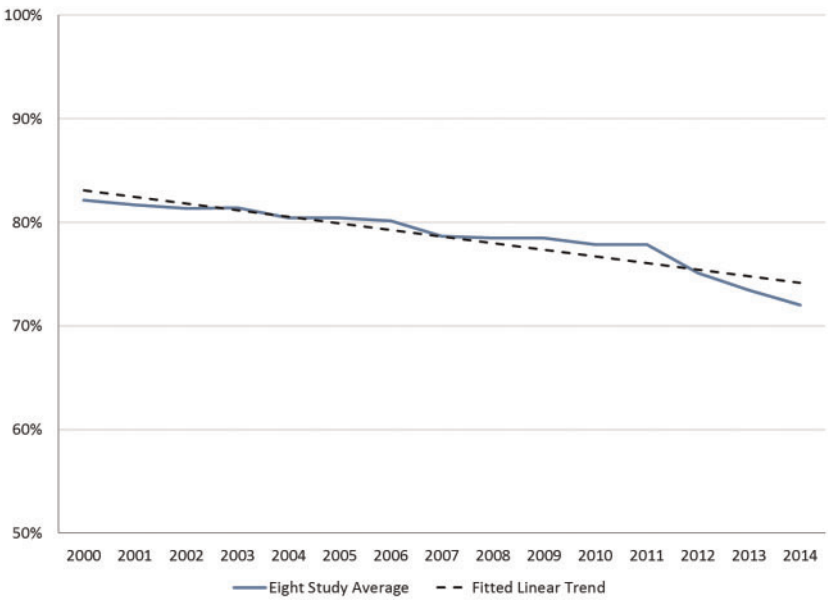
Figure 2. Response Average across Eight Studies for 2000 through 2014.

The figure shows that most surveys have suffered increased nonresponse over this time period, with the exception of the GSS, CPS-ASEC, and NSFG. Furthermore, the trend is not consistent over the period; from 2000 through 2005, the rates show some stability. Beginning around 2006, there is a steeper decline in the response rates for many studies, but the decline occurred later for some (e.g., the NCVS and CPS-ASEC). Lastly, the year-to-year trends are characterized by spikes and some steep declines.

Table 1 gives the slopes for fitted trend lines for each study. The first column confirms the increase in nonresponse over the entire time period discussed above, with three surveys showing an annual decline in response rates of nearly 1 percentage point. The year-to-year change for the CPS-ASEC, GSS, and NSFG are close to zero. For the 2000–2005 period, the response rates declined about 0.5 percentage points on average for those showing declines; for 2006–2014, the declines were greater, with four surveys having annual declines of more than 1 percent and one other with a decline of just under 1 percent. The MEPS does not follow the same pattern as the other surveys for these time periods. It is important to realize that even small annual declines in response rates cumulate over time and can contribute substantially to survey costs.

Turbulence in the rates for many studies is evident in figure 3. These peaks and valleys may be related to changes in the survey climate or methodological
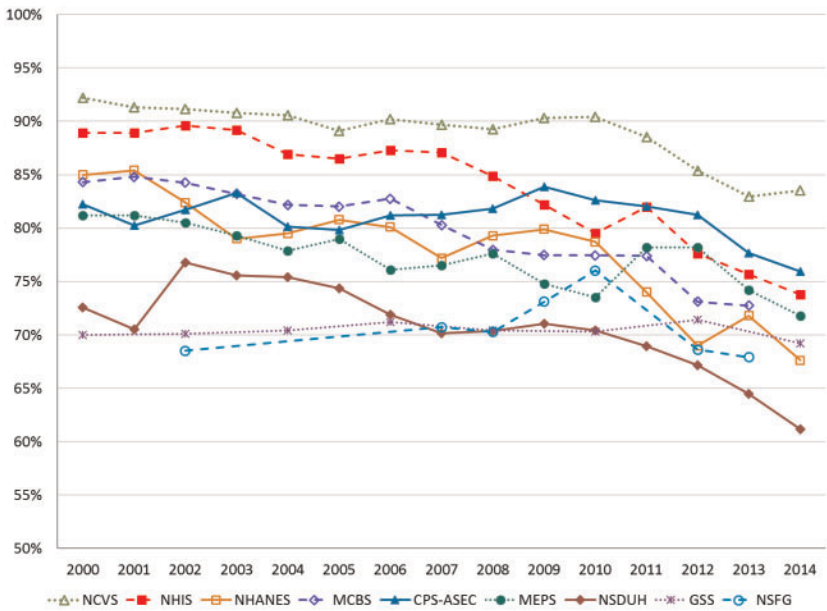
Figure 3. Response Rate Trends from 2000 through 2014. SOURCES. NHIS survey description notes (2000–2014); NSDUH data collection final report (2000–2014); GSS public use data file codebook, appendix A; NSFG rates were provided by the University of Michigan Survey Research Center; for all other studies, rates were calculated from project data files. NOTE. See footnote 3 for details on how the response rates were computed for each survey. CPS-ASEC, Current Population Survey-Annual Social and Economic Supplement; GSS, General Social Survey; MCBS, Medical Expenditure Panel Survey; MEPS, Medical Expenditure Panel Survey; NCVS, National Crime Victimization Survey; NHANES, National Health and Nutrition Examination Survey; NHIS, National Health Interview Survey; NSDUH, National Survey on Drug Use and Health; NSFG, National Survey of Family Growth.

changes in the survey such as the introduction of new technology or changes in procedures. Table 2 gives a few notable methodological changes for these studies during the time period. The timing of some changes appears to correspond with some of the spikes in the figure, but it is impossible to assess whether there is a causal relationship. While we have briefly explored survey-specific interventions, there are too few surveys to analyze survey-specific features that may explain difference in trends in response rates between studies. We do discuss some special causes at the survey level in the next section. Table 2 demonstrates the complexity of the relationship between response rates and factors that influence these rates. Despite differences between surveys, in general the trend is to lower response rates.

Table 1. Annual Slopes in Response Rates by Time Period (R^2 in Parentheses)

	2000–2014	2000–2005	2006–2014
NCVS	−0.53	−0.52	−0.94
	(0.70)	(0.88)	(0.73)
NHIS	−1.10	−0.53	− 1.72
	(0.88)	(0.57)	(0.94)
NHANES	−1.05	−1.20	−1.54
	(0.81)	(0.68)	(0.76)
CPS-ASEC	−0.18	−0.31	−0.58
	(0.15)	(0.18)	(0.41)
MCBS[a]	−0.92	−0.58	−1.28
	(0.92)	(0.85)	(0.89)
MEPS	−0.51	−0.63	−0.33
	(0.66)	(0.78)	(0.16)
NSDUH	−0.71	0.33	−1.03
	(0.74)	(0.21)	(0.80)
GSS[b]	—[c]	0.15	−0.12
		(0.86)	(0.25)
NSFG[a]	−0.02	—[d]	−0.51
	(0.00)		(0.15)

NOTE.— CPS-ASEC, Current Population Survey–Annual Social and Economic Supplement; GSS, General Social Survey; MCBS, Medical Expenditure Panel Survey; MEPS, Medical Expenditure Panel Survey; NCVS, National Crime Victimization Survey; NHANES, National Health and Nutrition Examination Survey; NHIS, National Health Interview Survey; NSDUH, National Survey on Drug Use and Health; NSFG, National Survey of Family Growth.
[a]Last year response rate data available is 2013.
[b]Includes interpolated values for intervening years where data collection did not occur.
[c]Approximates near zero.
[d]Not available due to only one data point for the time period.

3.4 Special Cases

Three studies are notable in the figures and tables above because their response rates do not have substantial declines like the other studies. The CPS-ASEC and GSS are nearly flat for the period. The decline began in 2010 for the CPS-ASEC; for the GSS, there is only a hint of a decline for the most recent year. The NSFG had increases in response rates to 2010 and then what appears to be a relatively sharp decline afterwards. We discuss some features of the NSFG and GSS surveys that might be responsible for these unique patterns. As shown in table 2, the CPS-ASEC had relatively few changes that would explain the overall stability observed.

Table 2. List of Potentially Impactful Methodological or Organizational Changes for Each Study

Project	Implemented change
National Health Interview Survey	2002–2009: Frequent sample size reductions occurred. 2006: Changes to sample design implemented.[a] 2011: Time available for interviewers to complete a case changed from 17 days to one month.
Current Population Survey–Annual Social and Economic Supplement	2002: Sample expansion to improve state estimates of children's health insurance. 2014: Redesigned questions for income and health insurance coverage.
National Survey on Drug Use and Health	2002: Study name changed from National Household Survey on Drug Abuse to current name; introduction of $30 promised incentive.[b]
General Social Survey	2002: Change from PAPI to CAPI administration. 2004: First-year interviewers were permitted to complete some interviews by telephone. 2006: Implementation of rotating panel design and sampling of Spanish-speaking respondents—prior to 2006, these were considered out of scope. 2008: Computer Assisted Audio-Recorded Interview added.
National Crime Victimization Study	2006: Converted to full CAPI instrument. 2011: Bureau of Census implemented refresher training focusing on data quality. Emphasis on interviewer performance focused less on response and more on a combination of response, interview timing, and other data quality measures.[c]
Medical Expenditure Panel Survey	2007: Promised incentive offered increased from $25 to $30. 2008: Implemented promised incentive experiment with monetary levels of $30, $50, and $75 (each level allocated to 1/3 of sample for Panel 13).[d] 2009: Returned to $30 promised incentive for all sample units. 2011: Increased promised incentive amount to $50. 2013: Implementation of Data Quality Initiative promoting higher expectations for use of medical records by participants.
Medicare Current Beneficiary Survey	2006: Changes to supplemental (first panel round) sample from previous year due to new Medicare benefit. The new sample was geographically different, with less sample drawn from large urban PSUs. 2008: Conversion to new software instrumentation.

Continued

Table 2. *Continued*

Project	Implemented change
	2008: Due to budget restrictions, interviewing was conducted only in English and Spanish; other languages were discontinued. This dramatically increased the proportion of language problem dispositions.
	2012: Interviewing was stopped for 4 to 6 weeks around the time of U.S. national elections, impacting field production.
National Health and Nutrition Examination Survey	2007–2010: Sample design for this period was changed to over-sample all Hispanics and not only Mexican-American Hispanics.
	2011–2014: Sample design for this period was changed to over-sample non-Hispanic Asians—the group demonstrating the lowest response.[e]
	2012: PSUs selected included areas expected to have low response propensities.
National Survey of Family Growth	2006: Moved from one-year sample design (last fielded in 2002) to four-year continuous sample design. This resulted in organizational and management changes, such as a large reduction in field staff and changes in interviewer hiring. This began real-time collection of paradata with a one-day lag between field activity and paradata availability.[f]

NOTE.— CAPI, computer-assisted personal interviewing; PAPI, paper-and-pencil interviewing; PSU, primary sampling unit.

[a]For more on changes to the NHIS sample design, see Parsons, Moriarity, Moore, Davis, and Tompkins (2014).

[b]Based on results from an incentive experiment conducted in 2001 (Office of Applied Studies 2002).

[c]For more on implementation of interviewer performance measures, see Bureau of Justice Statistics (2014).

[d]For detail on the methods and results of the incentive experiment, see Agency for Healthcare Research and Quality (2010).

[e]Additional detail on response by demographic group is available in Centers for Disease Control and Prevention (2013).

[f]For detail on organization, management, sample design changes, and implementation of responsive design, see Kirgis and Lepkowski (2013).

The seven NSFG data points in figure 2 correspond to three survey administrations—2002 (also referred to as Cycle 6), 2006 through 2010 (Cycle 7 corresponding to 2007 to 2010), and 2011 through 2015 (Cycle 8: only data for 2012 and 2013 currently available). Cycle 6 in 2002 used the earlier one-time survey design, fielded over an 11-month period. Cycle 7 switched to a continuous design, and this had major effects on the management and sample design,

Table 3. Percent of Respondents Receiving Incentives (Monetary or Other), by Year and Subsampling Group

Group	2004	2006	2008	2010	2012	2014
Before subsampling, %	34.5	39.0	52.4	60.6	50.1	47.1
After subsampling, %	97.3	94.5	91.0	94.2	97.7	98.6
Overall, %	44.5	49.8	60.2	66.9	61.8	63.4

NOTE.— Excludes cases for which information on incentive use is missing or not applicable.
SOURCE. GSS data file release 3, July 31, 2015.

including responsive survey design utilizing survey paradata (Groves et al. 2009; Kirgis and Lepkowski 2013). We could not find any explanation for why these changes resulted in response rate increases for the last two years of Cycle 7 rather than all years in Cycle 7. For the first two years of Cycle 8, the response rates fell, consistent with most of the other surveys.

The flatness of the response rates over time for the GSS is uncanny compared to other studies. Only the two most recent administrations in 2012 and 2014 show any hint of a decline. It is possible that some of the procedures listed in table 2 such as the start of subsampling nonrespondents may have contributed to the stability of the response rates (see Smith 2006). However, nonrespondent subsampling is done in other studies such as the NSFG (Kirgis and Lepkowski 2013).

Like many studies, the GSS utilizes incentives to encourage response. Although little has been published on the GSS incentive structure, it is unlike any of the other studies in this review. Smith (2011) indicated that all sampled cases were offered incentives in 2002, but case level detail on variable incentive amounts used did not begin until 2004, based on data[4] in the public release files. Interviewers, in consultation with their field supervisors, were allowed to determine if an incentive was needed and the amount necessary to obtain an interview. In 2014, field interviewers were given even more discretion over this determination (Fisher and Buha 2015).

Using the public use data files, we computed the percentage of respondents (only data for respondents are available) who received an incentive and the average amount paid, beginning in 2004. Table 3 shows the percentages separately for those who responded before and after nonrespondent subsampling. The percentage receiving an incentive increased steadily across both groups, and the "after subsampling" persons were almost always given an incentive.

---

4. Variables from the GSS public use data file include FEEUSED reporting whether an incentive was used and FEELEVEL reporting the amount of incentive used.

Table 4. Average Amount of Incentive Paid to Respondents Receiving a Monetary Incentive by Year and Subsampling Group

Group	2004	2006	2008	2010	2012	2014
Before subsampling, $	24.40	24.41	33.78	32.74	29.52	31.61
After subsampling, $	45.00	59.79	64.81	66.34	68.28	65.54
Overall, $	31.59	37.91	43.33	41.25	44.76	49.37

NOTE.— Excludes cases missing incentive amounts.
SOURCE. GSS data file release 3, July 31, 2015.

Table 4 shows the incentive amount paid (incentives above \$75 were categorized as "\$75+" and were treated as \$75 for our computations). The incentive amount increased over the time period for both groups, with a sizeable jump in 2008 for the "before subsampling" group and a steadier increase for the "after subsampling" group. The amount offered to initial respondents is nearly stable except for substantial increases in 2008 and 2010. For subsampled nonrespondents, the amount increased after 2004 but was relatively stable thereafter. While these data do not support making causal claims about the effect of the incentive, it is interesting that this unique approach is applied in the only survey with relatively stable response rates over the time period. We surmise that this approach to incentives is one of the active tools being used in the GSS to combat increases in nonresponse.

4. NONRESPONSE IN LONGITUDINAL SURVEYS

Three of the nine surveys in this analysis are longitudinal, but so far we have only analyzed the response rates for the base wave interviews for these surveys. Given the increases in nonresponse in cross-sectional face-to-face surveys, the question arises as to whether the same trends hold for nonresponse in later waves of a longitudinal survey.

Attrition nonresponse occurs when a sample unit that has previously cooperated fails to participate in later waves because the sample units move, cannot be contacted, or simply decide not to continue to cooperate. Attrition is difficult to measure across surveys, and we, like Schoeni et al. (2013), instead examine re-interview rates as a surrogate. The re-interview rate is the percentage of cases who completed the previous wave but not the current wave and is not equivalent to an AAPOR response rate. We compute re-interview rates for three longitudinal surveys in our review: the MCBS, MEPS, and NCVS.[5]

---

5. The GSS rotating sample design started in 2006, but there are too few points to include it in this analysis. The CPS is a rotating sample; however, we restrict our analysis to the March supplement, resulting in few data points.
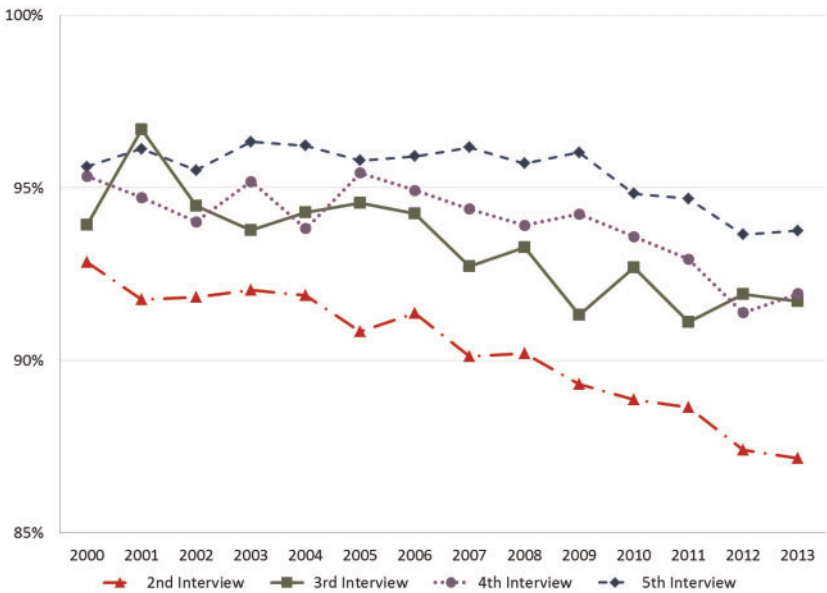
Figure 4. Re-interview Response Rates for MCBS from 2000 to 2013, by Wave.

These surveys all have fixed lengths, and those sampled in the base wave are in the sample for less than five years, a stark contrast with those in the Schoeni et al. (2013) comparisons. The surveys also use telephone interviewing to different extents for the re-interviews.

Figure 4 shows the re-interview response rates for the second, third, fourth, and fifth waves or rounds of MCBS from 2000 to 2013. The figure shows that the second wave rate drops about 5 percentage points over this time period, with the rates for the other waves more consistent although there is some drifting downward. Figure 5 shows the rates for MEPS for the same time period. The MEPS rates are relatively consistent over time, with no discernable decrease.

A similar chart is not given for the NCVS because it is a longitudinal sample of addresses rather than households and persons. The people who move into a sampled NCVS address are newly eligible for the survey. Thus, NCVS re-interview rates have a completely different interpretation compared to the other studies. To assess nonresponse in subsequent waves in the NCVS, we looked only at those addresses where the household composition did not change over the full seven interviews of the survey. For base waves that were selected in 2009 and 2010, about one-third of the addresses had no compositional changes over all seven waves. A high percentage of these addresses completed all seven interviews, but the percentage dropped by roughly 2 percentage points per year, from 81 percent for those first sampled in the beginning of 2009 to 75
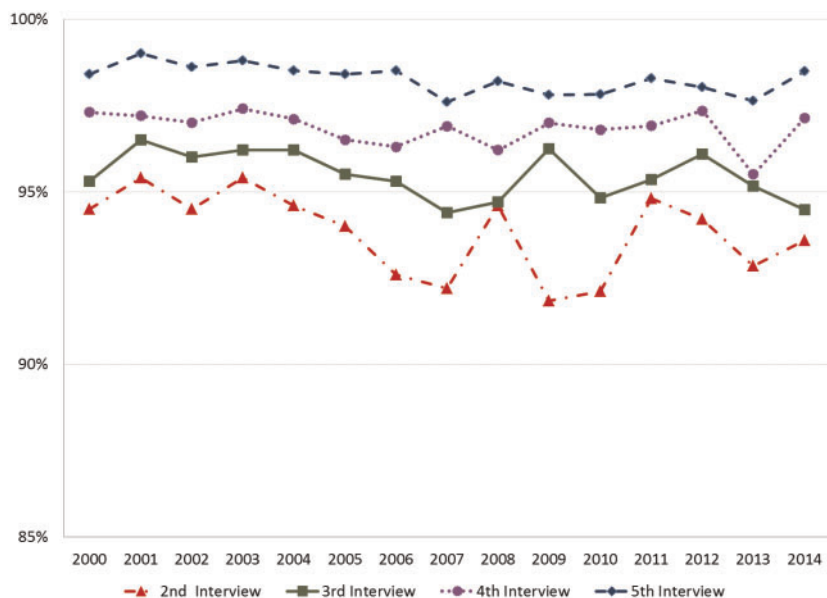
Figure 5. Re-interview Response Rates for MEPS from 2000 to 2014, by Wave.

percent for those sampled initially at the end of 2010. This finding suggests that nonresponse in the later waves of the NCVS may be increasing.

Based on these three surveys, there is mixed evidence about nonresponse in subsequent waves of longitudinal surveys. This finding differs from Schoeni et al. (2013), who found no evidence of falling re-interview rates over time. The MCBS and NCVS show increases in nonresponse for subsequent waves, but the MEPS does not. These inconsistent and ambiguous findings should not be very surprising because of the immense differences in the longitudinal features of the surveys. Just the difference between surveys that are very long-running such as those in Schoeni et al. (2013) and those of fixed length such as those we review could easily affect the rates. Other features such as levels of effort, periodicity, and following rules can also have large effects on the rates.

5. COMPONENTS OF NONRESPONSE

Any investigation into trends in survey response over time would be remiss without looking at the components of nonresponse. The components include refusals, noncontacts, and other noninterviews, with refusals and noncontacts making up the vast majority of nonresponse. We present the proportion of nonresponse due to refusals over time to make the presentation simpler. The supplementary materials provide separate component rates for all the surveys.
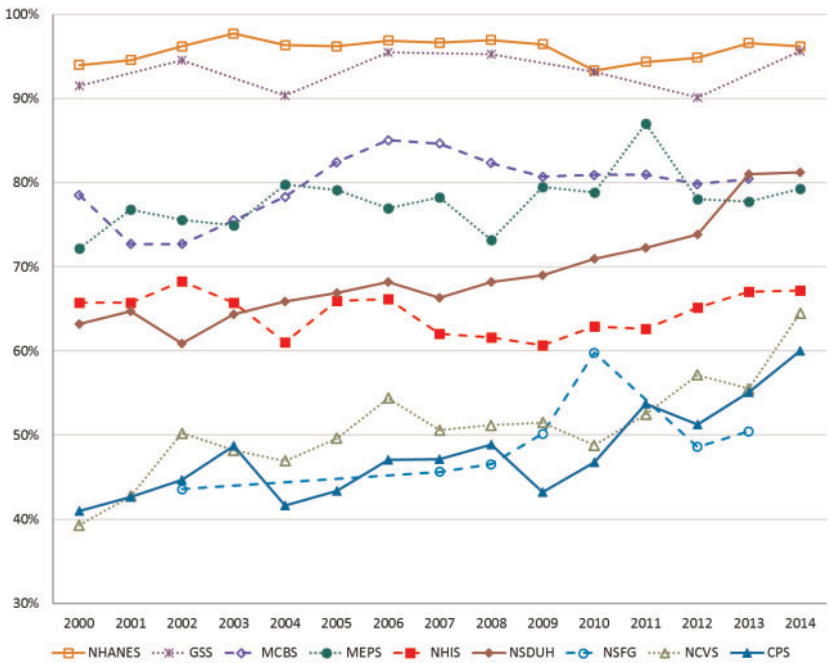
Figure 6. Percentage of Nonresponse due to Refusal from 2000 through 2014.
Note. NHANES and NSDUH include only interview-level nonresponse and exclude screener nonresponse. Nonresponse reported for CPS reflects CPS-BASIC, which is the core CPS interview. Data are not available for CPS-ASEC since nonresponse reported includes both refusals and noncontacts. Separate refusal and noncontact rates cannot be determined from the public files Sources. NHIS survey description notes (2000–2014); NSDUH data collection final report (2000–2014); GSS public use data file codebook, appendix A; NSFG rates were provided by the University of Michigan Survey Research Center; for all other studies, rates were calculated from project data files. CPS, Current Population Survey; GSS, General Social Survey; MCBS, Medicare Current Beneficiary Survey; MEPS, Medical Expenditure Panel Survey; NCVS, National Crime Victimization Survey; NHANES, National Health and Nutrition Examination Survey; NHIS, National Health Interview Survey; NSDUH, National Survey on Drug Use and Health; NSFG, National Survey of Family Growth.

Figure 6 shows the percentage of nonresponse due to refusals over time. While both overall nonresponse and nonresponse due to refusals have increased over time, the percentage of total nonresponse due to refusals has not varied greatly. Thus even though refusals account for the majority of total nonresponse, the other components (noncontacts and other nonresponse) are increasing at about the same rate in most of the studies. The three exceptions are the NSDUH, NCVS, and CPS. For the NSDUH, refusals increased from 63

percent to 81 percent of nonresponse during 2000 to 2014; for the NCVS, refusals increased from 39 percent to 65 percent of nonresponse during 2000 to 2014; and for the CPS, the increase was from 41 percent to 60 percent of nonresponse for the period.

Even though refusals are not solely responsible for the increases in nonresponse, they account for such a large component of nonresponse that the changes in refusals are very important. For example, in the NHIS, the percentage of both refusals (7.3 percent to 17.6 percent) and noncontacts (3.8 percent to 8.6 percent) has doubled since 2000, but the effect of doubling the refusals is more dramatic. Furthermore, it is possible that some noncontact and nonresponse are actually veiled refusals. Households sometimes deliberately avoid contact or use non–English speakers in the household instead of refusing directly.

6. TRENDS IN LEVEL OF EFFORT

As noted in the previous section, the noncontact rate for most of the surveys is so low, almost always under 5 percent, that it is likely that numerous contact attempts are being made. Four of the studies have information on number of contacts available, but the time period covered is limited. The information was difficult to recover from project archives for most of these studies. The published series of contact attempts for the NSDUH is the longest one available.[6]

We report on screening interview results where initial contact is made with the household. The published data give the actual number of attempts with categories beyond five attempts (5–9 and 10+). To plot these data, we use the midpoint for the 5–9 category and "10" for the 10+ category. The NHIS first released data on contacts and other paradata for 2006 (National Center for Health Statistics 2008). The contact data from the MEPS and MCBS were obtained through operational data files provided by the projects.

Figure 7 shows the ratio of the number of contact attempts (of all sample cases) to the number of completed cases for the four surveys. We use this measure because it reflects all contact efforts for the achieved response rate. It is similar to approaches used to examine changes in effort for other modes, such as telephone (e.g., Curtin, Presser, and Singer 2000). We also reviewed the mean number of contacts per sample unit, and it exhibited the same trend over time, so it is not shown here.

For all four studies, the ratio of contacts per completed interview increases, with the NHIS and MCBS having larger nominal increases. Some of these differences are probably due to procedures (e.g., the NSDUH asks interviewers to close a case after five contacts unless the field supervisor decides to do more), and some may be due to other features (e.g., the decrease for the MEPS in

---

6. Data obtained from the NSDUH Data Collection Final Report for years 2000 through 2014.
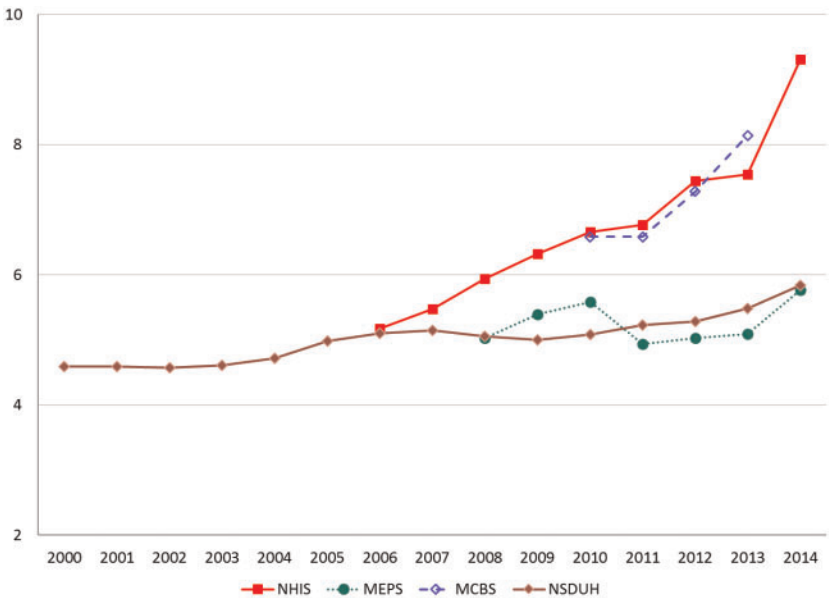
Figure 7. Ratio of Total Study Contact Attempts per Completed Interview. MCBS, Medicare Current Beneficiary Survey; MEPS, Medical Expenditure Panel Survey; NHIS, National Health Interview Survey; NSDUH, National Survey on Drug Use and Health.

2011 and 2012 may be due to changes to the sampling rates for domains). The sample source for the MEPS is from cooperating households (including partial completes) in the NHIS. This group is expected to have higher contact and response propensities.

7. CONCLUSION

This review shows increases in nonresponse in face-to-face surveys in the United States since 2000, continuing a long-standing trend. The magnitude of the decline in response rates differs by study; most studies have periods of decreases and other periods of stability. The data from the longitudinal surveys are not conclusive about trends in nonresponse for later waves, but there is at least some suggestion that re-interview rates too may be decreasing.

Researchers have noted declines for nearly four decades, yet until recently there was little cause for alarm since the rates for many face-to-face surveys remained relatively high and substantially higher than other modes could achieve. Now, nearly all the surveys in our analysis have rates of response lower than 80 percent, and several are below 70 percent. While our review is not exhaustive, we included a selection that covers multiple agencies,

organizations, and survey topics. These findings are consistent with other, more limited, investigations (see Meyer, Mok, and Sullivan 2015; Tourangeau and Plewes 2013). The problem of increasing nonresponse is widespread and a concern for our industry. Despite previous research that showed that the relationship between nonresponse and bias is not strong (Groves and Peytcheva 2008), there remains a strong desire to maximize response and control costs.

The increases in nonresponse that we observed were substantial, although not as dramatic as telephone surveys have experienced. However, nonresponse increases were greater for the later time period, suggesting that we have not yet reached a new plateau and that continued monitoring of nonresponse trends is warranted.

As with earlier efforts, surveys frequently made methodological and organizational changes during this period. We noted several situations where changes in methods coincided with changes in the rates. Many of these changes were interventions directed at addressing increases in nonresponse. Some, such as incentives, may have been effective. However, increases in response rates due to incentives are almost always temporary and are often followed by sharp declines. The innovative implementation of incentives for the GSS does not seem to follow this pattern. An alternative to giving the interviewers control over the incentive amount is to adjust incentive amounts frequently rather than waiting for response rates to fall to a worrisome level.

Refusals continue to be the main reason for nonresponse, but noncontact is also increasing. This suggests reluctance to participate, and increases in barriers to contact are at play. The increased availability of paradata from face-to-face surveys provided empirical evidence supporting the hypothesis that nonresponse is rising despite increased contact attempts. We posit that this increased effort may be fueling the increase in refusals. Increased contact resulting from these efforts does not necessarily reduce nonresponse bias (Fuchs, Bossert, and Stukowski 2013; Heerwegh, Abts, and Loosveldt 2007).

Although we report evidence of increased effort, the data are limited to four surveys, and the time covered is short. Data on level of effort are still difficult to obtain and generally not reported in survey methodological reports. We are encouraged that this is changing. Multiple surveys conducted by Census Bureau have adopted their contact history instrument, and we hope that other surveys will follow the NHIS example and make these data part of their public use data files.

Supplementary Materials

Supplementary materials are available online at academic.oup.com/jssam.

References

Agency for Healthcare Research and Quality (2010), *Respondent Payment Experiment with MEPS Panel 13*, Rockville, MD: Agency for Healthcare Research and Quality, Available at http://meps.ahrq.gov/mepsweb/data_files/publications/rpe_report/rpe_report_2010.shtml.

American Association for Public Opinion Research (1998), *Standard Definitions: Final Disposition of Case Codes and Outcome Rates for RDD Telephone Surveys and In-Person Household Surveys*, Ann Arbor, MI: AAPOR.

Atrostic, B. K., N. Bates, G. Burt, and A. Silberstein (2001), "Nonresponse in U.S. Government Household Surveys: Consistent Measures, Recent Trends, and New Insights," *Journal of Official Statistics*, 17, 209–226.

Barbier, S., G. Loosveldt, and A. Carton (2015), "The Flemish Survey Climate: An Analysis Based on the Survey of Social-Cultural Changes in Flanders," paper presented at the International Workshop on Household Survey Nonresponse, Leuven, Belgium.

Bates, N., J. Dahlhamer, P. Phipps, A. Safir, and L. Tan (2010), "Assessing Contact History Paradata Quality Across Several Federal Surveys," Proceedings of the Section on Survey Research Methods, pp. 91–105.

Biemer, P. P., P. Chen, and K. Wang (2013), "Using Level-of-Effort Paradata in Non-response Adjustments with Application to Field Surveys," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176, 147–168.

Blom, A. G. (2013), "Setting Priorities: Spurious Differences in Response Rates," *International Journal of Public Opinion Research*, 26, 245–255.

Bradburn, N. M. (1992), "Presidential Address: A Response to the Non-response Problem," *Public Opinion Quarterly*, 56, 391–398.

Brehm, J. (1994), "Stubbing Our Toes for a Foot in the Door? Prior Contact, Incentives, and Survey Response," *International Journal of Public Opinion Research*, 6, 45–63.

Brick, J. M., and D. Williams (2013), "Explaining Rising Nonresponse Rates in Cross-sectional Surveys," *The ANNALS of the American Academy of Political and Social Science*, 645, 36–59.

Bureau of Justice Statistics (2014), "National Crime Victimization Survey Technical Documentation: NCJ 247252," Available at http://www.bjs.gov/content/pub/pdf/ncvstd13.pdf.

Burton, J., H. Laurie, and P. Lynn (2006), "The Long-term Effectiveness of Refusal Conversion Procedures on Longitudinal Surveys," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 169, 459–478.

Centers for Disease Control and Prevention (2013), *National Health and Nutrition Examination Survey: Analytic Guidelines, 2011–2012*, Hyattsville, MD: National Center for Health Statistics.

Curtin, R., S. Presser, and E. Singer (2000), "The Effects of Response Rate Changes on the Index of Consumer Sentiment," *Public Opinion Quarterly*, 64, 413–428.

de Leeuw, E., and W. de Heer (2002), "Trends in Household Survey Nonresponse: A Longitudinal and International Comparison," in *Survey Nonresponse*, eds. R.M. Groves, D.A. Dillman, J.L. Eltinge, and R.J. A. Little, pp. 41–54, New York: Wiley.

Dillman, D. A., J. D. Smyth, and L. M. Christian (2014), Internet, Phone, Mail and Mixed-Mode Surveys: The Tailored Design Method. Hoboken, NJ: John Wiley and Sons, Inc.

Duncan, G. J., and G. Kalton (1987), "Issues of Design and Analysis of Surveys across Time," *International Statistical Review*, 55, 97–117.

Fisher, B., and M. Buha (2015), "Incentive Use Tracking and the Effect of Incentives on Interview Completion for the General Social Survey," paper presented at the 70th Annual Conference of the American Association for Public Opinion Research, Hollywood, FL.

Frankel, L. L., and D. S. Hillygus (2014), "Looking Beyond Demographics: Panel Attrition in the ANES and GSS," *Political Analysis*, 22(3), 336–353.

Fuchs, M., D. Bossert, and S. Stukowski (2013), "Response Rate and Nonresponse Bias-Impact of the Number of Contact Attempts on Data Quality in the European Social Survey," *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, 117, 26–45.

Groves, R. M., W. D. Mosher, J. M. Lepkowski, and N. G. Kirgis (2009), "Planning and Development of the Continuous National Survey of Family Growth," *Vital and Health Statistics*, 1(48), 1–64.

Groves, R. M., and E. Peytcheva (2008), "The Impact of Nonresponse Rates on Nonresponse Bias a Meta-analysis," *Public Opinion Quarterly*, 72, 167–189.

Heerwegh, D., K. Abts, and G. Loosveldt (2007), "Minimizing Survey Refusal and Noncontact Rates: Do Our Efforts Pay Off?" *Survey Research Methods*, 1, 3–10.

Kirgis, N. G., and J. M. Lepkowski (2013), "Design and Management Strategies for Paradata-Driven Responsive Design: Illustrations from the 2006–2010 National Survey of Family Growth," in *Improving Surveys with Paradata*, ed. F. Kreuter, Hoboken, NJ: John Wiley and Sons, Inc, 123–144.

Marquis, K. H. (1977), *Survey Response Rates: Some Trends, Causes and Correlates*, *The Rand Paper Series*, Santa Monica, CA: The Rand Corporation.

Meyer, B. D., W. K. C. Mok, and J. X. Sullivan (2015), "Household Surveys in Crisis," *Journal of Economic Perspectives*, 29, 199–226.

National Center for Health Statistics (2000), "1998 National Health Interview Survey (NHIS) public use data release survey description," Available at ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHIS/1998/srvydesc.pdf.

——— (2008), *Paradata Data File Documentation, National Health Interview Survey, 2006* (machine readable data file and documentation), Hyattsville, MD: National Center for Health Statistics, Centers for Disease Control and Prevention.

Office of Applied Studies (2002), *2001 National Household Survey on Drug Abuse: Incentive Experiment Combined Quarter 1 and Quarter 2 Analysis* (RTI 07190.388.100, prepared for the Office of Applied Studies, Substance Abuse and Mental Health Services Administration, by RTI under Contract No. 283-98-9008), Rockville, MD: Substance Abuse and Mental Health Services Administration.

Parsons, V. L., C. Moriarity, K. Jonas, T. Moore, K. E. Davis, and L. Tompkins (2014), "Design and Estimation for the National Health Interview Survey, 2006–2015," *Vital and Health Statistics*, 2(165), 1–53.

Pyy-Martikainen, M., and U. Rendtel (2008), "Assessing the Impact of Initial Nonresponse and Attrition in the Analysis of Unemployment Duration with Panel Surveys," *AStA Advances in Statistical Analysis*, 92, 297–318.

Schoeni, R. F., F. Stafford, K. A. McGonagle, and P. Andreski (2013), "Response Rates in National Panel Surveys," *The Annals of the American Academy of Political and Social Science*, 645, 60–87.

Shettle, C. F., P. Guenther, D. Kaspryzk, and M. E. Gonzalez (1994), "Investigating Nonresponse in Federal Surveys," Proceedings of the Section on Survey Research Methods, vol. II, American Statistical Association, pp. 972–976.

Smith, T. W. (1995), "Trends in Non-response Rates," *International Journal of Public Opinion Research*, 7, 157–171.

——— (2006), "The Subsampling of Nonrespondents on the 2004 General Social Survey, GSS Methodological Report No. 106," National Opinion Research Center/University of Chicago.

——— (2011), "A Review of the Use of Incentives in Surveys," NORC Report, National Opinion Research Center/University of Chicago.

Smith, T. W., P. V. Marsden, and M. Hout (2015), *General Social Surveys, 1972–2014, Codebook*, Chicago: National Opinion Research Center.

Steeh, C. G. (1981), "Trends in Nonresponse Rates, 1952–1979," *Public Opinion Quarterly*, 45, 40–57.

Stoop, I., J. Billiet, A. Koch, and R. Fitzgerald (2010), *Improving Survey Response: Lessons Learned from the European Social Survey*, Chichester, UK: John Wiley and Sons, Ltd.

Substance Abuse and Mental Health Services Administration (2014), National Survey on Drug Use and Health: Data Collection Final Report, Rockville, MD: Substance Abuse and Mental Health Services Administration, Available at https://www.samhsa.gov/data/population-data-nsduh/reports.

Synodinos, N. E., and Y. Shigeru (1999), "Public Opinion Surveys in Japan," *International Journal of Public Opinion Research*, 6, 118–138.

Tourangeau, R. (2004), "Survey Research and Societal Change," *Annual Review of Psychology*, 55, 775–801.

Tourangeau, R., and T. J. Plewes (2013), *Nonresponse in Social Science Surveys: A Research Agenda*, Washington, DC: The National Academies Press.

U.S. Census Bureau (2006), "Current Population Survey Design and Methodology," Technical Paper 66, Bureau of the Census, Washington, DC.

Virgile, M. (2016), "Measurement Error in American Community Survey Paradata and 2014 Redesign of the Contact History Instrument," Center for Statistical Research and Methodology Research Report Series (Survey Methodology #2016-01), U.S. Census Bureau, Available at https://www.census.gov/srd/papers/pdf/RSM2016-01.pdf.