An Evaluation of Prepaid and Promised Incentive Experiment in a Mail Push-to-Web Survey

Daifeng Han[1], Tien-Huan Lin[1], J. Michael Brick[1], David Kashihara[2], Sadeq Chowdhury[2]
[1]Westat, 1600 Research Blvd., Rockville, MD 20850
[2]Agency for Healthcare Research and Quality, 5600 Fishers Lane, Rockville, MD 20857

Abstract
Address-based sampling (ABS) with mail push-to-web data collection mode has become more widely used in the United States. In 2022, we tested an ABS mail push-to-web sequential mailing design with a stratified sample based on whether a household was likely to have a person of age 65 or older. The short survey included health measures and basic demographic questions. A random experiment of prepaid and promised monetary incentives was embedded in the design with four treatment groups: (1) no incentive; (2) prepaid $2 and promised $10; (3) prepaid $5 only; and (4) prepaid $5 and promised $10. The paper evaluates the effects of these incentives on overall response rates, timing of response and data collection cost, and composition of the responding sample. There is little research about the impact of the combined effect of prepaid and promised incentives in mail push-to-web surveys, making these findings of great interest to ABS designs with this data collection protocol.

1. Introduction

Mail push-to-web data collection has become widely used in the United States. By "mail push-to-web" we mean that the questionnaire modes are presented sequentially, whereby web is the only mode available initially, and then the paper questionnaire is introduced during nonresponse follow-up (NRFU). Compared with face-to-face mode, the cost for mail push-to-web data collection is typically much lower. For a national study, face-to-face data collection would require a clustered sample, but the mail push-to-web approach can be based on a single-stage address sample, eliminating the clustering effect at the household level. Also, the web and paper modes may give target respondents more privacy protection because no interviewer is required to administer the survey instrument.

Despite its advantages, two major concerns about the mail push-to-web data collection mode are lower response rate and higher potential nonresponse bias. The literature has shown that monetary incentives could boost cooperation and bring in people who otherwise might not respond (Groves, Singer, and Corning 2000; Singer and Ye 2013). The relationship between incentive amount and response rate tended to be nonlinear (Edwards et al., 2005). In the current social environment where response rates are dropping annually, it is unclear what the effective incentive levels should be. Although there has been consistent evidence that prepaid incentives increase response rates, particularly for mail surveys (Mercer et al., 2015), there has been little research on the combined effect and trade-off between prepaid and promised incentives.

To fill in these gaps in the literature, we conducted a random experiment involving both prepaid and promised monetary incentives through the American Household Medical Study (AHMS). The study was sponsored by the Agency for Healthcare Research and Quality and fielded by Westat during the fall of 2022. There were four incentive experiment conditions: (1) prepaid $0 and promised $0; (2) prepaid $5 and promised $0; (3) prepaid $2 and promised $10; and (4) prepaid $5 and promised $10. We used visible cash (i.e., cash that is visible from outside the envelope) for the prepaid incentives (Debell et al. 2019) because the literature suggested that including visible money could increase the response rate by several percentage points at no significant cost.

After describing the study design and incentive conditions in Section 2, we will discuss the impact of incentives on the final response rate, timing of response, data collection cost, and composition of the responding sample in Section 3. We will summarize our findings and identify future research areas in Section 4.

2. Study Design and Incentive Experiment

The purpose of the AHMS was not to generate any national estimates, but to test the feasibility of using a 10-minute, mail push-to-web survey to screen for some subpopulations of interest. The questionnaire included two sets of items. The first set of items was health related, with the goal to identify households with potentially higher than average healthcare utilization and expenditures. The second set of items was about race and ethnicity, with the inquiry made at the household level (through questions like "Is anyone in the household of Hispanic, Latino, or Spanish origin?") to avoid the burden of enumerating all the household members. As this was a pilot test, only the English-language questionnaire was provided. The web questionnaire was programmed with an optimized design so that it could be completed via a computer or a mobile device.

2.1 Sampling Method

Given the interest in targeting households with potentially high medical expenditures, we selected the sample in two phases using Medicare administrative data from the Centers for Medicare and Medicaid Services (CMS). The initial sample was a nationally representative simple random sample from the United States Postal Service (USPS) Computerized Delivery Sequence File (CDSF) containing standardized mailing address fields. We then conducted address matching between the CDSF sample and the Medicare beneficiary addresses from the CMS. As a result, the CDSF sample was sub-stratified into two groups — those matched to the Medicare beneficiary addresses from the CMS and those not. We will refer to these groups as CMS stratum and non-CMS stratum, respectively, in the remainder of this paper. Most of the addresses in the CMS stratum were expected to contain at least one person of age 65 or older, and most of the addresses in the non-CMS stratum were expected to contain only persons of age 64 or younger.[1] When selecting the final sample for data collection, we applied a higher sampling rate for the CMS stratum than for the non-CMS stratum so that households with persons of age 65 or older were oversampled.

[1] This was confirmed through the responding sample. Among the responding households in the CMS stratum, 85.8 percent of the households reported having at least one person of age 65 years or older. Among the responding households in the non-CMS stratum, only 16.0 percent of the households reported having at least one person of age 65 years or older.

2.2 Incentive Experiment

The incentive experiment involved both prepaid and promised incentives. The prepaid cash incentive included three conditions: $0, $2, and $5. Although the literature did not recommend the $0 option (Church 1993; Groves, Singer, and Corning 2000; Lesser et al. 2001), this condition served as a baseline for comparison. The promised incentive was either $0 or $10. We decided not to test the promised $5 condition because we suspected that the promised incentive would not be effective unless the amount was substantially higher than the prepaid incentive. Given the limited budget, we carefully chose the following four prepaid and promised incentive combinations to implement:

- prepaid $0 and promised $0 (referred to as $0/$0)
- prepaid $5 and promised $0 (referred to as $5/$0)
- prepaid $2 and promised $10 (referred to as $2/$10)
- prepaid $5 and promised $10 (referred to as $5/$10)

The addresses in both the CMS and the non-CMS strata were randomly assigned to the four incentive conditions described above. Table 1 shows the sample sizes by sampling stratum and incentive condition.

Table 1: Sample Sizes by Sampling Stratum and Incentive Condition

			Sample size		
Group	*Prepaid*	*Promised*	*CMS stratum*	*Non-CMS stratum*	*Total*
$0/$0	$0	$0	1,504	1,396	2,900
$5/$0	$5	$0	1,505	1,395	2,900
$2/$10	$2	$10	1,504	1,396	2,900
$5/$10	$5	$10	1,505	1,395	2,900

2.3 Mailing Protocol

The AHMS data collection was conducted from September to December 2022. The questionnaire modes were presented sequentially whereby web was the only mode available initially, and then the paper questionnaire was introduced during NRFU attempts. As shown in Figure 1, five contact attempts were made, with the first four through USPS First-Class mailing and the last through FedEx mailing:

- **Advance postcard:** This mailing consisted of a 6 x 9 postcard informing the household of the pending mail invitation to the web survey. For the addresses assigned to the $2 and $5 prepaid incentive conditions, the advance postcard also referred to the prepaid incentive amount to encourage the household to open the upcoming initial invitation.
- **Initial invitation:** The invitation letter included the website URL for access to the web survey. The letter and mailing package were customized according to incentive treatment, with 2,900 sampled addresses in each of the four groups: (1) no incentive and no mention of promised incentive; (2) $5 cash enclosed and no mention of promised incentive; (3) $2 cash enclosed and referencing $10 promised incentive; and (4) $5 cash enclosed and referencing $10 promised incentive. For those with the prepaid incentive, the cash was attached to the letter using a removable glue dot, and placed so that a small amount of the cash would show in the address window.

- **Reminder/thank-you postcard:** This was a sealed postcard reminding the household to complete the screener by web, or if completed, thanking them for their participation.
- **First NRFU mailing with paper questionnaire:** Sampled addresses from which no response had been received after the reminder/thank-you postcard attempt; these addressees were then mailed a follow-up package. Each letter included the web URL for accessing the web survey. A paper questionnaire and a business-reply envelope for returning the questionnaire were also included.
- **Second NRFU mailing with paper questionnaire:** Sampled addresses from which no response had been received after the first NRFU were mailed a final package. The content of the mailing package was similar to that for the first NRFU, except that the letter was phrased differently, and the package was mailed through FedEx to further motivate participation.

The reminder/thank-you postcard and the letters for the two NRFU mailings referenced the promised $10 gift for the incentive treatment groups $2/$10 and $5/$10 shown in Table 1.

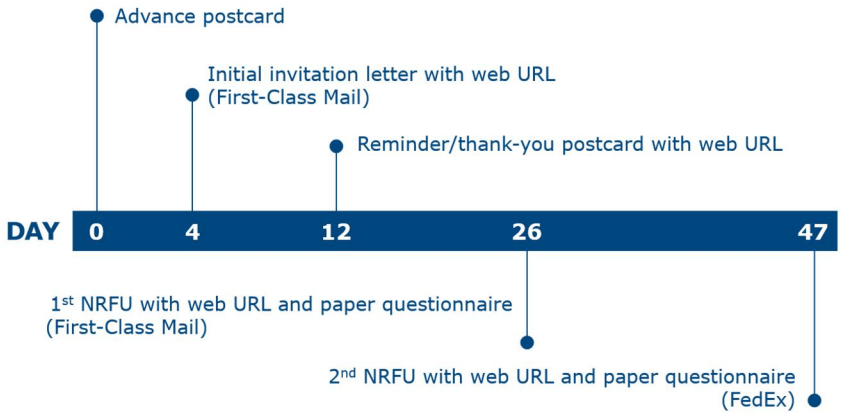

Figure 1: Mailing Attempts for AHMS Data Collection

3. Results

This section summarizes the key results from the AHMS incentive experiment. Section 3.1 examines the final weighted response rates by incentive condition. Section 3.2 evaluates the impact of incentives on the timing of response and data collection cost. Section 3.3 provides some preliminary results on how the incentive condition affected the composition of the responding sample. All the response rates presented in this section are weighted to account for the impact of oversampling in the CMS stratum.

3.1 Impacts on Final Weighted Response Rates

The first set of results focuses on the final weighted response rates. Table 2 shows the status of all the fielded addresses and the weighted response rates by incentive condition, both overall and for the CMS stratum and non-CMS stratum separately. The following overall treatment effects are statistically significant and have meaningful practical implications:

- Conditioning on no promised incentive, the response rate for the prepaid $5 group (23.4 percent) was 10.2 percentage points higher than that for the prepaid $0 group (13.2 percent).
- Conditioning on promised $10, the response rate for the prepaid $5 group (30.4 percent) was 5.3 percentage points higher than that for prepaid $2 group (25.1 percent).
- Conditioning on prepaid $5, the response rate for the promised $10 group (30.4 percent) was 7.0 percentage points higher than that for the promised $0 group (23.4 percent).

Table 2: Weighted Response Rates by Incentive Condition and Sampling Stratum

	Overall	*CMS stratum*	*Non-CMS stratum*
Prepaid $0 and Promised $0			
Fielded Addresses	2,900	1,504	1,396
Ineligible Addresses*	194	56	138
Responding Households	415	286	129
Weighted Response Rate	13.2%	19.7%	10.3%
Prepaid $5 and Promised $0			
Fielded Addresses	2,900	1,505	1,395
Ineligible Addresses*	192	51	141
Responding Households	718	478	240
Weighted Response Rate	23.4%	32.7%	19.1%
Prepaid $2 and Promised $10			
Fielded Addresses	2,900	1,504	1,396
Ineligible Addresses*	204	68	136
Responding Households	740	465	275
Weighted Response Rate	25.1%	32.3%	21.8%
Prepaid $5 and Promised $10			
Fielded Addresses	2,900	1,505	1,395
Ineligible Addresses*	182	42	140
Responding Households	920	587	333
Response Rate	30.4%	38.8%	26.5%

* Ineligible addresses include post office undeliverable addresses and non-residential addresses.

Figure 2 illustrates the point estimates and 95 percent confidence intervals of the weighted response rates for the four incentive conditions, both overall and by sampling stratum. The left panel shows the overall result. First, not providing any monetary incentive appeared to suppress response, as the response rate for the $0/$0 condition was only 13.2 percent. Second, the response rate for the $5/$10 condition (30.4 percent) was meaningfully higher than those for the conditions $5/$0 (23.4 percent) and $2/$10 (25.1 percent). Third, the response rate difference between the $5/$0 and $2/$10 conditions was not statistically significant.

The middle and right panels of Figure 2 show the response rates for the CMS stratum and non-CMS stratum separately. For each incentive condition, the households in the CMS stratum responded substantially better than those in the non-CMS stratum because the former were much more likely to include persons of age 65 or older than the latter.

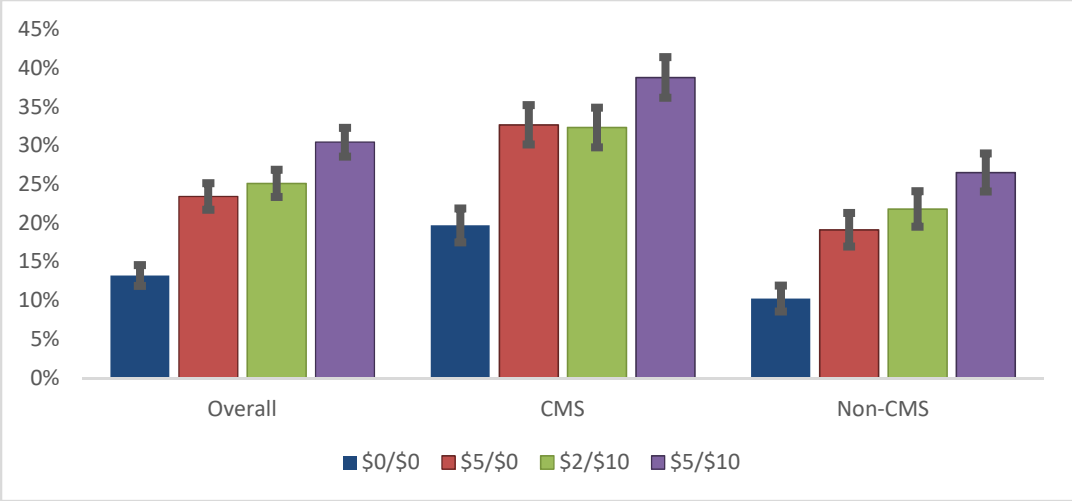

Figure 2: AHMS Weighted Response Rates by Incentive and Stratum

3.2 Impact on Timing of Response and Data Collection Cost

To examine the impact of incentive on the timing of response, we calculated the cumulative weighted response rates by mailing effort, as shown in Figure 3. For each incentive condition, there are four data points, corresponding to the initial invitation, reminder/thank-you postcard, first NRFU with paper questionnaire, and second NRFU with paper questionnaire, respectively. We connected the data points using a dotted line to show the change of the cumulative response rate due to each mailing attempt. The "slopes" for the four lines are very similar whereas the "intercepts" differ substantially across incentive conditions. The response rate to the initial invitation letter with web URL was about 1 percent for the $0/$0 group, 10 percent for the $5/$0 and $2/$10 groups, and 14 percent for the $5/$10 group, respectively. The monetary incentives not only increased the final response rates as shown in Section 3.1, but also motivated earlier response and helped reduce the NRFU cost. As a side note, Figure 3 also shows that the two NRFUs with paper questionnaire boosted the final response rates substantially for all the incentive conditions. This is because the topic of the survey probably appeared more salient to households with older persons (Groves, Singer, and Corning 2000), and the paper mode made it easier for them to respond than the web mode.

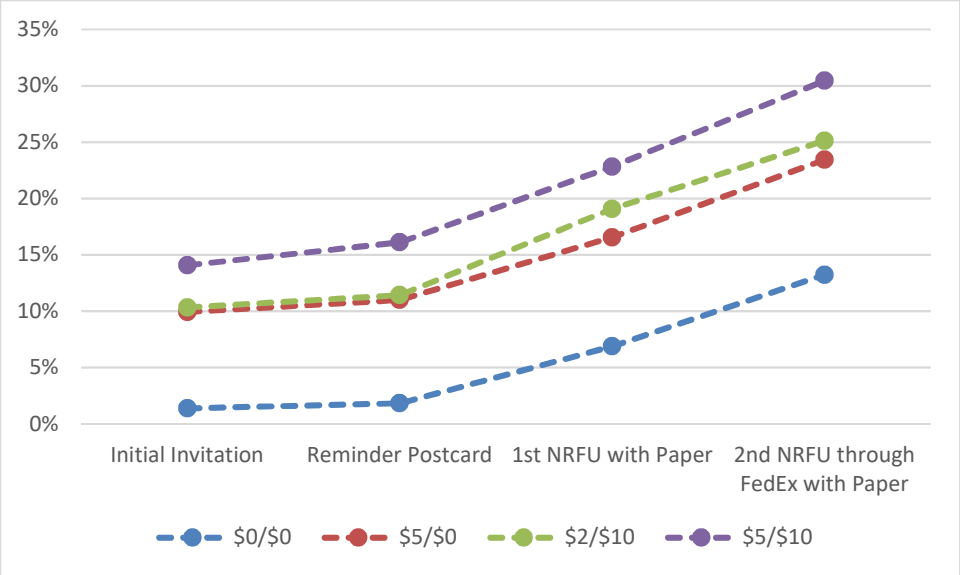

Figure 3: AHMS Cumulative Weighted Response Rates by Mailing Effort

Due to its impact on the final response rates and timing of response, the incentive experiment also has cost implications. For the purpose of this research, we accounted for the material, printing, mailing, and incentive costs. For each incentive condition, we calculated the average cost per complete (referred to as unit cost), and then computed the relative unit cost using the $2/$10 group (which had the lowest unit cost) as the comparison basis. That is, the relative unit cost for a particular incentive condition was the ratio of the unit cost for that condition versus the unit cost for the $2/$10 condition. Table 3 shows the overall relative unit cost for each incentive condition, and breaks it down into two components — the cost for monetary incentive, and the cost for materials, printing, and mailing. Interestingly, the highest overall relative unit cost was associated with the $0/$0 condition, indicating that the average cost per complete for the $0/$0 group was approximately 20 percent higher than for the $2/$10 group. This is due to the extremely low response to the initial invitation, which drove up the material, printing, and mailing costs for the NRFU attempts. In contrast, although the $5/$10 condition was costly from the perspective of incentives per se, its overall relative unit cost was approximately the same as for the $5/$0 condition. This is because the higher incentives not only boosted the final response rate but also motivated early response.

This cost analysis was based on the unweighted counts. We think that compared to the weighted analysis, the unweighted analysis underestimated the overall unit cost for the $0/$0 condition to a greater extent than for the $5/$10 condition. As described in Section 2.1, the proportion of CMS addresses in the sample was substantially higher than that in the population due to the oversampling in the CMS stratum. Table 2 in Section 3.1 shows that although the CMS stratum responded better than the non-CMS stratum for all the incentive conditions, their difference was more substantial under the $0/$0 condition than under the $5/$10 condition. That is, the ratios of the CMS-stratum response rates versus the non-CMS-stratum response rates were 1.9 (calculated as 19.7% divided by 10.3%) for the $0/$0 group and 1.5 (38.8% divided by 26.5%) for the $5/$10 group, respectively. Therefore, the weighted cost analysis is expected to reveal even greater cost-saving effects of the higher monetary incentive groups than the unweighted analysis.

Table 3: Impact of Incentive on Relative Unit Cost

	Relative unit cost compared to \$2/\$10		
	\$0/\$0	\$5/\$0	\$5/\$10
Overall cost	1.2	1.1	1.1
Material, printing, and mailing costs	1.6	1.1	1.0
Monetary incentive cost	0.0	1.1	1.4

3.3 Impact on Sample Composition

Our final research question was whether and how the incentive treatment affected the composition of the respondent sample. We calculated weighted estimates of selected characteristics of the responding households using the basic design weights. The first set of estimates were based on the questionnaire items, including:

- **65+ year old:** presence of a person of age 65 or older;
- **Cancer:** presence of a person who had ever been diagnosed with cancer;
- **Function limitation:** presence of a person with difficulties walking, climbing stairs, grasping objects, reaching overhead, lifting, bending or stooping, or standing for long periods of time due to impairment or physical or mental health problem;
- **Arthritis:** presence of a person who had ever been diagnosed with arthritis;
- **Heart disease:** presence of a person who had ever been diagnosed with heart disease; and
- **Mental health:** presence of a person who received mental health care in the past 12 months.

Figure 4 shows the comparison of these estimates across the incentive conditions. The \$0/\$0 group was associated with higher estimates for the presence of 65+ year olds, function limitation, cancer, arthritis, and heart disease than the non-zero incentive conditions. In contrast, the mental health care estimate was the lowest for the \$0/\$0 condition. These differences were not statistically significant (which was partly due to the small sample sizes) except for the presence of cancer. However, we think the nominal differences were in the expected direction and largely driven by the household age composition. That is, older households (i.e., those with a person of age 65 or older) responded better than younger households (i.e., those without a person of age 65 or older) under the \$0/\$0 condition. People in the older households tended to have more physical health problems but were less likely to seek mental health care than those in the younger households. There was limited evidence that the monetary incentives boosted the response from younger households.

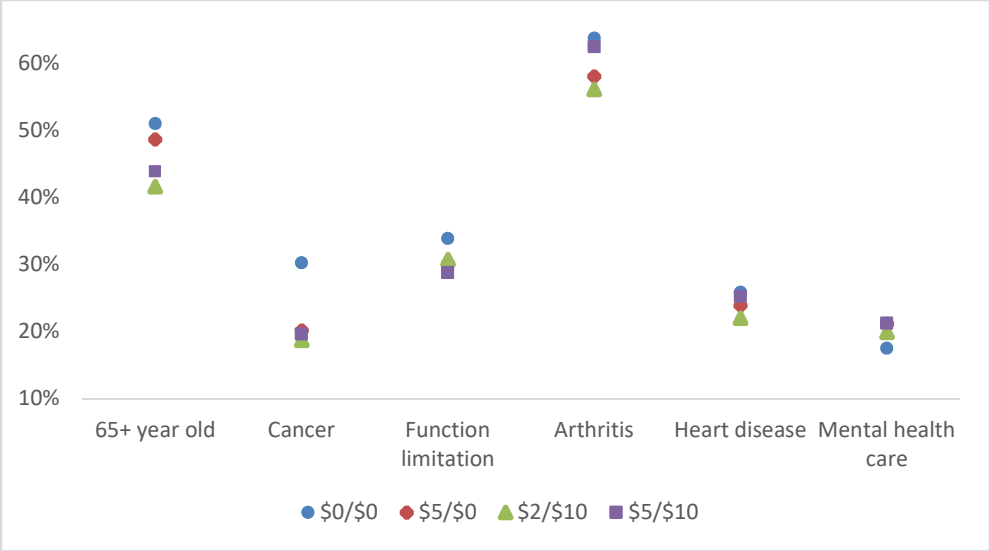

Figure 4: Characteristics of Responding Households by Incentive Condition

We also examined the composition of the responding sample using the appended variables from our ABS frame vendor, Marketing Systems Group. The appended variables came from many sources with varying levels of completeness (Roth, Caporaso, and DeMatteis, 2022; Roth, Han, and Montaquila, 2013). Although the accuracy rate might be low for some appended variables, we think that the status of missingness/non-missingness might be informative. The fact that the appended information was unavailable for some addresses probably meant that these households were harder to reach for various reasons. Figure 5 shows the base-weighted missing rates for some appended variables, including age of the head of the household (HoH), education of the HoH, tenure, marital status, and number of adults in household. The $0/$0 condition was associated with nominally lower missing rates than the non-zero incentive conditions, particularly for education, tenure, and marital status. There is limited evidence that the monetary incentives boosted response from the harder-to-reach households.

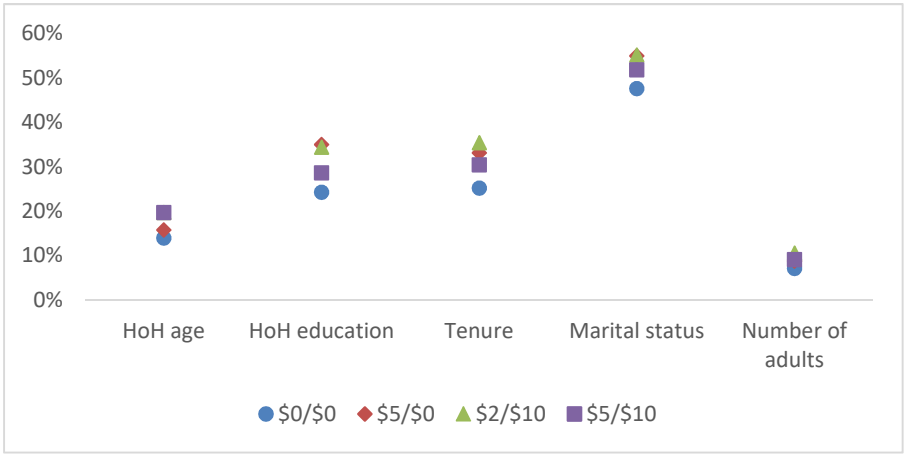

Figure 5: Proportion of Item Missingness for Vendor-Appended Frame Variables by Incentive Condition

4. Conclusions and Discussions

Our research established the significant treatment effects of some prepaid and promised incentive conditions for mail push-to-web data collection. For example, without any promised incentive, prepaid $5 improved the response rate by 10.2 percentage points from prepaid $0. Conditioning on promised $10, prepaid $5 boosted the response rate by 5.3 percentage points compared to prepaid $2. Conditioning on prepaid $5, the response rate was 7.0 percentage points higher for promised $10 than for promised $0. In addition, our cost analysis showed that by carefully choosing incentive levels and combinations, survey practitioners could boost response rates while containing or even lowering the average cost per complete. Finally, there was limited evidence that monetary incentives helped bring in the harder-to-reach population, including younger households.

We also have some thoughts about future research. First, it would be worth testing the effectiveness of a lower-level promised incentive (e.g., $5). Second, budget permitted, we would like to conduct an experiment with full factorial design to evaluate the interaction effect and tradeoff between prepaid and promised incentives. Lastly, most national surveys are more than 10-minute long, and it is unclear whether our findings would hold for a longer questionnaire. However, one application of a short questionnaire is to screen for rare subpopulations in a two-phase data collection. In the two-phase setting, it would be interesting to examine how the screener incentive condition might affect the main interview response rate and sample composition.

References

Church, A. H. 1993. Estimating the effect of incentives on mail survey response rates: A meta-analysis. *Public Opinion Quarterly*, 57, 62-79.

Debell, M., Maisel, N., Edwards, B., Amsbary, M., and Meldener, V. 2019. Improving survey response rates with visible money. *Journal of Survey Statistics and Methodology*, 8(5), 821–831. doi: 10.1093/jssam/smz038.

Edwards P., Cooper, R., Roberts, I., & Frost, C. 2005. Meta-analysis of randomized trials of monetary incentives and response to mailed questionnaires. *Journal of Epidemiology and Community Health*, 59, 987–999.

Groves, R. M., Singer, E., and Corning, A. 2000. Leverage-saliency theory of survey participation. *Public Opinion Quarterly*, 64, 299-308.

Lesser, V. M., Dillman, D. A., Carlson, J., Lorenz, F., Mason, R., and Willits, F. 2001. *Quantifying the influence of incentives on mail survey response rates and nonresponse bias.* Paper presented at the annual meeting of the American Statistical Association, Atlanta, GA.

Mercer, A., Caporaso, A., Cantor, D., & Townsend, R. 2015. How much gets you how much? Monetary incentives and response rates in household surveys. *Public Opinion Quarterly*, 79, 105–129.

Roth, S., Caporaso, A., and DeMatteis, J. 2022. Variables appended to ABS frames: Has their data quality improved? *PLoS ONE*, 17(11): e0269110. doi: 10.1371/journal.pone.0269110

Roth, S. B., Han, D., and Montaquila, J. 2013. The ABS frame: Quality and considerations. *Survey Practice*, 6(4).

Singer, E. and Ye, C. 2013. The use and effects of incentives in surveys. *The Annals of the American Academy of Political and Social Science*, 645, 112–141.