*Journal of Survey Statistics and Methodology* (2022) **00**, 1–22

# THE EFFECTS OF A TARGETED "EARLY BIRD" INCENTIVE STRATEGY ON RESPONSE RATES, FIELDWORK EFFORT, AND COSTS IN A NATIONAL PANEL STUDY

KATHERINE A. MCGONAGLE*
NARAYAN SASTRY
VICKI A. FREEDMAN

Adaptive survey designs are increasingly used by survey practitioners to counteract ongoing declines in household survey response rates and manage rising fieldwork costs. This paper reports findings from an evaluation of an early-bird incentive (EBI) experiment targeting high-effort respondents who participate in the 2019 wave of the US Panel Study of Income Dynamics. We identified a subgroup of high-effort respondents at risk of nonresponse based on their prior wave fieldwork effort and randomized them to a treatment offering an extra time-delimited monetary incentive for completing their interview within the first month of data collection (treatment group; $N = 800$) or the standard study incentive (control group; $N = 400$). In recent waves, we have found that the costs of the protracted fieldwork needed to complete interviews with high-effort cases in the form of interviewer contact attempts plus an increased incentive near the close of data collection are extremely high. By incentivizing early participation and reducing the number of interviewer contact attempts and fieldwork days to complete the interview, our goal was to manage both nonresponse and survey costs. We found that the EBI

KATHERINE A. MCGONAGLE is a Research Professor, NARAYAN SASTRY is a Research Professor, and VICKI A. FREEDMAN is a Research Professor with the Institute for Social Research, University of Michigan, 426 Thompson St., P.O. Box 1248, Ann Arbor, MI 48106, USA.This work was supported by the National Science Foundation [SES 1623864], the Eunice Kennedy Shriver National Institute of Child Health and Human Development [R01 HD069609], and the National Institute on Aging [R01 AG040213]. The authors are grateful for the contributions of Wen Chang, Minako Edgar, and Shonda Kruger-Ndiaye for implementing the experimental intervention and leading fieldwork activities. The deidentified data and code used for this article are available through the ICPSR PSID Data Repository (McGonagle, Sastry, and Freedman 2021).

*Address correspondence to Katherine A. McGonagle, Institute for Social Research, University of Michigan, 426 Thompson St., P.O. Box 1248, Ann Arbor, MI 48106, USA; E-mail: kmcgon@umich.edu.

doi: 10.1093/jssam/smab042

**Statement of Significance**

This paper reports findings from a targeted incentive experiment embedded in a long-running, nationally representative panel study of US families designed to reduce field-work effort and minimize nonresponse of high-effort respondents. Compared to the standard incentive protocol, offering a time-delimited extra incentive at the onset of fieldwork to respondents who required the highest levels of effort and fieldwork duration in the prior wave reduced fieldwork duration and interviewer effort. Response rates for respondents receiving the experimental treatment were significantly higher at the conclusion of fieldwork than for those receiving the standard incentive protocol. The incentive treatment had wide appeal across a range of respondent demographic, social, and economic characteristics and achieved slight cost savings by the end of data collection.

treatment increased response rates and reduced fieldwork effort and costs compared to a control group. We review several key findings and limitations, discuss their implications, and identify the next steps for future research.

KEYWORDS: Adaptive design; Fieldwork effort; Incentives; Nonresponse error; Panel study; Response rate; Survey costs.

## 1. INTRODUCTION

Survey organizations are increasingly using adaptive designs to counteract the steady decline in response rates of household surveys over the past quarter-century (e.g., National Research Council 2013; Beullens, Loosveldt, Vandenplas, and Stoop 2018; de Leeuw, Hox, and Luiten 2018; Williams and Brick 2018). In interviewer-administered surveys, such declines are attributable to increases in both refusals and difficulty making contact with respondents (e.g., Groves and Couper 1998) stemming from a variety of technological, societal, and cultural changes. Adaptive survey designs attempt to offset these trends by targeting subgroups of sample members for special data collection procedures with the key aims of maximizing response rates, reducing nonresponse bias, and effectively expending survey budgets (see Schouten, Calinescu, and Luiten 2013; Tourangeau, Brick, Lohr, and Li 2017).

In a targeted adaptive design, subgroups of sample members are selected to receive a variant of the standard data collection protocol with the goal of reducing survey error while balancing survey costs (see Lynn 2014, 2017). The focus of targeted designs is typically on the reduction of nonresponse error by providing special treatment to cases that have low response propensity (Lynn 2017). With data collected on the same respondents over multiple waves, longitudinal studies are especially well-suited for the use of targeted design elements. This emphasis reflects the importance of minimizing nonresponse in panel studies, because over time even minor attrition can lead to sample bias, a reduction in the number of observations made on each case, and potential loss in the scientific value of the data (Lugtig 2014).

A variety of design elements have been targeted to sample subgroups to reduce variation in response propensity, including extra interviewer contact attempts (Wagner, West, Kirgis, Lepkowski, Axinn, et al. 2012), interviewer bonuses (Peytchev, Riley, Rosen, Murphy, and Lindblad 2010), call window alternatives (Luiten and Schouten 2013; Kreuter and Müller 2015; McGonagle and Sastry 2020), and respondent messaging (Fumagalli, Laurie, and Lynn 2012; Lynn 2016). A strategy widely used in panel studies is tailoring the timing and amounts of study incentives. For instance, near the close of data collection several ongoing panel studies routinely target cases that remain incomplete for "end-game" strategies (e.g., McGonagle and Freedman 2017; see Schoeni, Stafford, McGonagle, and Andreski 2013). This may include special messages alerting respondents to a fixed study end date and an offer of a time-delimited extra incentive in exchange for completing the interview by the end of the study. While there is some evidence that this strategy is successful in reducing nonresponse, it does so at a high cost, targeting cases late in fieldwork after numerous and unsuccessful interviewer contact attempts have been expended (McGonagle 2020).

Offering time-delimited incentives to low-propensity cases earlier in fieldwork may achieve a better balance between the reduction of nonresponse error and survey costs. However, only a handful of studies have implemented such incentives, and much remains unknown about their effectiveness and efficacy. Fomby, Sastry, and McGonagle (2017) examined both nonresponse and survey costs using an experimental design to offer difficult-to-reach caregivers a time-delimited incentive in exchange for completing the interview during a traditionally slow holiday period. They found that the extra incentive was cost-efficient and provided a statistically significant increase in completed interviews with no negative effects on final response rates. Three other panel studies used experimental designs to test the effects of time-delimited incentives offered at the start of fieldwork, but none evaluated survey costs. Coopersmith, Vogel, Bruursema, and Feeney (2014) found that an extra incentive offered during the first three weeks of data collection to induce the early participation of a web-based panel of school principals significantly reduced fieldwork duration and had no negative effects on final response rates after the offer was withdrawn. In an experiment conducted on the UK Household Longitudinal Survey Innovation Panel, Brown and Calderwood (2014) found that an early-bird incentive (EBI) offer, especially when combined with a higher study incentive, was modestly successful in reducing fieldwork effort by encouraging respondents to schedule interview appointments. Finally, LeClere, Plummer, Vanicek, Amaya, and Carris (2012) offered a random sample of families participating in a mixed-mode community-based study an additional incentive for study completion within one week of interviewer contact and found significantly higher response rates both during the offer period and by the end of data collection.

While these few studies provide promising evidence for the effectiveness of targeting early time-delimited incentives to particular respondent subgroups, the generalizability of these findings is unclear. Moreover, there has been little attention in the literature to the balance between potential reductions in nonresponse error provided by a targeted intervention such as an extra incentive and the cost of its implementation (Tourangeau et al. 2017). In addition to implementation costs, the overall impact on survey costs will depend on the proportion of targeted respondents who respond to the extra incentive and the difference between the incentive costs and the amount of fieldwork effort saved. A driving factor for both nonresponse error and survey cost is the behavior of respondents who by choice or circumstance do not take up the extra incentive during the offer period. As others have suggested (Peycheva, Calderwood, and Wong 2019), these respondents may be discouraged by the withdrawal of the extra incentive and become even less receptive to interviewer attempts and other adaptive fieldwork strategies, such as an end-game incentive offer. As cases remain active, each subsequent interviewer attempt incurs a marginal cost, with each day requiring managerial resources independent of interviewer contact attempts, such as case review and monitoring. In this scenario, the additional fieldwork effort to minimize nonresponse could reduce or even outweigh any cost savings of an early extra incentive.

Our paper reports findings from an evaluation of an early-bird incentive (EBI) experiment embedded in the 2019 wave of the Panel Study of Income Dynamics (PSID). The experiment targeted reinterview cases (i.e., those who completed an interview in the prior wave) that were "high effort" in the past and hence represented a subgroup of sample members at increased risk for nonresponse. The treatment was to offer these sample members a time-delimited monetary incentive that was substantially higher than the standard incentive amount. In recent waves, we have found that the cost of protracted fieldwork needed to complete interviews with these cases—which typically includes numerous interviewer contact attempts over many months plus an additional incentive during end-game—is extremely high. By incentivizing early participation and reducing the number of interviewer contact attempts and fieldwork days to complete the interview, our goal was to manage both nonresponse and survey costs.

## 2. RESEARCH QUESTIONS

We examine four research questions with specific hypotheses.

RQ1: First, what is the impact of a time-delimited EBI offer on cumulative and conditional response rates at three phases of fieldwork: the offer period, main production, and the end of production? Consistent with the handful of studies finding positive effects of time-delimited incentive offers, we hypothesize that the cases assigned to receive the EBI offer (treatment group) will have higher response rates during the offer period than cases receiving the standard incentive (control group). By the end of fieldwork, we hypothesize that there will be no difference between the experimental conditions in response rates. This is based on our expectation that any "backlash effect" such that treatment cases become less cooperative when the EBI offer is withdrawn will be offset by the two opportunities for extra incentives over the course of data collection—during the EBI offer period, and again as an adaptive end-game strategy applied to all remaining cases shortly before the end of fieldwork.

RQ2: Second, what is the impact of the EBI on fieldwork effort as assessed by the total number of interview contact attempts and days in the field needed to finalize the interview? Building on our hypothesis that treatment cases will be more likely to respond during the EBI offer period, we expect that treatment cases will consequently have significantly fewer interviewer contact attempts and days of fieldwork than control cases over the course of data collection.

RQ3: Third, how does response to the EBI treatment affect the composition of the responding sample? We expect the EBI treatment to have no overall effect on sample composition, consistent with our expectation that there will be no overall treatment effect by the end of the field period on response rates.

RQ4: Finally, what are the cost implications of the EBI treatment? We compare the costs of the treatment group with the control group and hypothesize that the higher incentive costs expended for the treatment group will be offset by the savings in fieldwork resources due to the earlier response of treatment cases, yielding a net savings for the EBI intervention.

## 3. METHODS

### 3.1 Study Design of the PSID

The experiment was embedded in the PSID, a longitudinal study of a nationally representative sample of US families. The main interview is about 75 minutes on average and collects a variety of data on economic, health, and social behavior. The primary mode of data collection is computer-assisted telephone interview (CATI) by professional interviewers employed by the Survey Research Center at the University of Michigan. Families in the PSID have been interviewed annually 1968–1997, and biennially since 1997 (see McGonagle, Schoeni, Sastry, and Freedman [2012] for more information).

Data collection occurs in odd-numbered years between about March 1 and December 31 over the course of nearly 44 weeks. December 31 is a firm end date because the questionnaire references specific periods within the current calendar year. The study interviews one adult respondent in each family, typically the individual who is most knowledgeable about the family finances. During the 2019 wave, interviews were completed with 9,444 families with a 93 percent overall reinterview response rate.

Like most panel studies, postpaid monetary incentives are provided to respondents. The baseline incentive corresponds to a payment of approximately $1 per minute for the mean interview duration and is highlighted in an informational letter sent to sample members prior to the start of data collection. Subsequent interview requests, which are made through telephone calls, postal mail, email, and text message, frequently refer to the incentive. Starting in 2015, cases still nonfinal with six to eight weeks remaining in the field period have been offered double the baseline incentive. This end-game incentive is offered to approximately 17–20 percent of cases, with a slight increase in the percentage each wave reflecting the growing difficulty of making contact with respondents in telephone studies (e.g., de Leeuw, Hox, and Luiten 2018). While this end-game strategy has helped the study to achieve target response rates, fieldwork costs for these cases are substantial—in 2015, more than half accepted the $150 incentive offer—*after* receiving approximately eighty interviewer contact attempts on average (see McGonagle 2020). In 2019, in addition to implementing the baseline and end-game incentives, we experimentally incorporated a new approach that shifted the higher end-of-study incentive for difficult cases to the beginning of the field period.

## 3.2 Experimental Design

### *3.2.1 Sample selection.*

The EBI experiment targeted high-effort reinterview cases at risk of nonresponse drawn from two subgroups. The first group consisted of cases that in the prior wave required the highest number of interviewer contact attempts and received the end-game incentive. A second group comprised high-effort cases who were eligible for two upcoming supplemental studies of children and young adults by virtue of having age-eligible family members. These cases were of high priority as eligibility for the supplements depended upon successful completion of the main interview, with the supplementary studies benefitting from these cases completing their main interviews early. Thus, minimizing nonresponse in the main study for these cases increased the pool of eligible sample for the supplements.

Random assignment to the experimental conditions was conducted separately for the two groups. The first group consisted of 1,900 cases at or above the 75th percentile of average interviewer attempts in 2017. From this pool,

**Table 1. Assignment to Experimental Condition by Sample Type**

| Sample type | Total | Treatment | Control |
|---|---|---|---|
| Total | 1,189[a] | 792 | 397 |
| Not eligible for both supplements | 597 | 397 | 200 |
| Eligible for both supplements | 592 | 395 | 197 |

[a]$N = 11$ cases were found to be nonsample and excluded from the study.

600 cases were randomly selected for the experiment. The second group comprised 842 cases that were eligible for both of the two supplemental studies. From this pool, the 600 cases with the highest number of interviewer attempts in 2017 were drawn. The 1,200 cases from the two groups were pooled and randomly assigned to a treatment group ($N = 800$ cases) or a control group ($N = 400$ cases). We randomized a greater number to receive the treatment to reflect our expectation that the EBI would reduce fieldwork effort and duration. Following random assignment, $N = 11$ cases were found to be nonsample and excluded from the experiment (see table 1). An extra time-delimited incentive was offered to cases in the treatment group and the standard study incentive was offered to cases in the control group.

*3.2.2 Experimental manipulation.*

The experimental intervention consisted of a colorful note attached to the front of the advance letter to respondents for the treatment group only describing the


**Figure 1. Early-Bird Incentive Offer.**

increased (doubled) incentive that would be provided if the interview was completed within the first 31 days of the study (see figure 1; "Complete your interview by March 31 to receive an extra $75 for a total of $150!"). The letter sent to the control group was identical except that it did not include the additional early-bird note.

Interviewers were aware that an experiment was being conducted and not blinded to the assignment of cases to the treatment and control groups. However, interviewers were not aware of the nature or purpose of the experiment, only that some cases would be offered different incentive payments. During the main study training and in subsequent team meetings, interviewers were instructed to work all cases with the same level of effort.

## 3.3 Measures

### *3.3.1 Outcome measures.*

We examine three fieldwork-related outcomes: (1) *Response rates*, calculated as the percentage of eligible respondents completing an interview (based on definition RR6; The American Association for Public Opinion Research [AAPOR] 2016), measured at three points during fieldwork (by the end of the one-month EBI offer period, by the end of mid-production, and by the end of fieldwork); (2) *Total number of interviewer contact attempts* required to finalize the case (i.e., complete an interview or apply a final refusal disposition), constructed as the sum of telephone calls, emails, and text messages; and (3) *Fieldwork duration*, defined as the number of days from the release date of the case for an individual respondent (i.e., "case") to receive a final fieldwork disposition.

### *3.3.2 Respondent characteristics.*

We examined the characteristics of respondents assigned to the treatment and control groups to ensure randomization to the experimental conditions and to explore the impact of the EBI treatment on sample composition. The variables include: (i) *female respondent*, (ii) *age of the respondent* (in continuous years), (iii) *whether married or cohabiting* ("yes" = 1, "no" = 0); (iv) *whether children reside in the household* ("yes" = 1, "no" = 0); (v) a series of dummy variables using percentile cut-points for *2017 family income* (25th, 50th, 75th, 90th; "yes" = 1, "no" = 0); (vi) *mutually exclusive categories for respondent self-reported racial identity* ("white, non-Hispanic," "African-American, non-Hispanic," "Other race," "yes" = 1, "no" = 0), (vii) whether the respondent residence was in a *metropolitan location* as defined by the Beale-Ross Rural-Urban Continuum Code ("yes" = 1, "no" = 0), (viii) mutually exclusive categories for *sample type* for cases who were selected from a national probability sample ("yes" = 1, "no" = 0), a low-income oversample ("yes" = 1, "no" = 0), or an immigrant refresher sample ("yes" = 1, "no" = 0); and (ix) *low* *effort-level in 2017*, an indicator variable designed to capture cases at the lowest level of prior wave effort (i.e., the 25th percentile or lower of interviewer contact attempts), among the high-effort respondents in the experiment ("yes" = 1, "no" = 0). This variable was constructed separately for the group of respondents with prior wave effort at or above the 75th percentile, and the high-effort group of respondents who were eligible for both supplemental studies.

## 3.4 Analysis Strategy

Following confirmation that cases were correctly randomized to each experimental condition, we address RQ1 and RQ2 by testing for differences between the treatment group and control group in fieldwork-related outcomes, including response rates (RQ1) and the total number of interviewer contact attempts to finalize the case and the number of days of fieldwork (RQ2). We test mean differences using *t*-tests and median differences using quantile regression. We also test differences in time-to-completion of the interview (across the entire field period and by percentiles of case completion) using a Kaplan–Meier nonparametric survival model (Kaplan and Meier 1958).

We investigate RQ3 by examining heterogeneous treatment effects of the EBI intervention on sample composition and test mean differences using *t*-tests and estimate three separate multivariate logistic regression models predicting completion of the interview (1) by the end of the EBI offer period, (2) by the end of mid-production, and (3) by the end of fieldwork. Each model includes main effects for the set of respondent characteristics described in section 3.3.2, an indicator for assignment to the treatment group (versus control) and a multiplicative interaction term between the treatment indicator and each respondent characteristic to test if the chances of responding differ for treatment and control group members with different characteristics.

Finally, to address RQ4, we evaluate the cost implications of the EBI treatment by comparing major variable costs incurred during data collection separately by experimental condition, based on the response rates, average number of interviewer attempts, and incentive costs for each group. We provide estimates of interviewer contact attempt costs based on interviewer hourly wages and their expected productivity. In consultation with the study survey director, we assume that interviewers typically make four contact attempts per hour. We also assume that attempt types (telephone, email, text message) require approximately the same amount of time, including to dial and leave a voice-mail message or type, review, and send a text message or email message, as well as time spent reviewing notes summarizing the case history (e.g., information about the respondent and best times to make contact) and results of prior contact attempts (e.g., the number, dates, and outcomes of prior contact attempts) and postattempt updating of case history notes. We derive a per-attempt cost of

$6.56 based on the average hourly wage and fringe benefits of an interviewer (approximately $26) and four attempts per hour. Total interviewer contact attempt costs are derived by multiplying the cost per attempt and number of cases by the average number of total attempts based on separate calculations for completed interviews and nonresponse cases. We calculate incentive costs for completed interviews (nonresponse cases have zero incentive costs), for each phase of data collection by experimental condition. Finally, we generate grand total costs by summing the costs for interviewer attempts and incentives and provide a cost per case estimate by experimental condition.

Because there were no differences in the pattern of results based on analyses stratified by the two sample subgroups, analyses presented in this paper pool the two groups and refer to them as 'high-effort' cases.

## 4. RESULTS

### 4.1 Randomization Check

We checked to make sure that the randomization of respondents to the experimental groups was successful. Details of the tests we performed are provided in a supplementary appendix.

### 4.2 RQ1: What Is the Effect of the EBI Treatment on Response Rates?

The first research question examines the effect of the EBI treatment by comparing response rates for the experimental conditions across the three phases of production: the early-bird offer period, main production, and the end-game offer period. The top panel of table 2 shows the cumulative unconditional response rates for each group by the end of the three phases. As predicted, by the end of the EBI offer period, treatment cases had a significantly higher response rate than control cases, with nearly 56 percent of treatment cases completing the interview compared to 15.4 percent of control cases—a large and statistically significant difference of 40.4 percentage points ($p \leq .0001$). By the end of main production, the difference in cumulative response rates narrowed, with a response rate of 80.2 percent for treatment cases and 73.8 percent for control cases ($p \leq .05$). By the end of the field period, the response rate for the treatment group was 89.3 percent compared to 85.4 percent for the control group, a difference of nearly 4.0 percentage points ($p \leq .05$).

The middle panel of table 2 shows incremental response rates in each discrete phase of data collection (i.e., the proportion responding in each phase among all cases). During main production, nearly one-quarter of treatment cases completed an interview after the EBI expired compared to about 58 percent of control cases ($p \leq .0001$). The higher completion rate among control

**Table 2. Response Rates (RRs) by Experimental Condition**

| | Treatment group ($N = 792$) | | | Control group ($N = 397$) | | | Difference | |
|---|---|---|---|---|---|---|---|---|
| | No. of active cases | No. of completed interviews | RR (%) | No. of active cases | No. of completed interviews | RR (%) | RR % Δ | $p$ of Δ |
| Cumulative (unconditional) RRs by end of phase | | | | | | | | |
| By end of EBI offer period (2/28–3/31) | 792 | 442 | 55.8 | 397 | 61 | 15.4 | 40.4 | *** |
| By end of main production (2/28–11/14) | 792 | 635 | 80.2 | 397 | 293 | 73.8 | 6.4 | * |
| By end of fieldwork (2/28–12/31) | 792 | 707 | 89.3 | 397 | 339 | 85.4 | 3.9 | * |
| RRs during each phase | | | | | | | | |
| Incremental RRs | | | | | | | | |
| During EBI offer period (2/28–3/31) | 792 | 442 | 55.8 | 397 | 61 | 15.4 | 40.4 | *** |
| During main production (4/1–11/14) | 792 | 193 | 24.4 | 397 | 232 | 58.4 | −34.1 | *** |
| During end-game offer period (11/15–12/31) | 792 | 72 | 9.1 | 397 | 46 | 11.6 | −2.5 | ns |
| Conditional RRs | | | | | | | | |
| During EBI offer period (2/28–3/31) | 792 | 442 | 55.8 | 397 | 61 | 15.4 | 40.4 | *** |
| During main production (4/1–11/14) | 350 | 193 | 55.1 | 336 | 232 | 69.0 | −13.9 | ** |
| During end-game offer period (11/15–12/31) | 157 | 72 | 45.9 | 104 | 46 | 44.2 | 1.6 | ns |

*$p \leq .05$,
**$p \leq .001$,
***$p \leq .0001$.

cases during this phase is expected since a much higher percentage remained active after the EBI offer period compared to the treatment group. The incremental end-game response rates were not significantly different for the two groups (9.1 percent versus 11.6 percent).

The bottom panel of table 2 recalculates the response rates within each phase, using the number of cases still not complete at the beginning of the phase as the denominator for the conditional response rate. Among those who had not completed the interview by the beginning of the main production period, 55.1 percent of treatment cases and 69.0 percent of control cases completed the interview during main production ($p \leq .05$). The conditional end-game response rates were not significantly different for the two groups (45.9 percent versus 44.2 percent).

## 4.3 RQ2: What Is the Impact of the EBI Treatment on Fieldwork Effort?

We next examine our second RQ and hypothesis predicting lower fieldwork effort for treatment cases relative to control cases by comparing interviewer contact attempts and fieldwork duration.

### *4.3.1 Interviewer contact attempts.*

Table 3 compares the mean and median number of interviewer contact attempts for the experimental conditions across the phases of data collection. The top panel shows the cumulative unconditional interviewer attempts for each group by the end of each phase. By the end of the EBI offer period, treatment cases received significantly more interviewer attempts at the mean (3.0 more attempts, $p \leq .0001$) and median (3.0 more attempts, $p \leq .0001$) than control cases. This trend was reversed by the end of main production, with treatment cases receiving significantly fewer attempts than control cases (nearly 8.0 fewer attempts, $p \leq .0001$). As predicted, by the end of data collection, treatment cases received significantly fewer interviewer contact attempts than control cases (nearly 9.0 fewer mean attempts, $p \leq .0001$ and 12.0 fewer median attempts, $p \leq .0001$).

Increments of interviewer contact attempts by discrete phases of production are shown in the middle panel of table 3. On average, treatment cases received about 42 percent fewer attempts during main production (nearly 11.0 fewer attempts, $p \leq .0001$) and 23 percent fewer attempts during end-game (1.1 fewer attempts, $p \leq .05$) than control cases. There were no significant differences between the groups in median incremental attempts during end-game. Conditional interviewer attempts (bottom panel) among cases still active during main production and end-game were not significantly different for the two groups.

**Table 3. Interviewer Contact Attempts by Experimental Condition**

| | Treatment group ($N = 792$) | | | Control group ($N = 397$) | | | Difference | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | No. of active cases | Mean | Med | No. of active cases | Mean | Med | Mean Δ | *p* of Δ | Med Δ | *p* of Δ |
| Cumulative (unconditional) interviewer attempts by end of phase | | | | | | | | | | |
| By end of EBI offer period (2/28–3/31) | 791 | 9.5 | 9.0 | 393 | 6.5 | 6.0 | 3.0 | ** | 3.0 | ** |
| By end of main production (2/28–11/14) | 791 | 24.3 | 13.0 | 393 | 32.1 | 24.0 | −7.8 | ** | −11.0 | ** |
| By end of fieldwork (2/28–12/31) | 791 | 27.9 | 13.0 | 393 | 36.8 | 25.0 | −8.9 | ** | −12.0 | ** |
| Interviewer attempts during each phase | | | | | | | | | | |
| Incremental attempts | | | | | | | | | | |
| During EBI offer period (2/28–3/31) | 791 | 9.5 | 9.0 | 393 | 6.5 | 6.0 | 3.0 | ** | 3.0 | ** |
| During main production (4/1–11/14) | 791 | 14.8 | 0.0 | 393 | 25.5 | 17.0 | −10.7 | ** | −17.0 | ** |
| During end-game offer period (11/15–12/31) | 791 | 3.6 | 0.0 | 393 | 4.7 | 0.0 | −1.1 | * | 0.0 | ns |
| Conditional attempts | | | | | | | | | | |
| During EBI offer period (2/28–3/31) | 791 | 9.5 | 9.0 | 393 | 6.5 | 6.0 | 3.0 | ** | 3.0 | ** |
| During main production (4/1–11/14) | 349 | 33.5 | 31.0 | 332 | 30.2 | 25.0 | 3.3 | ns | 6.0 | ns |
| During endgame offer period (11/15–12/31) | 156 | 18.1 | 17.5 | 103 | 17.9 | 17.0 | 0.2 | ns | 0.5 | ns |

NOTE.—Data are missing for $n = 1$ treatment case and $n = 4$ control cases.
*$p \leq .05$,
**$p \leq .0001$.

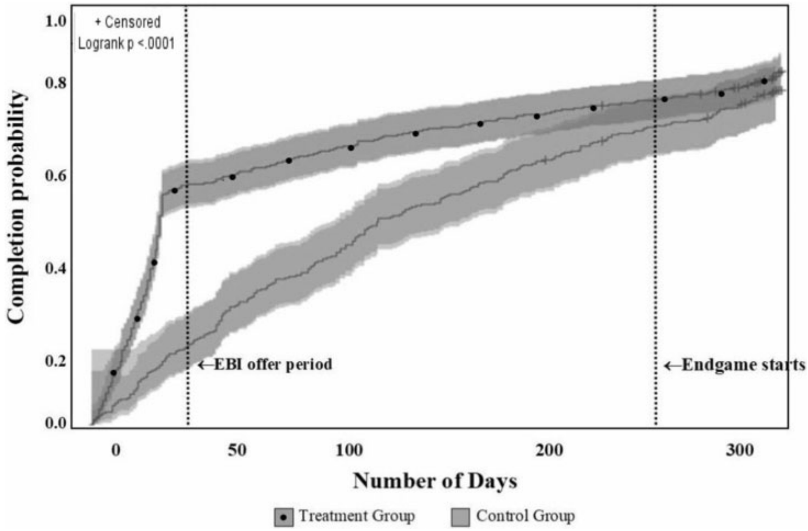


**Figure 2. Timing of Completing the Interview by Experimental Condition.**

*4.3.2 Days in the field.*

Consistent with our hypothesis, treatment cases required fewer days of fieldwork than control cases. The treatment group was in the field an average of nearly fifty fewer days (7.1 weeks) compared to the control group (101.0 versus 150.9 days, respectively, $p \leq .0001$). Tests of differences in Kaplan–Meier curves by experimental condition (see figure 2) are statistically significant (log-rank estimator $\chi^2 = 35.8(1)$, $p \leq .0001$), with the treatment group (dotted line) completing the interview at a significantly faster rates than the control group. More than 50 percent of treatment cases completed their interview in the first month (by day 31), compared to the fourth month (by day 123) for the control group ($p \leq .0001$). Nearly 9 months (by day 264) were needed for the control group to achieve the completion of 75 percent of its cases, significantly longer than the 6 months required for the treatment group (by day 185, $p \leq .001$). After day 185, the gap in completion rates between the groups narrowed.

## 4.4 RQ3: Does Response to the EBI Treatment Affect the Composition of the Responding Sample?

Our evaluation of potential heterogeneous effects of the EBI intervention revealed an especially large differential effect. High-income respondents (at or above the 90th percentile in the total sample) were highly responsive to the EBI, with a large, positive, and statistically significant interaction effect between high income and the EBI treatment (OR = 3.9, $p \leq .05$). As shown in

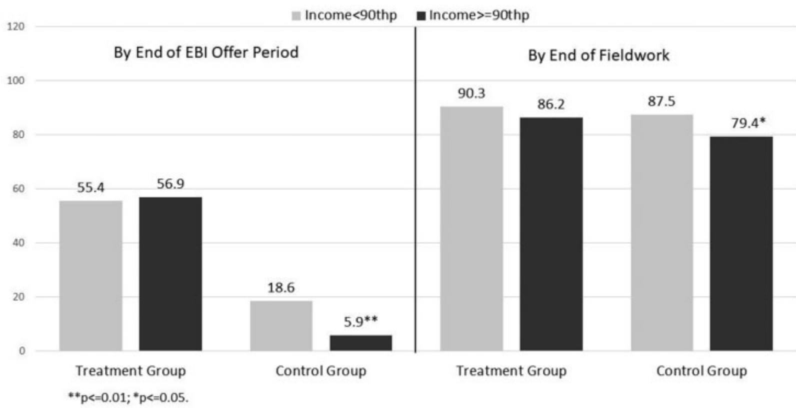


**Figure 3. Response Rates by Income Status and Experimental Condition.**

figure 3, the total effect of the EBI was similar for high-income and lower-income respondents, who had rates of early response that did not differ (56.9 percent versus 55.4 percent, respectively, $p = \text{NS}$). The significant interaction effect was due to the substantial difference within the control group, with early response by only 5.9 percent of high-income respondents compared to 18.6 percent of lower-income respondents ($p \leq .001$). By the end of fieldwork, the income difference within the control group remained, with significantly fewer high-income respondents completing an interview compared to lower-income respondents (79.4 percent versus 87.5 percent, $p \leq .05$). High-income treatment cases had a final response rate nearly seven percentage points greater than high-income control cases, although the difference was not statistically significant (86.2 percent versus 79.4 percent, $p = .13$). This effect may have achieved statistical significance with a larger sample size and more statistical power.

We evaluated alternative income percentile cut-points (25th, 50th, 75th) and the effects of income when specified as a continuous or a nonlinear variable to capture variation in responsiveness to the treatment by income. There was no evidence of variation in the effects of the treatment based on these alternative income specifications. None of the other respondent characteristics altered the effects of the EBI intervention during the offer period, by the end of mid-production, or by the end of the field period.

## 4.5 RQ5: What Are the Cost Implications of the EBI Treatment?

The final research question addresses the cost-effectiveness of the EBI strategy. Table 4 presents the results of a basic evaluation of the EBI intervention by comparing the two main sources of cost for completed interviews and nonresponse cases separately by experimental condition.

**Table 4. Cost Evaluation of the EBI Intervention**

| | Treatment group | Control group | Difference |
|---|---|---|---|
| Number of total cases | 791 | 393 | |
| Average cost per interviewer attempt | $6.56 | | |
| Cost based on completed interviews | $N = 706$ | $N = 335$ | |
| Interviewer contact attempts | | | |
| Interviewer attempts (mean) | 21.9 | 30.1 | |
| Total cost | $101,427 | $66,148 | $35,279 |
| Total cost per case | $144 | $197 | −$54 |
| Incentive costs | | | |
| Early-bird offer ($150 treatment, $75 control) | $66,300 | $4,575 | $61,725 |
| Main production ($75, all cases) | $14,400 | $17,400 | −$3,000 |
| End-game offer ($150, all cases) | $10,800 | $6,900 | $3,900 |
| Total cost | $91,500 | $28,875 | $62,626 |
| Total cost per case | $130 | $86 | $43 |
| Cost based on nonresponse cases | $N = 85$ | $N = 58$ | |
| Interviewer contact attempts | | | |
| Interviewer attempts (mean) | 77.7 | 75.5 | |
| Total cost | $43,326 | $28,726 | $14,599 |
| Total cost per case | $510 | $495 | $14 |
| Incentive costs | $0 | $0 | $0 |
| Grand total costs | | | |
| Interviewer contact attempts | | | |
| Total cost | $144,752 | $94,874 | $49,878 |
| Total cost per case, all cases | $183 | $241 | −$58 |
| Incentive costs | | | |
| Total cost | $91,500 | $28,875 | $62,625 |
| Total cost per case, all cases | $130 | $86 | $43 |
| Grand total costs | $236,252 | $123,749 | $112,503 |
| Grand total cost per case | | | |
| All cases | $299 | $315 | −$16 |
| Interview cases | $273 | $284 | −$11 |

Consistent with our hypothesis, the net difference between the costs due to interviewer contact attempts and the incentive yielded modest savings for the EBI intervention compared to the control group across all cases of about 5.4 percent or $16 per case ($299 versus $315) and about 4.0 percent or $11 per interview case ($273 versus $284). These savings were due to high early completion of the interview by the treatment group that led to fewer interviewer contact attempts the costs of which outweighed modestly higher incentive costs.

## 5. DISCUSSION

This paper describes an experimental evaluation of an adaptive incentive strategy embedded in a nationally representative panel study. We identified a subgroup of high-effort cases at risk of nonresponse based on their prior wave fieldwork effort and randomized them to a treatment offering an extra time-delimited monetary incentive or the standard study incentive at the beginning of fieldwork. The intervention goals were to save costs by reducing the number of interviewer contact attempts and active days of fieldwork while minimizing nonresponse. We review several key findings and limitations, discuss their implications, and identify next steps for future research.

First, consistent with our expectations, we found that the EBI treatment was successful in inducing high-effort respondents to participate at a much higher rate during the first month of data collection compared to those offered the standard incentive. Moreover, over the ensuing months of fieldwork, the targeted respondents remained cooperative, with a substantial proportion completing the interview for the standard incentive, and taking up the end-game incentive at rates as high as control cases. By the end of data collection, the final response rate of the treatment group exceeded the control group by four percentage points, an unanticipated and positive outcome of the intervention.

Second, as expected, the treatment improved the fieldwork outcomes of high-effort cases compared to those receiving the standard protocol. The average number of interviewer contact attempts to finalize treatment cases was substantially lower than for cases offered the standard incentive. Treatment effects on fieldwork duration were large with targeted cases completing the interview seven weeks faster on average than control cases, and more than half of targeted cases completing the interview within four weeks.

Third, our experiment had the key benefit of being embedded in a nationally representative panel study thereby allowing us to test differential treatment effects among high-effort cases across a variety of respondent characteristics. We found that high-income respondents were strongly responsive to the EBI, especially early in fieldwork. This was in contrast to the standard protocol where high-income respondents had dramatically lower rates of early response compared to lower-income respondents. Thus, while being effective for both groups, the EBI treatment had the unexpected and positive effect of removing the differential in response between higher- and lower-income respondents, thereby eliminating an important potential source of bias. As suggested by leverage-saliency theory (Groves, Singer, and Corning 2000), a sizeable incentive may be an especially influential benefit of the survey request for high-income respondents to the extent it is perceived as a better match to the value of their time than the standard incentive. One caveat to note is the specificity of this finding to high-effort high-income cases, which may not generalize to high-income respondents who require lower levels of fieldwork effort.

Fourth, unlike the typical outcome of most adaptive design interventions that tradeoff gains in one domain of survey error for losses in another (Tourangeau et al. 2017), the treatment in the current study was successful in achieving a balance between minimizing nonresponse and survey costs. We found that the overall cost difference between the treatment and the standard protocol was small, and slightly favored treatment cases. While a range of potential survey cost evaluation models exist (see Olson, Wagner, and Anderson 2020), our evaluation was based on both actual and estimated variable costs. Actual incentive costs were obtained from study records; contact attempt costs were derived from actual interviewer contact attempts and an estimate of interviewer productivity using actual interviewer hourly wages and fringe benefits. These costs were used largely because they account for the majority of all fieldwork costs and are relatively straightforward to quantify. Some types of costs that are difficult if not impossible to quantify were not included, such as variable costs due to time spent on case review and production monitoring for sample lines that remain active. However, these activities favor savings for the treatment group since more than half of them were finalized early, reducing interviewer caseloads. Moreover, because of its simplicity, the treatment incurred minimal additional costs other than the extra incentive. For example, no significant special interviewer training or additional supervision was required (other than communicating to interviewers that treatment cases should be treated equally), and there were minimal costs for developing the respondent postcard that was mailed with the advance letter.

General social psychological theories of decision-making can be applied to understand the strong treatment effects during the offer period. The decision-theory framework of regret regulation (Zeelenberg and Pieters 2007) posits that individuals are regret averse and thus motivated to behave in ways that lead them to avoid regret. Studies of consumer psychology have found that anticipatory regret plays a role in purchase intentions which become strengthened by situations that create a sense of urgency, such as time-limited sales and coupon expirations, motivating individuals to act in the immediate or near future (Swain, Hanna, and Abendroth 2006). Our design using a sizeable incentive combined with a relatively short offer period likely motivated respondents to complete the survey request to obtain the higher incentive, and to do so by acting before it was too late, avoiding the anticipated regret of missing out.

We found no evidence of a backlash effect such that respondents became less cooperative following the withdrawal of the extra incentive. Indeed, nearly one-quarter of treatment cases completed an interview for the standard incentive. Some respondents may simply have been too busy to notice the offer, or participate before the expiration, or may have been less intrinsically motivated by the large incentive. This heterogeneity is consistent with the framework of leverage-saliency theory which posits that there is individual variability in the value assigned to various features of a survey request (Groves et al. 2000). It is also possible that the treatment effects persisted even after the EBI expired. A handful of studies have found positive carryover effects of monetary incentives on panel retention over multiple waves (Mack, Huggins, Keathley, and Sundukchi 1998; Scherpenzeel, Zimmermann, Budowski, Tillmann, Wernli, et al. 2002; Jäckle and Lynn 2008; McGonagle 2020) potentially by underscoring the legitimacy of the study and the high importance of participation (Singer and Ye 2013). In the context of a long-running panel study, the positive messages conveyed by the original incentive offer may remain salient to loyal panel members.

A key strength of this study is the random assignment of high-effort respondents and the implementation of the experiment at the start of data collection. However, a limitation is that although interviewers were trained to treat all cases equally, it was not possible to blind them to which cases were eligible for the higher incentive offer. We found that interviewers made significantly more contact attempts on average to treatment cases during the offer period which could potentially explain why these cases were more cooperative. It is also possible that at least some of the additional contact attempts were due to the greater responsiveness of the treatment cases to the intervention. Since it is rare for respondents to complete the interview on the first attempt, cases that have signaled their cooperation may receive subsequent contact attempts to set an interview appointment, administer the interview itself, and for some respondents, take multiple sessions to finish the entire interview. Although beyond the scope of this analysis, future research should address the issue of interviewer reactivity to early incentive designs.

Several additional directions for future research to improve the design of EBI interventions are suggested by the current study. One is identifying the optimal length of the EBI offer and the amount of the incentive. We found that a relatively brief offer period combined large incentive amount that was double the standard incentive was highly effective and widely appealing to high-effort respondents. Additional research should examine how varying the length of the offer period affects the take-up rate. For example, longer offer periods may provide individuals with a greater opportunity to respond, but by reducing time pressure and a sense of urgency, they may be less successful in motivating individuals to complete an interview before it expires. Future research should also pinpoint whether lower incentive amounts are equally successful and hence more cost-effective. In the panel study context, the design of targeted incentive strategies must also consider principles such as the provision of fair and equitable tokens in exchange for the burden of survey participation, avoiding coercive strategies, and building goodwill to ensure the success of future requests (Singer and Ye 2013). While we found that the EBI led targeted respondents to respond at high rates even after the offer expired, an important question is whether sizeable time-delimited EBIs create expectations for future extra incentives that spillover to subsequent waves and if so, how targeted EBIs can be designed to manage carryover effects.

Another design issue is how we communicate to respondents about EBIs and other special treatments in targeted designs. The current design used a message that was framed in a positive manner describing a bonus that would be provided for taking up a special offer. Drawing on prospect theory (Kahneman and Tversky 1979) which predicts that individuals are more sensitive to losses than to gains of the same magnitude, recent experimental research has demonstrated that for some respondents, messages highlighting the negative consequences of nonparticipation may have stronger effects on gaining cooperation with survey requests than messages emphasizing benefits of participation (Tourangeau and Ye 2009; Kreuter, Sakshaug, and Tourangeau 2016; Lynn 2019). Alternative ways of framing respondent messages merit further exploration and may enhance the effectiveness of EBIs.

Our design of an EBI successfully reduced fieldwork effort and improved response rates while balancing costs among respondents targeted because they required extensive field resources in the prior wave. A more nuanced understanding of the reasons some respondents require extensive effort could improve our design and potentially do so with much lower cost. Finally, future research should examine how EBIs may be effectively targeted to other high-priority subgroups such as prior wave nonrespondents whose continued participation is of critical importance for maintaining the scientific value of panel data.

## Supplementary Materials

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

## REFERENCES

Beullens, K., G. Loosveldt, C. Vandenplas, and I. Stoop (2018), "Response Rates in the European Social Survey: Increasing, Decreasing, or a Matter of Fieldwork Efforts?," *Survey Methods: Insights from the Field*, available at https://surveyinsights.org/?p=9673.

Brown, M., and L. Calderwood (2014), "Can encouraging respondents to contact interviewers to make appointments reduce fieldwork effort? Evidence from a randomized experiment in the UK," *Journal of Survey Statistics and Methodology*, 2, 484–497.

Coopersmith, J., L. K. Vogel, T. Bruursema, and K. Feeney (2014), "Effects of Incentive Amounts and Type on Web Survey Response Rates," Survey Practice, 9, 10.29115/SP-2016-0002.

de Leeuw, E., J. Hox, and A. Luiten (2018), "International Nonresponse Trends across Countries and Years: An Analysis of 36 Years of Labour Force Survey Data," *Survey Insights: Methods from the Field*, available at https://surveyinsights.org/?p=10452.

Fomby, P., N. Sastry, and K. A. McGonagle (2017), "Effectiveness of a Time-Limited Incentive on Participation by Hard-to-Reach Respondents in a Panel Study," *Field Methods*, 29, 238–251.

Fumagalli, L., H. Laurie, and P. Lynn (2012), "Experiments with Methods to Reduce Attrition in Longitudinal Surveys," *Journal of the Royal Statistical Society Statistics in Society, Series A*, 176, 499–519.

Groves, R. M., and M. P. Couper (1998), *Nonresponse in Household Interview Surveys*. New York: John Wiley.

Groves, R. M., E. Singer, and A. Corning (2000), "Leverage-Saliency Theory of Survey Participation: Description and an Illustration," *Public Opinion Quarterly*, 64, 299–308.

Jäckle, A., and P. Lynn (2008), "Respondent Incentives in a Multi-Mode Panel Survey: Cumulative Effects on Nonresponse and Bias," *Survey Methodology*, 34, 105–117.

Kahneman, D., and A. Tversky (1979), "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47, 263–291.

Kaplan, E. L., and P. Meier (1958), "Nonparametric Estimation from Incomplete Observations," *Journal of the American Statistical Association*, 53, 457–481.

Kreuter, F., and G. Müller (2015), "A Note on Improving Process Efficiency in Panel Surveys with Paradata," *Field Methods*, 27, 55–65.

Kreuter, F., J. Sakshaug, and R. Tourangeau (2016), "The Framing of the Record Linkage Consent Question," *International Journal of Public Opinion Research*, 28, 142–152.

LeClere, F., S. Plummer, J. Vanicek, A. Amaya, and K. Carris (2012), "Household Early Bird Incentives: Leveraging Family Influence to Improve Household Response Rates," American Statistical Association Joint Statistical Meetings, Proceedings of the Survey Research Section, pp. 4156–4165, available at http://www.asasrms.org/Proceedings/y2012/Files/304811_73772.pdf.

Lugtig, P. (2014), "Panel Attrition: Separating Stayers, Fast Attriters, Gradual Attriters, and Lurkers," *Sociological Methods and Research*, 43, 699–723.

Luiten, A., and B. Schouten (2013), "Tailored Fieldwork Design to Increase Representative Household Survey Response: An Experiment in the Survey of Consumer Satisfaction," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176, 169–189.

Lynn, P. (2014), "Targeted Response Inducement Strategies on Longitudinal Surveys," in *Improving Survey Methods: Lessons from Recent Research*, ed. U. Engel, pp. 322–338, New York: Routledge.

———. (2016), "Targeted Appeals for Participation in Letters to Panel Survey Members," *Public Opinion Quarterly*, 80, 771–782.

———. (2017), "From Standardised to Targeted Survey Procedures for Tackling Nonresponse and Attrition," *Survey Research Methods*, 11, 93–103.

———. (2019), "Applying Prospect Theory to Participation in a CAPI/Web Panel Survey," *Public Opinion Quarterly*, 83, 559–567.

Mack, S., V. Huggins, D. Keathley, and M. Sundukchi (1998), "Do Monetary Incentives Improve Response Rates in the Survey of Income and Program Participation?," Proceedings of the Section on Survey Methodology, American Statistical Association, pp. 529–534.

McGonagle, K. A. (2020), "The Effects of an Incentive Boost on Response Rates, Fieldwork Effort, and Costs across Two Waves of a Panel Study," *Methods, Data, Analyses*, 14, 241–250.

McGonagle, K. A., and N. Sastry (2020), "An Experimental Evaluation of an Online Interview Scheduler: Effects on Fieldwork Outcomes," *Journal of Survey Statistics and Methodology*, 9, 412–428.

McGonagle, K. A., and V. A. Freedman (2017), "The Effects of a Delayed Incentive on Response Rates, Response Mode, Data Quality, and Sample Bias in a Nationally Representative Mixed Mode Study," *Field Methods*, 29, 221–237.

McGonagle, K. A., N. Sastry, and V. A. Freedman (2021), *Early-Bird Incentive Strategy*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2021-07-22. 10.3886/E145761V2.

McGonagle, K. A., R. F. Schoeni, N. Sastry, and V. A. Freedman (2012), "The Panel Study of Income Dynamics: Overview, Recent Innovations, and Potential for Life Course Research," *Longitudinal and Life Course Studies*, 3, 268–284.

National Research Council (2013), "Nonresponse in Social Science Surveys: A Research Agenda," in *Panel on a Research Agenda for the Future of Social Science Data Collection*, eds. Tourangeau R. and T.J. Plewes. Committee on National Statistics, Division of Behavioral and Social Sciences and Education, National Research Council. Washington, DC: The National Academies Press.

Olson, K., J. Wagner, and R. Anderson (2020), "Survey Costs: Where Are We and What Is the Way Forward?," *Journal of Survey Statistics and Methodology*, 0, 1–22. https://doi.org/10.1093/jssam/smaa014.

Peycheva, D., L. Calderwood, and E. Wong (2019), Effects of a Time-Limited Push-To-Web Incentive in a Mixed-Mode Longitudinal Study of Young Adults, available at https://esrc.ukri.org/files/funding/funding-opportunities/darina-peycheva-lisa-calderwood-and-erica-wong-effects-of-a-time-limited-push-to-web-incentive-in-a-mixed-mode-longitudinal-study-of-young-adults/.

Peytchev, A., S. Riley, J. Rosen, J. Murphy, and M. Lindblad (2010). "Reduction of Nonresponse Bias in Surveys through Case Prioritization," *Survey Research Methods*, 8, 67–80.

Scherpenzeel, A., E. Zimmermann, M. Budowski, R. Tillmann, B. Wernli, and A. Gabadinho (2002), Experimental Pre-Test of the Biographical Questionnaire, Neuchatel: Swiss Household Panel. Working Paper, No. 502. Available at: http://aresoas.unil.ch/workingpapers/WP5_02.pdf. Accessed May 31, 2012.

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

Schouten, B., M. Calinescu, and A. Luiten (2013), "Optimizing Quality of Response through Adaptive Survey Design," *Survey Methodology*, 39, 29–58.

Singer, E., and C. Ye (2013), "The Use and Effects of Incentives in Surveys," *The Annals of the American Academy of Political and Social Science*, 645, 112–141.

Swain, S. D., R. Hanna, and L. J. Abendroth (2006), "How Time Restrictions Work: The Roles of Urgency, Anticipated Regret, and Deal Evaluations," in *NA—Advances in Consumer Research* (Vol. 33), eds. C. Pechmann and L. Price, pp. 523–525, Duluth, MN: Association for Consumer Research.

The American Association for Public Opinion Research (AAPOR) (2016), *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys* (9th ed.), Oakbrook Terrace, IL: AAPOR.

Tourangeau, R., M. Brick, S. Lohr, and J. Li (2017), "Adaptive and Responsive Survey Designs: A Review and Assessment," *Journal of the Royal Statistical Society (A)*, 180, 203–223.

Tourangeau, R., and C. Ye (2009), "The Framing of the Survey Request and Panel Attrition," *Public Opinion Quarterly*, 73, 338–348.

Wagner, J., B. T. West, N. Kirgis, J. M. Lepkowski, W. G. Axinn, and S. Kruger-Ndiaye (2012), "Use of Paradata in a Responsive Design Framework to Manage a Field Data Collection," *Journal of Official Statistics*, 28, 477–499.

Williams, D., and J. M. Brick (2018), "Trends in U.S. Face-to-Face Household Survey Nonresponse and Level of Effort," *Journal of Survey Statistics and Methodology*, 6, 186–211.

Zeelenberg, M., and R. Pieters (2007), "A Theory of Regret Regulation 1.0," *Journal of Consumer Psychology*, 17, 3–18.