



Visible cash, a second incentive, and Priority Mail? An experimental evaluation of mailing strategies for a screening questionnaire in a national push-to-web/mail survey

Journal:	<i>Journal of Survey Statistics and Methodology</i>
Manuscript ID	JSSAM-2022-089.R1
Manuscript Type:	Article
Keywords:	push-to-web/mail survey, incentive, Priority Mail, response rate, respondent composition, mailing strategy

SCHOLARONE™
Manuscripts

1
2
3
4 Visible cash, a second incentive, and Priority Mail? An experimental evaluation of mailing
5
6
7 strategies for a screening questionnaire in a national push-to-web/mail survey
8
9
10
11
12

13 **Abstract**
14

15
16
17 In push-to-web/mail surveys that use mailing to contact sampled cases, participation is
18
19
20 contingent on the mail being opened and the survey invitation being delivered. The design of the
21
22
23 mailings is crucial to the success of the survey. We address the question of how to design
24
25
26 invitation mailings that can grab potential respondents' attention and sway them to be interested
27
28
29 in the survey in a short window of time. In the household screening stage of a national survey,
30
31
32 the American Family Health Study, we experimentally tested three mailing design techniques for
33
34
35 recruiting respondents: 1) a visible cash incentive in the initial mailing, 2) a second incentive for
36
37
38 initial nonrespondents, and 3) use of Priority Mail in the nonresponse follow-up mailing. We
39
40
41
42
43
44 evaluated the three techniques' overall effects on response rates, as well as how they
45
46
47
48
49
50
51 differentially attracted respondents with different characteristics. We found that all three
52
53
54
55
56
57
58
59
60 techniques were useful in increasing the screening response rates, but there was little evidence

1
2
3 that they had differential effects on sample subgroups that could help to reduce nonresponse
4
5
6
7 biases.

13 **Statement of significance**

14
15
16
17 Many scientists are increasingly attracted to using push-to-web/mail surveys to collect social
18
19
20 data. Compared to alternative modes (e.g., the face-to-face interview mode), web surveys have
21
22
23 lower costs, require less infrastructure, and implement questionnaires in a more private way.
24
25
26
27 However, two major concerns about web surveys are the low response rates and respondent
28
29
30 representativeness, both of which are likely related to the brief and impersonal recruitment
31
32
33 interaction with potential respondents through mailing. In the household screening stage of a
34
35
36 national push-to-web/mail survey, the American Family Health Study, we experimentally
37
38
39 evaluated how response rates and respondent compositions were affected by three mailing
40
41
42 techniques: 1) a visible cash incentive in the initial mailing, 2) a second incentive for initial
43
44
45 nonrespondents, and 3) the use of Priority Mail in the nonresponse follow-up mailing. Our
46
47
48
49
50 results contribute to the development of cost-efficient mailing strategies for large-scale web/mail
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

surveys and to understanding the influences of recruitment strategies on sample
representativeness and nonresponse biases.

For Review Only

Introduction

Push-to-web/mail surveys with address-based probability samples have become increasingly attractive for many reasons. Today, a large portion of the U.S. population can participate in surveys via the web as internet coverage has reached over 90% among adults living in the U.S. (Pew Research Center, 2021). The ability to answer questions anytime, anywhere, and without the presence of interviewers helps to facilitate the confidentiality and privacy that are crucial to surveys asking sensitive and personal questions (Heerwegh, 2009; West, Axinn, Couper, Gatny, & Schroeder, 2022). Also, compared to interviewer-administered surveys, conducting self-administered push-to-web/mail surveys typically costs less and requires less infrastructure (Battaglia et al., 2016; Biemer, Murphy, Zimmer, Berry, Deng, & Lewis, 2018; Couper, 2008). For example, compared to hiring and training interviewers all across the country for face-to-face interviews, it is much easier to purchase an address-based sample and use postal mailings to invite cases to complete a web survey (Murphy, Biemer, & Berry, 2018). As an additional advantage, the address-based samples used for push-to-web/mail surveys typically produce lower design effects than the cluster samples employed by face-to-face surveys (West et al., 2022).

1
2
3
4 Despite these strengths of push-to-web/mail surveys, two major concerns about survey
5
6
7 quality are response rates and respondent representativeness. Face-to-face household surveys can
8
9
10 reach 60%-70% response rates (Williams & Brick, 2018), whereas many push-to-web/mail
11
12
13 surveys yield response rates below 30% (Braekman et al., 2022; Daikeler et al., 2020; Link &
14
15
16 Burks, 2013; Mercer, Caporaso, Cantor, & Townsend, 2015). Lower response rates risk
17
18
19 introducing nonresponse biases into survey estimates (Brick & Tourangeau, 2017). In terms of
20
21
22 respondent representativeness, previous studies have shown that self-administered push-to-
23
24
25 web/mail surveys recruit relatively fewer male, low-educated, and racial and ethnic minority
26
27
28 respondents [than suggested by population benchmarks](#) (Baker et al., 2010; Jackson, McPhee,
29
30
31 Lavrakas, 2020; Luijkx et al., 2021; Messer & Dillman, 2011). To a certain extent, the problems
32
33
34 of low response rates and unbalanced respondent compositions are interrelated. [Surveys with low](#)
35
36
37 [response rates get easier-to-recruit respondents, and these people tend to have different](#)
38
39
40 [characteristics than those who are harder to survey](#) (Brick & Tourangeau, 2017).
41
42
43
44
45
46

47 One potential reason for the low response rates in push-to-web/mail surveys is the brief
48
49
50 and impersonal recruitment interaction with potential respondents. [According to social and](#)
51
52
53 [cognitive psychology theories of decision-making, survey participation is likely to be decided](#)
54
55
56
57
58
59
60

1
2
3 through a “peripheral” route using heuristics and minimal cognitive effort, because people do not
4
5
6
7 take great interest in survey participation (Groves, Cialdini, & Couper, 1992; Keusch, 2015).
8
9

10 This heuristic-based processing can make it particularly challenging to recruit respondents by
11
12
13 mail. When people sort through their mail (a large proportion of which is third class advertising),
14
15
16
17 the decisions about whether to open and review each mailing are made quickly (Lesscher,
18
19
20 Lobschat, & Verhoef, 2021; Scardaville, Dias, Paek, Pugliese, Medway, & Zukerberg, 2021). If
21
22
23 the mail does not appear interesting or important from the look of the envelope, it can be
24
25
26
27 discarded without ever being opened and the invitation would never be delivered (Feld, Frenzen,
28
29
30 Krafft, Peters, & Verhoef, 2013; Medway, Paek, Battle, & McPhee, 2021). Even if the mail is
31
32
33
34 opened, potential respondents may not review the request in detail if the content does not seem
35
36
37 attractive.
38
39

40 Because push-to-web/mail survey participation is contingent on mail being opened and
41
42
43 an invitation being read, the design of the mailings is crucial to the success of these surveys. We
44
45
46
47 were motivated by the challenge of how to design invitation mailings that can grab potential
48
49
50 respondents’ attention and convince them to participate. To this end, we experimentally tested
51
52
53
54 three design techniques that might affect impressions of the mailings and trigger interest in the
55
56
57
58
59
60

1
2
3 content from those receiving and opening the mail. Two techniques changed the appearance of
4
5
6 the mailings by making a cash incentive visible in the initial mailing and using Priority Mail in
7
8
9 the nonresponse follow-up mailing (instead of regular First Class mail); the third technique
10
11
12 added a second incentive in the nonresponse follow-up mailing. After implementing our
13
14
15
16
17 experiment in a large national probability sample of addresses, we documented how these design
18
19
20 techniques influenced the response rate to the screening invitation and how they differentially
21
22
23 attracted respondents with different features. Our experiments were conducted as part of an effort
24
25
26 to invite household members to complete a screening questionnaire for a national survey; the
27
28
29 results are generally relevant to both the screening stage of two-phase surveys and single-phase
30
31
32 surveys.
33
34
35
36
37
38
39

40 **Background**

41 *Visible Cash in Initial Mail Invitations*

42
43
44 One concern about mailing out prepaid incentives on a large scale is that people are not aware of
45
46
47 the cash and the envelopes are discarded without ever being opened, which wastes funds and
48
49
50
51 curbs the effects of prepaid incentives in motivating survey participation (DeBell, Maisel,
52
53
54
55
56
57
58
59
60

1
2
3
4 Edwards, Amsbary, Meldener, 2020). To highlight the presence of the incentives, survey
5
6
7 practitioners have previously experimented with putting messages on the envelopes to indicate
8
9
10 the existence of the cash gift, but had limited success (Dykema et al., 2015). To go one step
11
12
13 further, a recently-developed technique directly makes the cash visible via a transparent window
14
15
16 on the envelope (Bilgen, Dutwin, Singh, & Hendarwan, 2021; DeBell et al., 2020; Sherr &
17
18
19 Wells, 2021). In theory, making a cash incentive one of the first noticeable things would increase
20
21
22 the salience of the incentive, which in turn facilitates mail opening and cooperation by intriguing
23
24
25 potential respondents with the mailing and setting them in a positive frame of mind (DeBell et
26
27
28 al., 2020).
29
30
31

32
33
34 A few studies have indeed shown that the visibility of cash boosted the positive effect of
35
36
37 prepaid incentives. The American National Election Studies (ANES) experimented with making
38
39
40 a \$5 prepaid incentive visible in a *nonresponse follow-up study* in 2017. The visible incentive
41
42
43 significantly increased the response rate from 42.6% to 46.9% (DeBell et al., 2020). Likewise,
44
45
46 the California Health Interview Survey (CHIS) tested a \$2 visible incentive in the screening
47
48
49 stage of a push-to-web/telephone survey. The visible cash condition had a significantly higher
50
51
52 cooperation rate than the non-visible condition; this effect was carried over to the main interview
53
54
55
56
57
58
59
60

1
2
3 stage, resulting in a significantly higher number of main responses (Sherr & Wells, 2021; S.
4
5
6
7 Sherr, personal communication, May 2, 2022). NORC at the University of Chicago's
8
9
10 AmeriSpeak panel also experimented with the prepaid cash incentive amount and the
11
12
13 presentation of the cash as part of initial recruitment efforts for the panel. Making the incentive
14
15
16
17 visible increased the odds of converting a household member to a panelist by 1.2 times relative
18
19
20 to using a regular envelope. The positive effect of visible cash was strengthened by sending a \$5
21
22
23 incentive as opposed to \$2 (1.46 times more effective) and showing the amount of the dollar bill
24
25
26 through the window as opposed to showing the image/face (1.39 times more effective; Bilgen et
27
28
29 al., 2021).
30
31
32

33
34 Although the overall effect of a visible incentive on response rates seems consistently
35
36
37 positive, less is known about its differential impact on response rates in different sample
38
39
40 subgroups and the types of respondents that it effectively recruits. One desirable possibility is
41
42
43 that the visible cash not only increases the response rate, but also contributes to sample balance
44
45
46
47 by motivating participation from low-response-propensity cases to a larger extent. Because
48
49
50 incentives may compensate for the lack of interest and bring in people who otherwise may not
51
52
53 respond (Singer & Ye, 2013), the idea is that making incentives visible may further increase the
54
55
56
57
58
59
60

1
2
3
4 salience of the request for those whom other features of the request may not provide sufficient
5
6
7 leverage (Groves, Singer, & Corning, 2000). There has been some initial evidence of these
8
9
10 additional positive features of the visible incentive. Based on the AmeriSpeak experiment, the
11
12
13 visible incentive recruited more panelists who were Hispanic, younger (18-24), older (50+), and
14
15
16 with-children (Bilgen et al., 2021). Also, based on the experiment in the ANES nonresponse
17
18
19 follow-up study (which invited both respondents and non-respondents to the original 2016 study;
20
21
22 DeBell & Maisel, 2018), the visible incentive seemed to be particularly useful among harder-to-
23
24
25 recruit individuals. It was shown to increase the response rate among the addresses that were
26
27
28 invited but never responded to the original study to a larger extent (27.2% to 32.5%) than among
29
30
31 the addresses that had previously responded to the survey request (58% to 60.8%; DeBell et al.,
32
33
34 2020). In the current study, we systematically evaluated the effect of a visible incentive on
35
36
37 sample subgroups with different response propensities and on the demographic composition of
38
39
40 the respondents.
41
42
43
44
45
46
47
48
49
50

51 *A Second Prepaid Incentive and Use of Priority Mailing in Nonresponse Follow-up*

52
53
54
55
56
57
58
59
60

1
2
3
4 If sampled cases do not respond to early contact attempts, nonresponse follow-up mailings can
5
6
7 incorporate other elements to further motivate potential respondents and highlight the importance
8
9
10 of the survey request. One possible technique is to use a second prepaid incentive. The positive
11
12
13 effect of a single small prepaid incentive on survey response has been well established by a large
14
15
16 number of studies (Mercer et al., 2015; Singer & Ye, 2013). The prepaid token urges people to
17
18
19 complete the survey by putting the invitation in a positive light and acting on a norm of
20
21
22 reciprocity (Dillman, Smyth, & Christian, 2014). This logic can be extended to justify the use of
23
24
25 a second incentive in a nonresponse follow-up contact, because this may invoke another round of
26
27
28 reciprocity and/or catch some cases who did not open the initial mailing and missed the first
29
30
31 incentive (Dykema et al., 2015). However, how well a second incentive works in practice has not
32
33
34 been widely tested, because offering a second incentive is controversial for rewarding
35
36
37 nonresponse behaviors and is not typically implemented in surveys (Dillman et al., 2014).
38
39
40
41
42
43

44 Recently, a small number of studies have experimented with a second incentive ranging from \$1
45
46
47 to \$10 (Bucks, Couper, & Fulford, 2020; Dirks, Lavrakas, Lusskin, Ponce, & Felstead, 2017;
48
49
50 Dykema et al., 2015; Messer & Dillman, 2011; Wagner et al., 2022). In general, the second
51
52
53 incentive was found to be effective in motivating nonrespondents (Bucks et al., 2020; Dirks et
54
55
56
57
58
59
60

1
2
3
4 al., 2017; Dykema et al., 2015; Messer & Dillman, 2011), but there was some evidence that the
5
6
7 amount needed to be larger than the initial incentive (Dykema et al., 2015).
8
9

10 Repeatedly sending the same invitation may result in the same nonresponse, and thus the
11
12
13 follow-up mailing is more likely to reach and obtain responses from initial nonrespondents if it
14
15
16 includes a noticeable change from the previously unsuccessful mailings (Dillman, 2020; Messer
17
18
19 & Dillman, 2011). Therefore, another technique involves sending the follow-up invitation via a
20
21
22
23 Priority Mail envelope as opposed to a First Class envelope. Priority Mail uses a paperboard
24
25
26 envelope and delivers in 1-3 business days (**Figure 1** left), whereas First Class mail uses a paper
27
28
29 envelope and delivers in 1-5 business days (**Figure 1** right). The special packaging of Priority
30
31
32
33 Mail conveys the importance and legitimacy of the survey request; it also differentiates the final
34
35
36 contact from the previous mailings, which hopefully breaks the cycle of nonresponse (Dillman et
37
38
39 al., 2014; Messer & Dillman, 2011). However, the benefit of using Priority Mail for nonresponse
40
41
42
43 follow-up has not been consistently demonstrated in the literature. In the nonresponse follow-up
44
45
46 study conducted by the ANES (DeBell et al., 2020) and a local survey in Washington (Messer &
47
48
49 Dillman, 2011), Priority Mail did not result in a higher response rate than First Class mail. In
50
51
52
53 contrast, in the 2011 field test of the National Household Education Survey (Han, Montaquila, &
54
55
56
57
58
59
60

1
2
3 Brick, 2013) and the first replicate of the American Family Health Study (Wagner et al., 2022),
4
5
6
7 there was evidence that Priority Mail might have increased response rates compared to First
8
9
10 Class mail; but in both cases, the effect of Priority Mail could not be isolated because it was
11
12
13 confounded by the use of additional incentives.
14
15
16

17 We sought to find an effective and cost-efficient way to combine the use of the second
18
19
20 incentive and Priority Mail in nonresponse follow-up mailings. Both techniques are costly to
21
22
23 implement on a large sample with a relatively low response rate. Our questions were as follows:
24
25
26 Which technique was more effective, if at all, in motivating nonrespondents? When implemented
27
28
29 simultaneously, could their individual effects be added to improve response rates? Or was it
30
31
32 redundant to implement both techniques because only one was sufficient in making the follow-
33
34
35 up contact salient and thus incorporating the other only resulted in diminishing marginal returns?
36
37
38 The experiment reported here aimed to tease apart the individual effects and combined effect of
39
40
41 the second incentive and Priority Mail.
42
43
44
45
46

47 In addition, we examined not only how the second incentive and Priority Mail change
48
49
50 response rates, but also how they influence respondent composition. Particularly, we were
51
52
53 interested in whether the additional design elements might have differentially motivated
54
55
56
57
58
59
60

1
2
3 participation from subgroups that had lower response propensity or were under-represented (e.g.,
4
5
6
7 ethnic minority groups). The reasoning is that the second incentive may urge people who are less
8
9
10 intrinsically interested in the survey to respond (Dillman et al., 2014), and the Priority Mail
11
12
13 provides a change of stimulus that may appeal to people who have not been induced to open or
14
15
16 read earlier mailings (Messer & Dillman, 2011). Thus, in both cases, the additional elements may
17
18
19 contribute to bringing in people with previously under-represented characteristics.
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39



40 **Figure 1.** Left: Priority Mail envelope (sent during the nonresponse follow-up mailing). Right:

41
42
43 First Class envelope with the visible cash incentive (sent during the initial mailing).

44
45
46 Note. The transparent window was placed on the front of the envelope so it would have the
47
48
49 greatest immediate visibility alongside the address. The size of the transparent window was
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

constrained by the need to accommodate other elements, including the return address, mailing address, study logo, and stamp/metered postage.

For Review Only

Methods

The experiments were implemented in the second national sample replicate of the American Family Health Study (AFHS; see afhs.isr.umich.edu for additional details), fielded between September 2021 and April 2022. The AFHS is a two-phase survey, involving a screening stage to identify households with age-eligible persons (age 18-49), and a main stage containing questions on family life, marriage and divorce, reproductive health and health care, and related topics. The questionnaire content of the AFHS closely resembles that of a face-to-face national survey: the National Survey of Family Growth (Centers for Disease Control and Prevention, 2020). The current analysis focuses on outcomes from the screening phase of the survey. The data collection was based on a national address-based probability sample of more than 22,000 addresses. The sample was stratified; addresses predicted to have a higher likelihood of having age-eligible persons present and located in high-density African-American or Hispanic areas were oversampled.

The AFHS screening phase used a sequential mixed-mode push-to-web/mail protocol. Sampled households received up to four mailings in both English and Spanish, as summarized in

Figure 2. The first mailing announced the study and invited an adult in the household to

1
2
3 complete a short survey online (English/Spanish), in which respondents were asked to list all
4
5
6
7 household residents aged 18 or older and provide everyone's name, gender, date of birth,
8
9
10 Hispanic ethnicity, and race. The questionnaire also collected an email address and telephone
11
12
13 number of the screening respondent and asked for permission to text if a mobile number was
14
15
16 provided. [The screening questionnaire is included in Appendix I.](#) The visible cash experiment
17
18
19 was implemented in this first [mailing](#). For a random half of the sample ($n = 11,193$), a \$2 bill
20
21
22 was made visible through a transparent window on the front of the envelope (**Figure 1**, right);
23
24
25 and for the other random half-sample ($n = 11,188$), the \$2 was not visible from the envelope.
26
27
28
29

30 The second mailing was a reminder postcard. The third mailing included both a web link
31
32
33 and a paper version of the screening questionnaire. The paper questionnaire was in English and
34
35
36 was printed as an eight-page booklet; a Spanish version was available upon request by telephone.
37
38
39

40 The fourth mailing was the final nonresponse follow-up contact attempt, again including
41
42
43 both the web link and the paper questionnaire. The experiment examining the effects of the
44
45
46 second incentive and Priority Mail was implemented in this fourth [mailing](#). At this time, roughly
47
48
49 7% of the sample had already responded, and the remaining nonrespondents were randomized
50
51
52
53 into four groups by a full-factorial experimental design (i.e., fully crossed with the earlier visible
54
55
56
57
58
59
60

1
2
3
4 cash experimental assignment): 10% (n = 2,081) of the nonrespondents were sent the final
5
6
7 mailing as First Class mail with no additional incentive; 10% (n = 2,110) were sent Priority Mail
8
9
10 (Figure 1, left) with no additional incentive; 10% (n = 2,087) were sent First Class mail with a \$5
11
12
13 cash incentive included (not visible); and 70% (n = 14,570) were sent Priority Mail with a \$5
14
15
16 incentive included (not visible). Most cases were assigned to the Priority Mail and \$5 condition
17
18
19 for the purpose of maximizing the response rate, as the first national sample replicate of AFHS
20
21
22 found that sending a nonresponse follow-up Priority Mail with \$5 led to a large increase in the
23
24
25 screening questionnaire response rate compared to no additional mailing (13.8% vs. 5.3%;
26
27
28
29
30 Wagner et al., 2022).

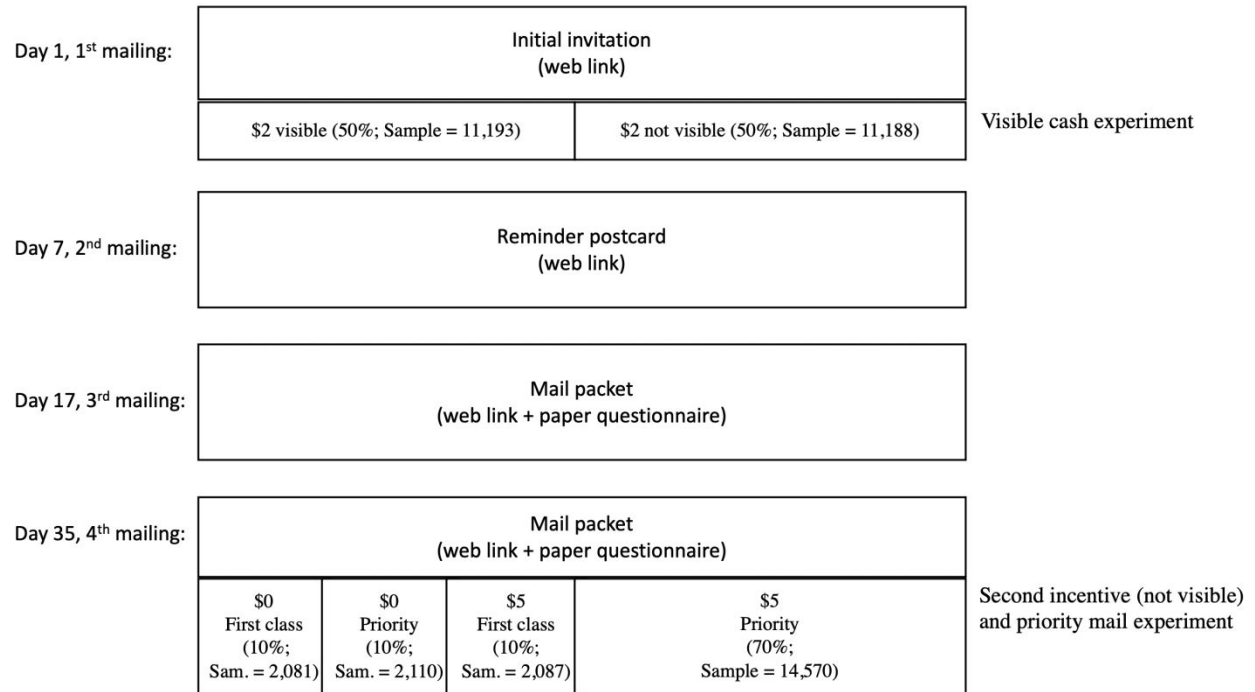


Figure 2. Four mailings and the assignments of the two experiments.

Analytical Procedure

We evaluated the results of the two experiments in four ways. First, we computed how the mailing design techniques changed response rates to the screening questionnaire.

Second, we compared the screening respondents in the different experimental conditions in terms of gender, age, race and ethnicity, household size, and household eligibility. The purpose of the comparison was to understand how people of different characteristics were more

1
2
3 or less influenced by the different mailing techniques. We note that the demographic
4
5
6 composition of these screening respondents was different from that of final respondents because
7
8
9
10 AFHS is a two-stage survey, and thus not all screening respondents were eligible or selected for
11
12
13 the main interview.
14
15

16
17 Third, we evaluated the effects of the mailing techniques for different sample subgroups.
18
19 One type of differential effect that we were particularly interested in was whether the techniques
20
21 motivated low-response-propensity cases to a larger extent. To this end, we divided the sample
22
23 into three subgroups with different response propensities using their block group-level feature—
24
25 the *Low Response Score* (LRS) developed by the Census Bureau. The LRS summarizes
26
27 sociodemographic characteristics associated with low Census mail return rates (e.g., high
28
29 proportion of renter-occupied units, female-headed households, people aged 18-24, and people
30
31 below poverty; Erdman & Bates, 2017). We took tertiles of LRS to correspond to high-,
32
33 medium-, and low-response-propensity subgroups.
34
35
36
37
38
39
40
41
42
43
44
45

46
47 Another type of effect that we were interested in was whether the mailing techniques
48
49 worked differently for subgroups defined by variables related to mailing operations. From
50
51 sampling frame variables, we identified a few of such variables that potentially have implications
52
53
54
55
56
57
58
59
60

1
2
3
4 for recruitment planning. *Probability of being an undeliverable address* captured the probability
5
6
7 of the sampled addresses being undeliverable. This variable was not directly available on the
8
9
10 sampling frame, but was predicted based on a model fitted using predictors such as Census
11
12
13 Division, percentage of vacant units at the block group level, whether the address was matched
14
15
16 to a commercial data source Marketing Systems Group (MSG), and whether the address was
17
18
19 marked as seasonal by MSG (Nishimura, Gatward, West, & Saw, 2021). Using the probabilities,
20
21
22 we divided the sample into two groups: one containing cases with the highest 10% of predicted
23
24
25 probabilities and the other containing the remaining cases. We focused on the predicted
26
27
28 probabilities instead of the actual undeliverable outcomes, because the probabilities can be
29
30
31 estimated before the data collection and thus are potentially useful for informing the design of
32
33
34 recruitment procedures. *Type of address* indicated whether the sampled address was a single-
35
36
37 family, multi-family, or PO Box. The information came from commercial data purchased from
38
39
40 MSG. *Census Region* indicated whether the address was in the West, Midwest, South, or
41
42
43
44
45
46
47 Northeast.

1
2
3
4 Finally, we were interested in the cost-efficiency of using the visible incentive, the
5
6
7 second incentive, and Priority Mail. We calculated the cost of each mailing technique per
8
9
10 respondent to evaluate whether the increase in response was worth the increase in costs.
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Review Only

Results

Visible Cash Experiment in Initial Mailing

Table 1 reports the screening and main response rates observed for the visible incentive experiment. Compared to the non-visible incentive, making the \$2 prepaid incentive visible through the envelope significantly increased the screening response rate from 15.5% to 16.8%.

To get an idea of how the visible incentive might have affected the timing of the responses, we plot the cumulative screening response rates of web respondents by the experimental condition in Figure 3. We focused only on the web respondents because we do not know when the paper questionnaires were completed and posted. The line graph shows that the difference between visible and non-visible conditions peaked around the third week and persisted until the end of the data collection.

We further examined whether the increases in screening response translated to more completed main interviews. Contingent on screening responses, the main-stage conditional response rates for the two experimental conditions were almost identical (62.6% vs. 62.3%). Thus, the visible incentive resulted in a higher yield of main interviews by bringing in more respondents during the screening stage.

Table 1. Screening and main response rates (RR) for the visible cash experiment

	Not visible	Visible	
Sample size	11,188	11,193	
Screening respondent	1,733	1,884	
Screening RR (AAPOR Response Rate 1)	15.5%	16.8%	$\chi^2(1) = 7.340, p = 0.007$
Invited for main	1,011	1,148	
Main respondent	633	715	
Main RR (conditional)	62.6%	62.3%	$\chi^2(1) = 0.013, p = 0.910$

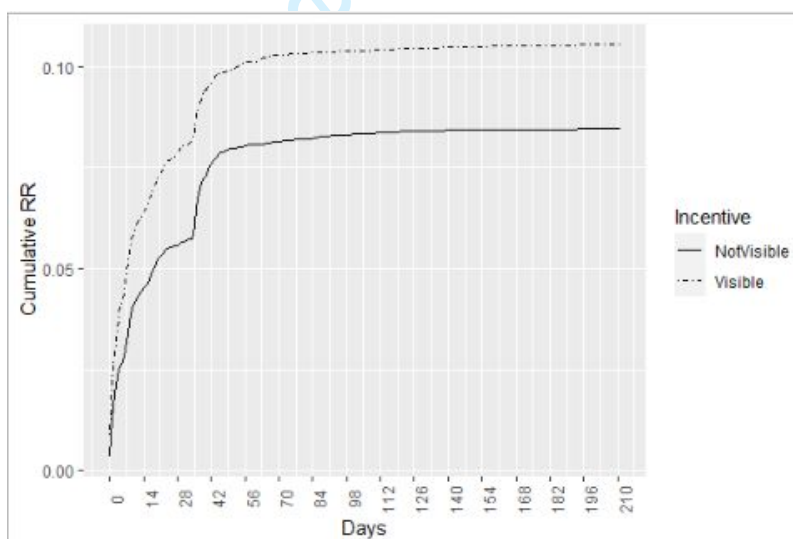


Figure 3. Cumulative response rates of *web* respondents by visible incentive treatment

1
2
3
4 **Table 2** reports the characteristics of respondents by experimental condition. Compared
5
6
7 to respondents whose incentive was not visible, respondents in the visible incentive condition
8
9
10 tended to be younger (<50). There is also evidence that the visible incentive recruited more
11
12
13 Hispanic respondents (14.4% vs. 12.0%) and fewer non-Hispanic White respondents (66.5% vs.
14
15
16 62.7%). Other than these differences, making the incentive visible did not influence distributions
17
18
19 in terms of gender, household size, and household eligibility.
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 2. Demographic characteristics of respondents by visible cash experimental condition

	Not visible		Visible		Diff (p.p.)	Diff p-value ^a	Chi-squared test of independence
	n	%	n	%			
Gender							
Female	1,062	62.1%	1,119	60.4%	-1.8	0.290	$\chi^2(1) = 1.1,$ p = 0.290
Male	647	37.9%	735	39.6%	-	-	
Age							
<25	85	5.0%	119	6.4%	1.4	0.082	$\chi^2(3) = 7.5,$ p = 0.058
25-34	288	16.9%	312	16.8%	-0.1	0.955	
35-49	435	25.5%	520	27.9%	2.4	0.110	
50+	897	52.6%	910	48.9%	-3.7	0.029*	
Race/ethnicity							
Hispanic	205	12.0%	268	14.4%	2.4	0.040*	$\chi^2(3) = 7.0,$ p = 0.071
non-H Black	230	13.5%	257	13.8%	0.3	0.807	
non-H Other	138	8.1%	170	9.1%	1.1	0.290	
non-H White	1,136	66.5%	1,168	62.7%	-3.8	0.020*	
# of adults							
1	461	26.9%	509	27.4%	0.5	0.765	$\chi^2(3) = 0.7,$ p = 0.875
2	864	50.3%	942	50.6%	0.3	0.886	
3	225	13.1%	227	12.2%	-0.9	0.443	
4+	166	9.7%	182	9.8%	0.1	0.956	
Household Eligibility							
Not eligible	722	41.7%	736	39.1%	-2.6	0.120	$\chi^2(1) = 2.4,$ p = 0.120
Eligible	1,011	58.3%	1,148	60.9%	-	-	

^a The p-values are based on Chi-squared tests that compare whether the proportions of the demographic categories in the visible and non-visible conditions (e.g., 62.1% vs. 60.4%) are significantly different.

Table 3 reports response rates by sample subgroups defined by response propensity, probability of having an undeliverable address, type of address, and Census region. The right-

1
2
3
4 most column reports tests of joint independence, namely, whether the joint distribution between
5
6
7 the experimental conditions (visible vs. not visible) and survey outcomes (response vs.
8
9
10 nonresponse) is independent of the sample subgroups (e.g., tertiles of LRS). All the joint
11
12
13 independence tests are statistically significant, suggesting that the effect of the visible incentive
14
15
16 differed by sample subgroups. To break down the effect of the visible incentive by sample
17
18
19 subgroup, we report the difference in response rates between receiving visible and not visible
20
21
22 incentives in the “Diff RR” column and the corresponding p-value in the “Diff p-value” column.
23
24
25

26
27 Across sample subgroups that had different response propensities, the effect was the
28
29
30 largest among cases that have moderate response propensity (difference in response rates = 1.8
31
32
33 percentage points in the 2nd tertile of LRS) and the weakest among cases that have low response
34
35
36 propensity (difference in response rates = 1.1 p.p. in the 3rd tertile of LRS). These results suggest
37
38
39 that making the incentive visible might be the most effective among people who were on the
40
41
42 fence about responding, but its effect might be weaker on the two ends of the spectrum among
43
44
45 those who were already inclined to respond and those who were hard to recruit.
46
47
48

49
50 For the other sample subgroups defined by operational variables, the effect of the visible
51
52
53 incentive was weaker among the 10% of addresses that were predicted to be the most likely
54
55
56

1
2
3 undeliverable (0.8 p.p.), compared to the rest of the sample (1.4 p.p.). The results suggest that we
4
5
6
7 correctly identified addresses that were likely undeliverable, which lowered the response rate and
8
9
10 diminished the effect of the visible incentive. In fact, we confirmed that the *actual* undeliverable
11
12
13 rate was higher among the addresses *predicted* to be the most likely undeliverable—28% of them
14
15
16 had at least one piece of mail undelivered, compared to 6% in the remaining sample. We
17
18
19 alternatively evaluated the effect of the visible incentive among addresses that had different
20
21
22 delivery outcomes. The visible incentive had no effect on the 1818 addresses that had at least one
23
24
25 mailing undelivered (3.1% response rate for non-visible incentive vs. 3.1% for visible incentive),
26
27
28 in contrast to the significant positive effect on the remaining addresses where all mail was
29
30
31 delivered (16.6% response rate for non-visible incentive vs. 18.0% for visible incentive).
32
33
34
35
36

37 In addition, the visible incentive seemed to be less effective for PO box addresses (-2.6
38
39 p.p.), though the sample size of PO boxes was too small to draw any definite conclusions. Across
40
41
42 the geographic regions, we observed that the visible incentive was the most effective in the West
43
44
45 (3.5 p.p.) and less effective in the Midwest (0.3 p.p.) and the South (0.4 p.p.). Reasons for these
46
47
48 differences across the regions were unclear.
49
50
51
52
53

54 **Table 3.** Response rates (RR) by visible incentive experimental condition and sample subgroup

55	56	57	58	59	60
	Not visible	Visible			

	Sample	RR	Sample	RR	Diff RR (p.p.)	Diff p-value ^a	Joint independence
<i>Tertile of LRS</i>							
low LRS (high propensity)	3,681	18.0%	3,747	19.4%	1.4	0.140	
medium LRS (moderate propensity)	3,794	15.8%	3,634	17.6%	1.8	0.041*	$\chi^2(6) = 91.8,$ $p < 0.0001$
high LRS (low propensity)	3,653	12.6%	3,775	13.7%	1.1	0.171	
<i>Probability of undeliverable</i>							
highest 10%	1,089	12.2%	1,149	13.1%	0.8	0.592	$\chi^2(3) = 24.8,$
remaining 90%	10,099	15.8%	10,044	17.3%	1.4	0.007**	$p < 0.0001$
<i>Type of address</i>							
Single family	7,906	16.2%	7,790	17.6%	1.4	0.005**	$\chi^2(6) = 24.3,$
Multi-family	3,203	13.7%	3,331	15.2%	1.4	0.110	$p = 0.0004$
PO Box	79	16.5%	72	13.9%	-2.6	0.832	
<i>Census region</i>							
West	2,544	13.8%	2,543	17.2%	3.5	<0.001***	
Midwest	2,057	19.0%	2,058	19.2%	0.3	0.849	$\chi^2(9) = 40.7,$
South	4,915	15.1%	4,917	15.5%	0.4	0.620	$p < 0.0001$
Northeast	1,672	14.9%	1,675	17.1%	2.2	0.085	

^a The p-values are based on Chi-squared tests that compare whether the response rates for the visible and non-visible conditions (e.g., 18.0% vs. 19.4%) are significantly different.

Second Incentive and Priority Mail Experiment in Nonresponse Follow-up Mailings

Table 4 reports the results of the second incentive and Priority Mail experiment. The upper panel presents the screening and main response rates for the four experimental conditions. Both the Priority Mail and the \$5 second incentive were useful in boosting the screening response rate, but

1
2
3 the effect of the second incentive was significantly stronger than that of Priority Mail (i.e., 8.6%
4
5
6
7 is significantly higher than 6.7%, $p = 0.046$). Combining the two techniques yielded the highest
8
9
10 response (11.5%), compared to using either one of them alone. The difference in screening
11
12
13 response rates between implementing neither of the techniques and implementing both mailing
14
15
16 techniques is consistent with the findings in the first replicate of AFHS, where the response rate
17
18
19 among nonrespondents who were sent a reminder in a Priority Mail envelope that also included
20
21
22 an additional \$5 was 13.8%, compared to 5.3% with no additional mailing (Wagner et al., 2022).
23
24
25
26 Furthermore, when conditioning on responses to the screening questionnaire, the main-stage
27
28
29 response rates were not significantly different across the four experimental groups. Thus, just
30
31
32 like the use of a visible incentive, sending the second incentive and Priority Mail also resulted in
33
34
35 a higher yield of main survey responses by bringing in more cases during the screening stage.
36
37

38 The lower panel of **Table 4** presents the results of a logistic regression model, fitted as an
39
40
41 alternative way to evaluate the experiment. We specified the second incentive and Priority Mail
42
43
44 as binary predictors and survey response as the outcome variable. The combined effect of the
45
46
47 two techniques was tested as an interaction term in the model. The significant main effects and
48
49
50 the nonsignificant interaction term suggest that the effects of the second incentive and Priority
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Mail were distinct and additive, and the two techniques did not reinforce or undermine the effect of one another.

For Review Only

Table 4. Survey outcomes of the second incentive and Priority Mail experiment

Response rate (RR) by experimental condition					
	\$0 First class	\$0 Priority	\$5 First class	\$5 Priority	
Sample size	2,081	2,110	2,087	14,570	
Screening respondent	98	141	179	1,682	
Screening RR (AAPOR Response Rate 1)	4.7%	6.7%	8.6%	11.5%	$\chi^2(3) = 132.850,$ $p < 0.0001$
Invited for Main	52	71	91	917	
Main respondent	26	31	44	489	
Main RR	50%	43.7%	48.4%	53.3%	$\chi^2(3) = 3.181,$ $p = 0.365$
Logistic regression results (response ~ experimental treatment)¹					
	Coefficient	Odds ratio	SE	p-value	
Intercept	-3.007	0.049	0.103	<0.001	***
Priority (vs. First class)	0.371	1.449	0.135	0.006	**
\$5 (vs. \$0)	0.641	1.899	0.130	<0.001	***
Priority * \$5	-0.041	0.960	0.158	0.797	

Note. ¹This regression model does not control for the first experiment on visible incentive because the assignment of second experiment (i.e., Priority Mail and the second incentive) was fully crossed with the conditions of the first experiment. We tested whether adjusting for the use of a visible incentive would affect the model results and found no influence on the estimated effects of Priority Mail and the second incentive.

We compared respondent characteristics by experimental condition in terms of gender, age, race/ethnicity, household size, and eligibility (i.e., same categories as in **Table 2**). There were no significant differences among respondents on any of the characteristics compared (see

1
2
3
4 **Appendix II**). The results suggest that the second incentive and Priority Mail increased the yield,
5
6
7 but did not change the sociodemographic compositions of the respondents. However, the
8
9
10 comparisons have low statistical power because the numbers of respondents were small in some
11
12
13 categories due to the unequal experimental assignment.
14
15

16
17 Using characteristics of the sampled cases, we further examined whether the effects of
18
19 the second incentive and Priority Mail varied across sample subgroups defined by response
20
21 propensity, the probability of having an undeliverable address, type of address, and Census
22
23 region. We studied the potential varying effects of these two mailing techniques (i.e., the second
24
25 incentive and Priority Mail) by examining whether interaction terms between them and sample
26
27 subgroup indicators were significant in logistic regression models. To be specific, the logistic
28
29 regression model reported in **Table 4**, with only the main effects of the second incentive and
30
31 Priority Mail, was taken as the *restricted model*. Then, for comparison on each sample
32
33 characteristic (e.g., Do the effects of the two techniques vary by LRS tertiles?), a *full model* was
34
35 fitted further including two sets of interaction terms, one set between the sample subgroups and
36
37 the second incentive indicator (e.g., tertiles of LRS * \$5) and the other set between the subgroups
38
39 and the Priority Mail indicator (e.g., tertiles of LRS * Priority Mail). Likelihood ratio tests were
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 used to compare the *restricted model* to the *full models*. If the effects of the two mailing
4
5
6 techniques vary by sample subgroups, the full model should have a significantly better fit than
7
8
9 the restricted model. This was not the case. None of the likelihood ratio tests were statistically
10
11
12 significant, suggesting that the full models did not improve the model fit beyond the restricted
13
14
15 model. (Results of the full models are reported in Appendix III.) Thus, the effects of the second
16
17
18 incentive and Priority Mail did not significantly differ across the sample subgroups.
19
20
21
22

23 *Costs*

24
25
26
27 **Table 5** reports the costs of the three mailing design techniques. With regard to making the
28
29
30 incentive visible, because the envelope with a transparent window was more expensive than a
31
32
33 standard envelope, the envelope cost averaged \$1.25 per respondent in the visible condition and
34
35
36 \$0.67 in the non-visible condition. Even though the cost appeared to be twice as high, since
37
38
39 envelopes are generally inexpensive, the total difference for over 11,000 sampled addresses was
40
41
42
43
44 only \$1,232.
45
46

47
48 In the second incentive and Priority Mail experiment, the cost of the additional incentive
49
50 was \$5; the envelop and postage of Priority Mail was \$8.25, compared to \$1.39 for the First
51
52
53 Class mail. Implementing these two techniques made a big difference in the total cost. Had the
54
55
56
57
58
59
60

second incentive and Priority Mail been sent to all nonrespondents (i.e., no experimental randomization), it would increase the cost by \$247,257, compared to simply using the First Class mail with no additional incentive. Furthermore, the differences, on average for each respondent, were also substantial across the four experimental conditions—the most expensive condition (\$0 Priority; \$123) cost four times more than the cheapest condition (\$0 First Class; \$30).

Table 5. Cost calculation for the two experiments

<i>Visible incentive experiment (1st mailing)</i>				
Envelope cost	Visible \$0.21		Not visible \$0.10	
Sample	11,193		11,188	
# respondents	1,884		1,733	
Cost/respondent	\$1.25		\$0.65	
<i>Second incentive + Priority Mail experiment (nonresponse follow-up mailing)</i>				
Second incentive	\$0	\$0	\$5	\$5
Envelope + Postage	First Class \$1.39	Priority \$8.25	First Class \$1.39	Priority \$8.25
Sample	2,081	2,110	2,087	14,570
# respondents	98	141	179	1,682
Cost/respondent	\$30	\$123	\$75	\$115

Discussion

Summary of the Findings

We experimentally evaluated the effect of a visible incentive in the initial mailing of a screening questionnaire and the effect of a second incentive and Priority Mail in a nonresponse follow-up

1
2
3 mailing, as a part of a national push-to-web/mail data collection. We found that all three
4
5
6
7 techniques were useful in increasing response rates to the screening questionnaire.
8
9

10 The visible incentive increased response rates by 1.3 percentage points over the non-
11
12
13 visible condition. This positive effect echoed findings of previous research (Bilgen et al., 2021;
14
15
16
17 DeBell et al., 2020; Sherr & Wells, 2021). In terms of respondent composition, the visible
18
19
20 incentive recruited more Hispanic and younger (<50) respondents. However, the effect on the
21
22
23 age of the screening respondents did not translate into age-based household eligibility (18-49), as
24
25
26
27 there was no difference in the eligibility rate between the visible and not visible conditions.
28
29

30 We were also interested in how the visible incentive differentially attracted cases with
31
32
33 different response propensities. The results showed that the visible incentive was the most
34
35
36
37 effective in swaying cases with moderate response propensities to participate. For those who had
38
39
40 high and low response propensities, the effect of the visible incentive was weaker. This could be
41
42
43 because households with high- and low-response propensity had certain mail sorting behaviors
44
45
46
47 (e.g., opening all mail, opening only personally relevant mail) that were not changed by the
48
49
50 visible incentive. It could also be because these households' participation decisions were driven
51
52
53
54 by other reasons and could not be easily affected by whether they opened the mail. Given the
55
56
57
58
59
60

1
2
3 results, the visible incentive seemed to be useful for increasing yield, but it probably had limited
4
5
6
7 effect in balancing response propensities and reducing the risk of nonresponse bias.
8
9

10 For the variables related to mailing operations, we found that the effect of the visible
11
12
13 incentive was smaller among cases with higher probabilities of being undeliverable. At the same
14
15
16
17 time, the effect was also smaller among cases in the Midwest and the South than in the West and
18
19
20 the Northeast; reasons for this regional variation warrant future investigation. Given these
21
22
23 differential effects and the higher cost of the windowed envelope, one strategy with improved
24
25
26
27 cost efficiency is to use the visible incentive only on addresses predicted to be deliverable. We
28
29
30 encourage future research to replicate our results and incorporate these operational variables
31
32
33 when examining the effects of alternative mailing techniques.
34
35
36

37 With regard to the use of the second incentive and Priority Mail, they additively
38
39
40 increased response rates, suggesting that it is beneficial to incorporate both techniques in the
41
42
43
44 nonresponse follow-up mailing if a survey budget allows for it. However, the Priority Mail had a
45
46
47 significantly weaker effect on response rates than the \$5 second incentive, despite being more
48
49
50
51 expensive (each piece of Priority Mail costs \$6.86 more than First Class mail). Thus, the
52
53
54 incentive would be a better investment if only one of the two techniques were to be chosen. In
55
56
57
58
59
60

1
2
3
4 fact, when weighing the yield against the cost, we found the use of Priority Mail to be
5
6
7 inefficient—it cost \$123 per respondent, a high price compared to \$30 per respondent using the
8
9
10 regular First Class mail.

11
12
13 In terms of respondent composition, we observed that the second incentive and Priority
14
15
16 Mail did not significantly influence the demographic distributions of the respondents, nor did
17
18
19 their effects differ across the sample subgroups. Thus, the two nonresponse mailing techniques
20
21
22 were effective in improving response rates, but we failed to find evidence that they affected the
23
24
25 risk of nonresponse biases. We note that the lack of influence on representativeness should be
26
27
28 interpreted with caution, because the unequal experimental assignment renders the response data
29
30
31 low statistical power in detecting potential differences.
32
33
34
35
36
37
38
39
40

41 *Practical Implications and Future Research*

42
43 Making the prepaid incentive visible cost an extra \$1,232 (for over 11,000 sample cases), a price
44
45
46 that we considered acceptable when it resulted in a 1.3 p.p. increase in the response rate to the
47
48
49 screening questionnaire. One concern about making the cash visible is that the mail may have a
50
51
52 higher risk of being intercepted. One way to test this concern is to compare the proportion of
53
54
55
56
57
58
59
60

1
2
3 returned mail in the visible and non-visible conditions. Assuming that the actual undeliverable
4
5
6 rates were the same, a lower mail return rate in the visible condition suggests potential mail
7
8
9 interception. We found that the undeliverable rate was 3.9% in the visible condition, compared to
10
11
12 4.2% in the non-visible condition. Though this difference is statistically significant ($p= 0.046$)
13
14
15 due to the large sample size and the high statistical power, it is substantively small and thus we
16
17
18 do not consider it as a major risk.
19
20
21
22

23
24 For nonresponse follow-up mailings, both the second incentive and Priority Mail were
25
26 found to be effective for increasing response rates, but they did not affect respondent
27
28 composition. Because the costs, particularly of using Priority Mail, were substantial, adding
29
30 these additional design elements did not seem to be cost efficient for the follow-up mailing, at
31
32 least in the current form of implementation. However, some modifications might improve the
33
34 performance of the two mailing techniques. For example, because our current finding
35
36 demonstrated the benefit of making incentives visible, the effect of the second incentive might be
37
38 further boosted by making the cash visible. Furthermore, to reduce the cost of Priority Mail and
39
40 still preserve its benefit, it might be worth testing a form of mailing that has the look of Priority
41
42 Mail but is posted with First Class service, because the effect of Priority Mail likely comes from
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4 its appearance rather than the speed of delivery. These steady advances in optimizing the mailing
5
6
7 process of push-to-web/mail surveys may encourage their increased use for creating more and
8
9
10 better data about the social world.
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Review Only

References

- American Association for Public Opinion Research. 2016. Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys, (Ninth Edition). Retrieved from: [https://www.aapor.org/Standards-Ethics/Standard-Definitions-\(1\).aspx](https://www.aapor.org/Standards-Ethics/Standard-Definitions-(1).aspx).
- Baker, R., Blumberg, S. J., Brick, J. M., Couper, M. P., Courtright, M., ... & Zahs, D. (2010). Research synthesis: AAPOR report on online panels. *Public Opinion Quarterly*, 74(4), 711-781.
- Biemer, P. P., Murphy, J., Zimmer, S., Berry, C., Deng, G., & Lewis, K. (2018). Using bonus monetary incentives to encourage web response in mixed-mode household surveys. *Journal of Survey Statistics and Methodology*, 6(2), 240-261.
- Bilgen, I., Dutwin, D., Singh, R., & Hendarwan, E. (2021). Peekaboo! The impact of displaying cash via window envelope during mail contact when recruiting to a probability-based panel. Paper presented at the annual conference of the Midwest Association for Public Opinion Research, Chicago, IL, November.
- Braekman, E., Demarest, S., Charafeddine, R., Drieskens, S., Berete, F., Gisle, L., ... & Van Hal, G. (2022). Unit Response and Costs in Web Versus Face-To-Face Data Collection:

1
2
3 Comparison of Two Cross-sectional Health Surveys. *Journal of Medical Internet*
4
5
6
7 *Research*, 24(1), e26299.
8
9

10 Brick, J. M., & Tourangeau, R. (2017). Responsive survey designs for reducing nonresponse
11
12
13 bias. *Journal of Official Statistics*, 33(3), 735-752.
14
15

16
17 Bucks, B., Couper, M. P., & Fulford, S. L. (2020). A mixed-mode and incentive experiment
18
19
20 using administrative data. *Journal of Survey Statistics and Methodology*, 8(2), 352-369.
21
22

23 Centers for Disease Control and Prevention. (2020). National Survey of Family Growth
24
25
26 [electronic resources]. <https://www.cdc.gov/nchs/nsfg/index.htm>.
27
28

29
30 Couper, M. P. (2008). *Designing Effective Web Surveys*. New York: Cambridge University
31
32
33 Press.
34
35

36
37 Daikeler, J., Bošnjak, M., & Lozar Manfreda, K. (2020). Web versus other survey modes: an
38
39
40 updated and extended meta-analysis comparing response rates. *Journal of Survey*
41
42
43 *Statistics and Methodology*, 8(3), 513-539.
44
45

46
47 DeBell, M., & Maisel, N. (2018). Methodology of the ANES 2016 non-response follow-up
48
49
50 study. *American National Election Studies*. Retrieved from:
51
52
53
54 <https://electionstudies.org/wp->
55
56
57

1
2
3
4 [content/uploads/2018/12/anes_timeseries_2016_nonresponse_follow-](#)
5
6
7 [up_study_methodology.pdf](#)
8
9

10 Debell, M., Maisel, N., Edwards, B., Amsbary, M., & Meldener, V. (2020). Improving Survey
11
12
13 Response Rates with Visible Money. *Journal of Survey Statistics and Methodology*, 8(5),
14
15
16
17 821-831.
18

19
20 [Dillman, D. A. \(2020\). Towards survey response rate theories that no longer pass each other like](#)
21
22
23 [strangers in the night](#). In P. S. Brenner (Ed.). *Understanding Survey Methodology:*
24
25
26
27 *Sociological theory and applications* (pp. 15-44). Springer, Cham.
28
29

30 Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode*
31
32
33 *surveys: the tailored design method*. John Wiley & Sons.
34
35
36

37 Dirksz, G., Lavrakas, P.J., Lusskin, L., Ponce, B., and Felstead, P. (2017). Testing the Effects of
38
39
40 Token Noncontingent Incentives in Follow-up Mailings in a Mixed Mode Survey. Paper
41
42
43
44 presented at the annual conference of the American Association for Public Opinion
45
46
47 Research, New Orleans, LA, May.
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Dykema, J., Jaques, K., Cyffka, K., Assad, N., Hammers, R. G., Elver, K., ... & Stevenson, J.

4
5
6
7 (2015). Effects of sequential prepaid incentives and envelope messaging in mail

8
9
10 surveys. *Public Opinion Quarterly*, 79(4), 906-931.

11
12
13 Erdman, C., & Bates, N. (2017). The Low Response Score (LRS) A Metric to Locate, Predict,

14
15
16 and Manage Hard-to-Survey Populations. *Public Opinion Quarterly*, 81(1), 144-156.

17
18
19
20 Feld, S., Frenzen, H., Krafft, M., Peters, K., & Verhoef, P. C. (2013). The effects of mailing

21
22
23 design characteristics on direct mail campaign performance. *International Journal of*

24
25
26
27 *Research in Marketing*, 30(2), 143-159.

28
29
30 Groves, R. M., & McGonagle, K. A. (2001). A Theory-Guided Interviewer Training Protocol

31
32
33 Regarding'survey Participation. *Journal of Official Statistics*, 17(2), 249.

34
35
36 Groves, R. M., Cialdini, R. B., & Couper, M. P. (1992). Understanding the decision to

37
38
39 participate in a survey. *Public Opinion Quarterly*, 56(4), 475-495.

40
41
42
43 Han, D., Montaquila, J. M., & Brick, J. M. (2013). An evaluation of incentive experiments in a

44
45
46
47 two-phase address-based sample mail survey. *Survey Research Methods*, 7(3), 207-218.

1
2
3
4 Heerwegh, D. (2009). Mode differences between face-to-face and web surveys: an experimental
5
6
7 investigation of data quality and social desirability effects. *International Journal of Public*
8
9
10 *Opinion Research*, 21(1), 111-121.

11
12
13 Jackson, M. T., McPhee, C. B., & Lavrakas, P. J. (2020). Using response propensity modeling to
14
15
16 allocate noncontingent incentives in an address-based sample: Evidence from a national
17
18
19 experiment. *Journal of Survey Statistics and Methodology*, 8(2), 385-411.

20
21
22
23
24 Keusch, F. (2015). Why do people participate in Web surveys? Applying survey participation
25
26
27 theory to Internet survey data collection. *Management Review Quarterly*, 65(3), 183-216.

28
29
30 Kolsaker, A., Görtz, D., & Gilbert, D. (2016). Making sense of metaphor: The impact of target-
31
32
33 group-specific headlines in direct mail. *Journal of Marketing Communications*, 22(1), 56-
34
35
36 82.

37
38
39
40 Lesscher, L., Lobschat, L., & Verhoef, P. C. (2021). Do offline and online go hand in hand?
41
42
43 Cross-Channel and synergy effects of direct mailing and display
44
45
46 advertising. *International Journal of Research in Marketing*, 38(3), 678-697.

1
2
3
4 Link, M. W., & Burks, A. T. (2013). Leveraging auxiliary data, differential incentives, and
5
6
7 survey mode to target hard-to-reach groups in an address-based sample design. *Public*
8
9
10 *Opinion Quarterly*, 77(3), 696-713.

11
12
13 Lujkx, R., Jónsdóttir, G. A., Gummer, T., Ernst Stähli, M., Frederiksen, M., Ketola, K., ... &

14
15
16
17 Wolf, C. (2021). The European Values Study 2017: On the way to the future using
18
19
20 mixed-modes. *European Sociological Review*, 37(2), 330-346.

21
22
23 Medway, R., Paek, C., Battle, D., & McPhee, C. (2021). Exploring experiences with and
24
25
26
27 attitudes toward survey research: Results of qualitative interviews with nonrespondents to
28
29
30 a national household survey. Presented at the annual conference of the American
31
32
33
34 Association for Public Opinion Research, Virtual Conference, May.

35
36
37 Mercer, A., Caporaso, A., Cantor, D., & Townsend, R. (2015). How much gets you how much?
38
39
40 Monetary incentives and response rates in household surveys. *Public Opinion*
41
42
43 *Quarterly*, 79(1), 105-129.

44
45
46
47 Messer, B. L., & Dillman, D. A. (2011). Surveying the general public over the internet using
48
49
50 address-based sampling and mail contact procedures. *Public Opinion Quarterly*, 75(3),
51
52
53
54 429-457.

1
2
3
4 Murphy, J., Biemer, P., & Berry, C. (2018). Transitioning a survey to self-administration using
5
6
7 adaptive, responsive, and tailored (ART) design principles and data visualization. *Journal*
8
9
10 *of Official Statistics*, 34(3), 625-648.

11
12
13 Nishimura, R., Gatward, R., West, B.T., and Saw, H. (2021). USPS Mail Delivery Patterns for a
14
15 National Address-Based Sample in the Time of COVID-19. Paper presented at the annual
16
17 conference of the American Association for Public Opinion Research, Virtual
18
19 Conference, May.

20
21
22
23 Pew Research Center. (2021). *Internet/Broadband fact sheet*. Retrieved from:

24
25
26 <https://www.pewresearch.org/internet/fact-sheet/internet-broadband/>
27

28
29
30 Scardaville, M., Dias, P., Paek, C., Pugliese, C., Medway, R., & Zukerberg, A. (2021). What
31
32 drives survey nonresponse? Using qualitative research to develop typologies of
33
34 nonresponse. Presented at the annual conference of the American Association for Public
35
36 Opinion Research, Virtual Conference, May.
37
38
39
40
41
42

43 Sherr, S. A. & Wells, B. M. (2021). What you see is what you get: Evaluating the use of visible
44
45 incentives in the California Health Interview Survey. Paper presented at the annual
46
47 conference of the American Association for Public Opinion Research, Virtual
48
49 Conference, May.
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Singer, E., & Ye, C. (2013). The use and effects of incentives in surveys. *The ANNALS of the*
4
5
6
7 *American Academy of Political and Social Science*, 645(1), 112-141.

8
9
10 Wagner, J., West, B. T., Couper, M. P., Zhang, S., Gatward, R., Nishimura, R., Saw, H. W.
11
12
13 (2022). An experimental evaluation of two approaches for improving response to
14
15
16 household screening efforts in national mail/web surveys. *Journal of Survey Statistics and*
17
18
19 *Methodology*.

20
21
22
23 West, B. T., Axinn, W. G., Couper, M. P., Gatny, H., & Schroeder, H. (2022). A Web-based
24
25
26
27 Event History Calendar Approach for Measuring Contraceptive Use Behavior. *Field*
28
29
30 *methods*, 34(1), 3-19.

31
32
33
34 West, B., Couper, M. P., Axinn, W. G., Wagner, J., Gatward, R., Saw, H. W., & Zhang, S.
35
36
37 (2022). Evaluating a web/mail alternative to a national face-to-face survey: Initial results
38
39
40 from the American Family Health Study. Paper presented at the annual conference of the
41
42
43 American Association for Public Opinion Research, Chicago, IL, May.

44
45
46
47 Williams, D., & Brick, J. M. (2018). Trends in US face-to-face household survey nonresponse
48
49
50 and level of effort. *Journal of Survey Statistics and Methodology*, 6(2), 186-211.

Appendix I

American Family Health Study Web Screener Questionnaire

What language would you like to complete this survey in?

1. English
2. Spanish

In order to select someone in the household to complete the full AFHS survey, please provide the names (or initials) of the people who usually live here. **Please include any unmarried children aged 18 or older who are away from home living on or near a college campus.**

First, including yourself, how many adults aged 18 or older live in this household (including those living away)?

Let's start with you; what is your name?

Enter your name or initials in the box below

First Name, Last Name

What is your gender?

1. Male
2. Female
3. Other

[online respondents ONLY - IF gender not given]

Please provide gender if possible. All answers will be kept confidential

In what year were you born?

1
2
3
4 **[online respondents ONLY - IF birth year not provided]**

5 Please provide birth year if possible. All answers will be kept confidential.
6
7

8 We need to know a little more to determine if you are eligible for the survey
9

10 What is your birth month?
11
12
13
14

15 What is your birth day?
16
17
18
19

20 Do you consider yourself to be Hispanic or Latino?
21
22

- 23 1. Yes
24 2. No
25
26
27
28
29

30 **[online respondents ONLY - IF Ethnicity not given]**
31

32 Please provide ethnicity if possible. All answers will be kept confidential.
33
34
35
36

37 What race do you consider yourself to be?
38

39 Please select all options that apply
40
41

- 42 1. American Indian or Alaska Native
43 2. Asian
44 3. Native Hawaiian or other Pacific Islander
45 4. Black or African American
46 47 5. White
48
49
50
51

52 **[online respondents ONLY - IF Race not given]**
53

54 Please provide race if possible. All answers will be kept confidential.
55
56
57
58
59
60

1
2
3
4
5 **[The following questions are repeated as necessary for household members 18 years + living at**
6 **the address]**
7
8
9

10 Now that you have answered each of these questions for yourself, please answer the same
11 questions for each member of your household **aged 18 and above that usually lives here.**
12
13
14

15
16
17 What is [THIS / THE FIRST/SECOND ...] person's name?"
18
19

20
21
22 First Name
23

24
25 Last Name
26
27

28
29
30 What is [NAME]'s gender?
31

32 Select one option below
33

- 34
35 1. Male
36
37 2. Female
38
39 3. Other
40

41
42 **[online respondents ONLY - IF gender not given]**
43

44 Please provide gender if possible. All answers will be kept confidential
45
46

47 What is [NAME]'s birth year?
48
49

50 **[online respondents ONLY - IF birth year not provided]**
51

52 Please provide birth year if possible. All answers will be kept confidential.
53
54

55 We need to know a little more to determine if [NAME] is eligible for the survey
56
57
58
59

1
2
3 What is [NAME] birth month?
4
5
6
7

8 What is [NAME] birth day?
9
10
11
12

13 Is [NAME] Hispanic or Latino?
14
15

- 16 1. Yes
 - 17 2. No
- 18
19
20
21
22

23 **[online respondents ONLY - IF ethnicity not given]**
24

25 Please provide ethnicity if possible. All answers will be kept confidential.
26
27
28
29

30 What race is [NAME]?
31

32 Please select all options that apply
33
34

- 35 1. American Indian or Alaska Native
 - 36 2. Asian
 - 37 3. Native Hawaiian or other Pacific Islander
 - 38 4. Black or African American
 - 39 5. White
- 40
41
42
43
44
45
46

47 **[online respondents ONLY - IF Race not given]**
48

49 Please provide race if possible. All answers will be kept confidential.
50
51
52
53

54 Thank you for completing this questionnaire. [One of your household members/You] may be
55 selected to participate in our main survey; in return for [their/your] participation, we will give
56
57
58
59
60

1
2
3 [them/you] up to \$70 as a token of our appreciation. If chosen, we will make contact with more
4 information.
5
6

7
8 For now, please enter the best email at which you can be reached below. We will only send
9 emails related to this survey.
10
11

12
13
14
15 Next, what is your telephone number?

16 We will only call you about this survey, and your number will not be shared with anyone else.
17
18

19
20
21
22 This is a...

- 23
24 1. mobile telephone number
25 2. home telephone number
26
27
28
29

30 Do you give us permission to send you text messages related to the survey?

31 We will not use this number for any other purposes.

- 32
33 1. Yes
34 2. No
35
36
37

38 We would also like to collect this information for [NAME selected R].
39
40

41
42 Please enter [NAME selected R]'s email address. We will only send emails related to this
43 survey.
44
45

46
47 Next, what is [NAME selected R]'s telephone number?

48 We will only call them about this survey, and their number will not be shared with anyone else.
49
50

51
52
53 This is a...

- 54 1. mobile telephone number
55
56
57
58
59
60

2. home telephone number
3. I don't know

For Review Only

Appendix II

This appendix reports respondents' demographic composition by the second incentive and Priority Mail experimental condition.

Table II.1. Demographic characteristics of respondents by the second incentive and Priority Mail experimental condition

	\$0 First Class		\$0 Priority		\$5 First Class		\$5 Priority		
	n	%	n	%	n	%	n	%	
Gender									
Female	61	64.2%	98	70.5%	114	64.8%	1018	61.8%	$\chi^2(3)=4.6,$ p=0.204
Male	34	35.8%	41	29.5%	62	35.2%	629	38.2%	
Age									
<25	5	5.2%	5	3.6%	7	4.0%	82	5.0%	$\chi^2(9)=7.6,$ p=0.572
25-34	10	10.4%	17	12.2%	17	9.8%	244	14.8%	
35-49	22	22.9%	32	23.0%	50	28.7%	399	24.3%	
50+	59	61.5%	85	61.2%	100	57.5%	919	55.9%	
Race/ethnicity									
Hispanic	14	14.6%	18	12.9%	25	14.4%	206	12.5%	$\chi^2(9)=4.8,$ p=0.842
Non-H Black	11	11.5%	16	11.5%	28	16.1%	261	15.8%	
Non-H Other	8	8.3%	15	10.8%	17	9.8%	138	8.4%	
Non-H White	63	65.6%	90	64.7%	104	59.8%	1046	63.4%	
# of adults									
1	28	28.6%	47	33.3%	53	29.6%	487	29.2%	$\chi^2(9)=6.6,$ p=0.678
2	52	53.1%	61	43.3%	82	45.8%	772	46.4%	
3	11	11.2%	22	15.6%	23	12.8%	209	12.6%	
4+	7	7.1%	11	7.8%	21	11.7%	197	11.8%	
Eligibility									
Not eligible	46	46.9%	70	49.6%	88	49.2%	765	45.5%	$\chi^2(3)=1.7,$ p=0.643
Eligible	52	53.1%	71	50.4%	91	50.8%	917	54.5%	

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Review Only

Appendix III

This appendix reports the *full models* that tested the interaction effects between the experimental treatments (i.e., Priority Mail and \$5) and sample subgroups. As stated in the main text, none of the full models was significantly better fitted than the *restricted model* in Table 4. Correspondingly, none of the interaction terms was statistically significant.

Table III.1. Logistic regression results: full models

<i>Subgroup = tertile of LRS</i>	Estimate	Std. Error	p-value
Intercept	-2.843	0.128	<0.0001
Priority (vs. First Class)	0.500	0.120	<0.0001
\$5 (vs. \$0)	0.440	0.118	<0.0001
Medium LRS (vs. Low LRS)	-0.065	0.182	0.721
High LRS (vs. Low LRS)	-0.471	0.202	0.020
Priority * Medium LRS	-0.272	0.166	0.100
Priority * High LRS	-0.176	0.183	0.335
\$5 * Medium LRS	0.277	0.173	0.109
\$5 * High LRS	0.286	0.191	0.136
<i>Subgroup = undeliverable probability</i>	Estimate	Std. Error	p-value
Intercept	-2.983	0.082	<0.0001
Priority (vs. First Class)	0.378	0.074	<0.0001
\$5 (vs. \$0)	0.604	0.077	<0.0001
High undeliverable (vs. remaining)	-0.111	0.283	0.695
Priority * High undeliverable	-0.432	0.245	0.078
\$5 * High undeliverable	0.162	0.290	0.577
<i>Subgroup = type of address</i>	Estimate	Std. Error	p-value
Intercept	-2.445	0.778	0.002
Priority (vs. First Class)	-0.589	0.791	0.457
\$5 (vs. \$0)	0.946	0.903	0.295
Multi-family (vs. PO Box)	-0.388	0.790	0.623
Single-family (vs. PO Box)	-0.619	0.784	0.429
Priority * Multi-family	0.770	0.802	0.337
Priority * Single-family	1.004	0.796	0.207
\$5 * Multi-family	-0.512	0.914	0.575

\$5 * Single-family	-0.260	0.908	0.774
<i>Subgroup = Census region</i>	Estimate	Std. Error	p-value
Intercept	-3.005	0.182	<0.0001
Priority (vs. First Class)	0.594	0.168	<0.0001
\$5 (vs. \$0)	0.563	0.165	0.001
Northeast (vs. Midwest)	-0.066	0.277	0.812
South (vs. Midwest)	-0.034	0.219	0.875
West (vs. Midwest)	0.134	0.241	0.578
Priority * Northeast	-0.196	0.251	0.433
Priority * South	-0.415	0.197	0.036
Priority * West	-0.185	0.225	0.411
\$5 * Northeast	0.059	0.257	0.819
\$5 * South	0.222	0.203	0.275
\$5 * West	-0.211	0.222	0.341

For Review Only