Global Poverty Research Lab Research Methods Notes

Messaging to Improve Response Rates: Effectiveness of Pre-Survey SMS Messages

Andrew Dillon
Global Poverty Research Lab, Northwestern University Kellogg School of Management

Steven Glazerman
Innovations for Poverty Action

Michael Rosenbaum
Innovations for Poverty Action

Working paper no. 21-106
April 2021

Keywords: research methods, phone surveys, response rates

This note is part of the Research Methods Initiative, which is a collaboration between Northwestern University's Global Poverty Research Lab (GPRL) and Innovations for Poverty Action (IPA). More information is available at GPRL's website and IPA's website.

Northwestern
GLOBAL POVERTY RESEARCH LAB

ipa
INNOVATIONS FOR POVERTY ACTION

Northwestern
GLOBAL POVERTY
RESEARCH LAB

ipa
INNOVATIONS FOR POVERTY ACTION

Messaging to Improve Response Rates

Effectiveness of Pre-Survey SMS Messages

Response rates remain low in phone surveys compared to face-to-face data collection (see here for a similar brief on response rates). This is especially true for random digit dial (RDD) or similar "cold call" phone surveys, which are necessary in the absence of a sample frame of reliable phone numbers. This brief presents early evidence from a series of experiments IPA conducted in 4 countries during 2020 to learn whether pre-survey messages, typically SMS texts, improve the rates at which respondents answer the phone, and complete the interview.

We find that, on average, SMS messages improve the rate at which respondents complete the survey relative to no message. This change is not driven by the rate at which respondents answer the phone, but by survey completion conditional on starting the survey. This has the potential to increase the productivity of phone surveys.

We also test how random variation in message *content* affects the rate at which respondents answer and complete phone surveys. We find no statistically significant differences between types of messaging content, but these initial findings are only based on tests of closely related messages such as one emphasizing the importance of survey participation to "research" versus to "government." We also compare an appeal to monetary incentives to a reminder of when the survey would occur, again finding no statistically significant differences.

Finally, we examine whether messaging and message content affect the *composition* of the sample, not just response rates. We find that most sample characteristics are similar for the experimental groups. We test for differences by age, gender, education, and poverty status. Out of 42 tests for the combined sample, 3 are statistically significant. For example, the group assigned to receive an SMS reminding them of the incentive is on average 1.3 years younger (0.7 years SE) and on average 2 percentage points less likely to be predicted to be in poverty (0.9 pp SE) than the other groups in the experiment. There are no detectable differences in terms of gender, educational attainment, household size, employment status, or estimated probability of being in poverty for groups assigned to receive different types of message content.

Motivation

A vast literature exists in varied contexts, including many low- and middle-income countries (LMICs), on the size and format of incentives that appeal to extrinsic motivations to participate in surveys. Most evidence is based on self-administered surveys, but is consistent with enumerator-administered surveys.[1] The takeaway is consistent: monetary incentives are important, but there are rapidly diminishing marginal returns to the size of the incentive. There is far less evidence on how messaging and framing may motivate respondents to

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

This note is part of a series investigating survey implementation using computer-assisted telephone interviewing (CATI) and other remote survey modes by Northwestern University's Global Poverty Research Lab (GPRL) and Innovations for Poverty Action (IPA). It was prepared by Andrew Dillon, Steve Glazerman, and Michael Rosenbaum with helpful input from Navishti Das, Dean Karlan, Chris Udry, and Shana Warren. These methods notes are made possible with the generous support from GPRL. More information is available on IPA's website about phone survey methods and GPRL & IPA's Research Methods Initiative.

complete surveys. Messaging via SMS is both less expensive than monetary incentives and easier to implement.

Low response rates mean high costs of interviewer time spent on unsuccessful attempts. This also carries over to data quality, potentially resulting in unrepresentative samples, as respondent pools are made up of only the most available, compliant individuals and those with working phones at the time of the survey. Respondents who keep up with payments, have access to electricity, and have access to reliable signal may be unrepresentative of the target population. Investigators can potentially increase the likelihood that respondents participate by informing them of the survey or the beneficial interventions that a study may provide, either through advertising or direct appeals to individuals, such as pre-survey SMS contacts.

Literature

In LMICs, various research teams have tested pre-survey SMS contact, finding substantive and statistically significant increases in response rates.[2] However, costs to send bulk SMS at the time of the studies, the mid-2010s, was high enough that it was not cost effective to increase response rates.[3] Some of these results may be context-specific, insofar as timing of the SMS and timing of the subsequent attempt had substantively different effects on response rates.[4]

This research does not suggest what content the SMS should contain. Reviews of the evidence in higher-income countries define specific pathways which may incentivize respondents to complete surveys: (1) intrinsic motivations, (2) extrinsic motivation, and (3) survey characteristics.[5] These pathways provide insights into what incentives, monetary or not, encourage survey completion.

This research is based on leverage-salience theory, a formal theory on survey response that defines how individual characteristics of a survey interact with respondent-specific weighting of interests.[6] However, the majority of this work is based on self-administered surveys from widely-known institutions in OECD countries. There is less evidence on survey incentives in LMICs. One study suggests this theory may be externally valid based on comparisons of different formats of monetary incentives in Ghana, finding partial replications of how respondents value prepaid and promised incentives.[7]

Data and Methods

The analysis in this brief comes from four RDD surveys that IPA conducted between April and September 2020 in Colombia, Mexico, the Philippines, and Rwanda, resulting in 5,719 complete surveys from 30,744

[2] Leo, B., & Morello, R. (2016). *Practical Considerations with Using Mobile Phone Survey Incentives: Experiences in Ghana and Tanzania* (SSRN Scholarly Paper ID 2841010). Social Science Research Network; Amaya, A., Lau, C., Owusu-Amoah, Y., & Light, J. (2018). Evaluation of Gaining Cooperation Methods for IVR Surveys in Low- and Middle-income Countries. *Survey Methods: Insights from the Field (SMIF)*;; Kasy, M., & Sautmann, A. (2021). Adaptive Treatment Assignment in Experiments for Policy Choice. *Econometrica*, *89*(1), 113–132.

[3] Leo, B., & Morello, R. (2016). *Practical Considerations with Using Mobile Phone Survey Incentives: Experiences in Ghana and Tanzania* (SSRN Scholarly Paper ID 2841010). Social Science Research Network

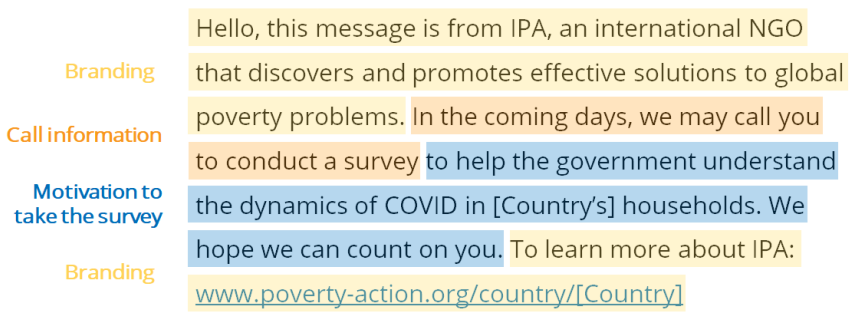
[4] Kasy, M., & Sautmann, A. (2021). Adaptive Treatment Assignment in Experiments for Policy Choice. *Econometrica*, *89*(1), 113–132. h

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

[6] Groves, R. M., Singer, E., & Corning, A. (2000). Leverage-Saliency Theory of Survey Participation. *Public Opinion Quarterly*, *64*(3), 299–308.

[7] Meuleman, B., Langer, A., & Blom, A. G. (2018). Can Incentive Effects in Web Surveys Be Generalized to Non-Western Countries? Conditional and Unconditional Cash Incentives in a Web Survey of Ghanaian University Students. *Social Science Computer Review*, *36*(2), 231–250.

attempted respondents. These data are representative of the cell-phone using population in each country, except in Mexico where the survey is limited only to Mexico City area codes.



In each of these sites, IPA randomly assigned respondents to receive a particular SMS message or no message one day before or on the day of the first call attempt. All SMS were sent in the language with the highest rates of literacy in the country (see Table 1). The SMS did not come from the same number from which interviewers would eventually place their calls.

These SMS contained multiple components: (1) notification of survey day, (2) branding on the survey firm, and (3) a reason to take the survey. An example format, translated from Spanish, is shown above.

Although we cannot isolate the effect of branding or call information, branding and informational components are necessary to ensure that the SMS messaging was viable and consistent with principles of informed consent required for protection of human subjects of research. Therefore, we only modified motivations to take the survey.

Table 1 displays the experimental variations in each site and the proportion of respondents that were sent an SMS. Survey branding and timing were adapted slightly to each country context based on feedback by project teams to ensure that messages were colloquial. These variations included two types of motivations to complete the survey: an intrinsic motivation, that the survey would help either "researchers" or the "government", and extrinsic motivation, that respondents would receive an incentive.

Table 1: Experimental conditions for SMS Content, with percentage of cases assigned to each

Country	Message Language	No SMS	Message content			
			Basic notification	Intrinsic motivation		Extrinsic motivation
				Research	Government	Incentives*
Colombia	Spanish	50%		25%	25%	
Mexico City	Spanish	50%		25%	25%	
Philippines	English	33%	33%			33%
Rwanda	Kinyarwanda	50%	50%			

Note: All SMS included information on IPA (branding) as well as information on which day calls for the survey would begin.
**Incentive text included survey information. This treatment arm can be interpreted as a comparison between including extra information on an incentive versus an SMS with just information on IPA and the survey.*

We measure productivity in two ways: the proportion of respondents who answered any call and the proportion of respondents who completed the survey over any call. We define outcomes as the best result across all attempts, where a respondent is coded as "answered and not complete" if they picked up once across all call attempts and verbally refused to take the survey. We also investigate if there are differences in sample composition on observable demographics such as age, gender identity, educational attainment, employment status, and predicted poverty probability measured through the poverty probability index (PPI) among completed surveys.

Not all SMS were successfully delivered. In our analysis, we estimate all effects as intent-to-treat, the effect of being assigned to receive an SMS, and don't account for SMS messages that were not delivered. The estimates of the treatment effects below should be interpreted as an underestimate of the true effect of receiving an SMS, as even respondents who did not receive an SMS are assumed to have received one.

The SMS marketing providers used by IPA had varying levels of success in delivering SMS. This can be due to a number of factors including whether the number is recognized by the network during the survey period or if the marketer is flagged by the network as sending spam. The delivery rates ranged from 19 to 73 percent of attempted SMS. Given this experience of SMS that failed to send, we recommend using providers like Twilio which allow for direct control of SMS delivery and tracking.

Findings

We find that sending a pre-survey SMS message *reduces* pickup rates by 1.0 percentage points (standard error of 0.6 percentage points, abbreviated hereafter as "pp SE") compared to an average of 55.7 percent pickup when not receiving an SMS. This is a statistically significant change at the $p < 0.10$ level. These effects are shown in Panel 1 of Table 2. This is driven by a larger reduction in the Philippines, where pickup rates are reduced by 3.7 percentage points (1.5 pp SE). In other sites the effect on pickup rates was not statistically significant. It is not clear what would explain this pattern, although receiving an SMS may lead respondents to screen their calls, and thus not answer. An experiment using pre-visit flyers during door-to-door solicitations of charitable giving found similar effects on answering the door.[8]

We find that sending a pre-survey SMS results in a statistically significant increase in survey completion of 1.1 percentage points (0.4 pp SE) across all sites. These effects are shown in Panel 2 of Table 2. This effect is driven by increases in three sites: 2.4 percentage points (1.1 pp SE) in Colombia, 1.2 percentage points (0.4 pp SE) in Mexico City, and 3.8 percentage points (1.8 pp SE) in Rwanda. In the Philippines, the negative effect on answering the phone translates into a statistically insignificant, but negative, effect on response rates.

Table 2: Effects of Receiving an SMS

	Colombia	Mexico City	Rwanda	Philippines	All sites
Panel 1: Answered					
Any SMS	-0.010	-0.008	0.009	-0.037***	-0.010*
No SMS	0.627	0.549	0.582	0.479	0.557
Panel 2: Complete					
Any SMS	0.024**	0.012***	0.038**	-0.014	0.011***
No SMS	0.238	0.057	0.422	0.186	0.172
SMS delivery rate	76.9%	36.8%	-	80.8%	63.3%
N	6018	13392	3339	7995	30744

*Note: The sample is restricted to numbers which could be dialed at any point. Each column presents results from a single OLS regression of the dependent variable described in the column heading on the variables with enumerator fixed effects and covariates. Covariates are included for day of week and time of day. Each group of columns reflects individual sites. The "All sites" column includes project fixed effects. SMS effects are intent-to-treat and do not account for SMS messages that were not delivered. SMS delivery rate for Rwanda was unavailable at the time of writing. Robust standard errors were used to calculate statistical significance. *** p<0.01, ** p<0.05, * p<0.10*

[8] DellaVigna, S., List, J. A., & Malmendier, U. (2012). Testing for altruism and social pressure in charitable giving. *The quarterly journal of economics*, *127*(1), 1-56.

Although, the results in Table 3 indicate a significant difference between SMS messages with various content variations and respondents who did not receive an SMS, tests of the differences between different content variations are all statistically insignificant. Similar to our findings from any SMS, variants in the content of the SMS result in reduced rates in answering the phone in the Philippines across all content variants and significant increases in completion for the majority of content variants in all other sites. This suggests that the content variations that we tested do not affect response behavior. However, these content variations are relatively subtle, sometimes changing only a single word.

Table 3: Effects of Receiving an SMS by Content Type

	Colombia	Mexico City	Rwanda	Philippines	All sites
Panel 1: Answered					
[1] SMS (Researchers)	-0.012	-0.002			-0.005
[2] SMS (Government)	-0.008	-0.013			-0.011
[3] SMS (Information only)			0.009	-0.044***	-0.015
[4] SMS (Incentive)				-0.028*	-0.010
p-value: [1] - [2]	0.837	0.327	-	-	0.544
p-value: [3] - [4]	-	-	-	0.243	0.668
No SMS	0.627	0.549	0.582	0.479	0.557
Panel 2: Complete					
[1] SMS (Researchers)	0.026*	0.014***			0.017***
[2] SMS (Government)	0.022	0.010*			0.014**
[3] SMS (Information only)			0.038**	-0.022*	0.001
[4] SMS (Incentive)				-0.006	0.003
p-value: [1] - [2]	0.827	0.557	-	-	0.561
p-value: [3] - [4]	-	-	-	0.126	0.814
No SMS	0.238	0.057	0.422	0.186	0.173
SMS delivery rate	76.9%	36.8%	-	80.8%	63.3%
N	6,018	13,392	3,339	7,995	30,744

*Note: The sample is restricted to numbers which could be dialed at any point. Each column presents results from a single OLS regression of the dependent variable described in the column heading on the variables with enumerator fixed effects and covariates for day of week and time of day. Each group of columns reflects individual sites. The "All sites" column includes project fixed effects. SMS effects are intent-to-treat and do not account for SMS messages that were not delivered. SMS delivery rate for Rwanda was unavailable at the time of writing. Robust standard errors were used to calculate statistical significance. *** p<0.01, ** p<0.05, * p<0.10*

Table 4 shows that the messaging strategies produced respondent samples that in most cases had similar characteristics, comparing self-reported age, gender, educational attainment (more versus less than secondary), employment status, or poverty probability. Of 24 tests comparing a specific messaging treatment arm to the no SMS condition, only 2 are statistically significant. For example, across all sites, respondents from the group assigned to receive a message with a reminder of the incentive are 1.3 years younger (0.7 years SE) and 2 percentage points lower (0.9 pp SE) on the poverty probability index, on average, than respondents in the group assigned to receive no message. Of 12 tests comparing different message strategies to each other, only the incentive message verus information-only message is statistically significant. The significant results are from the Philippines only. It might still be the result of chance due to the number of hypothesis tests performed.

Table 4. Sample Composition by Treatment Arm

	Age (years)	Female	Secondary Education	Household Size	Employed	Poverty Probability
Panel 1: Predicted Sample Mean by Treatment Arm						
No SMS	33.6	53%	70%	4.6	41%	21%
Any SMS	33.7	53%	71%	4.5	42%	20%**
[1] SMS (Researchers)	33.4	52%	69%	4.4	40%	20%
[2] SMS (Government)	34.5	55%	72%	4.5	39%	20%
[3] SMS (Information only)	33.8	52%	71%	4.7	44%	20%
[4] SMS (Incentive)	32.3**	56%	72%	4.5	41%	19%**
Panel 2: P-value for test:						
[1] - [2]	0.337	0.396	0.398	0.676	0.878	0.807
[3] - [4]	0.020**	0.172	0.763	0.286	0.421	0.353
N	4,762	5,218	5,210	5,211	5,220	4,890

*Note: The sample is restricted to complete surveys. Each column presents results from a single OLS regression of the dependent variable described in the column heading on the SMS treatment variables with enumerator and project fixed effects and covariates for day of week and time of day. SMS effects are intent-to-treat and do not account for SMS messages that were not delivered. Robust standard errors were used to calculate statistical significance. Poverty probability is the predicted probability from the PPI, estimating that the respondent is below each country's national poverty line. Employed indicates that the respondent worked for one or more hours in the 7 days prior to the survey. Tests in the upper panel compare each SMS treatment group to "No SMS." *** p<0.01, ** p<0.05, * p<0.10*

Implications

The marginal cost of sending SMS messages is typically low, less than one cent per message in some countries. The evidence in this brief argues for building such messages into survey plans. Despite no statistically significant changes in the rate at which respondents answer the phone, completion rates are significantly higher due to receiving an SMS. There are other considerations, like whether sample members are likely to be literate in the dominant language or languages. Messages can be sent in many languages, but with RDD in multilingual societies, there is no information about the respondent to predict what language a message should be written in. Another consideration is whether the survey is part of an RCT where there is an intervention being delivered via text message. Using the same platform for intervention and evaluation of the intervention could risk confusing study participants or contributing to messaging overload, which could weaken the intervention itself. This would be an undesirable by-product of SMS.

Given that sending a message is recommended as a default and there are a limited number of characters available for such a message, there may be gains based on messaging content. Based on our tests of making promised incentives salient and encouraging civic responsibility to respond, we do not find significant differences in message content on response rates, or meaningful patterns in sample composition changes. This evidence suggests that small contrasts in message content are insignificant, not that the message content in itself is not important.

Further research should expand on identifying which content is successful to encourage survey participation and for which types of respondents. Although our experimentation finds promising results of sending pre-survey SMS to RDD samples, specific samples may respond to different information. We are investigating this question with variations in messaging content to encourage panel retention among existing respondents to RDD surveys.