Global Poverty Research Lab Research Methods Notes

Assessing Repeated and Rescheduled Attempts in Random Digit Dial Survey

Biljana Bogicevic
Innovations for Poverty Action

Navishti Das
Global Poverty Research Lab, Northwestern University Kellogg School of Management

Emma Davies
Global Poverty Research Lab, Northwestern University Kellogg School of Management

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

Steve Glazerman
Innovations for Poverty Action

Michael Rosenbaum
Innovations for Poverty Action

Working paper no. 21-110
September 2021

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

Assessing Repeated and Rescheduled Attempts in Random Digit Dial Surveys

A central challenge to telephone surveys is low response rates. Nonresponse reduces survey efficiency (cost per complete) and could introduce bias if nonrespondents differ from respondents in ways that affect the parameters we seek to measure. This is particularly true with random digit dial (RDD) surveys, where the response rate can be especially low.[1] For researchers designing RDD survey protocols, there is a clear tradeoff between achieving a higher response rate, by calling fewer numbers repeatedly, and achieving a sample size target with less effort, by calling more numbers less intensively.

This brief explores this tradeoff by measuring the effects of (i) repeated attempts per case, and (ii) rescheduling a call, on completion rates and sample composition. In doing so, we generate evidence on whether repeated attempts and rescheduled appointments are effective in increasing response rates. Effectiveness of these two protocols is measured via: survey contact and completion rates. We also measure the effects on sample composition using a set of observed characteristics of respondents.

Using data from nine low- and middle-income countries (LMICs), we find that repeated and rescheduled attempts result in lower completion rates than new attempts, on average. However, the respondents who complete the survey in later attempts or after rescheduling have statistically significant differences in observable characteristics. This suggests that more call attempts may be needed to adequately represent the respondents who are harder to interview, even if those call attempts produce fewer completions per case.

Literature

There is no clear consensus on the optimal number of call attempts to balance costs and response rates in the survey methods literature. Research from higher-income countries (HIC) finds the optimal call attempt cutoff ranges from four to as many as twelve or fourteen calls.[2] Defining optimal call protocols may not be possible, as they may depend on respondent characteristics. Data on these characteristics in similar samples may not exist.[3]

[1] Dillon, A., Glazerman, S., & Rosenbaum, M. (2021). Understanding Response Rates in Random Digit Dial Surveys. Global Poverty Research Lab Working Paper No. 21-105

[2] Stec, J., P. Lavrakas, & C. Shuttles. (2004). Gaining efficiencies in scheduling call-backs in large RDD national surveys. In Proceedings of the Survey Research Methods Section of the American Statistical Association, 4430–37.; Srinath, K. et. al. (2001). Balancing cost and mean squared error in RDD telephone survey: the national immunization survey. In Proceedings of the Annual Meeting of the American Statistical Association.; Harpuder, B. & Stec, J. A. (1999). Achieving an optimum number of callback attempts: cost-savings vs. non-response error due to non-contacts in RDD surveys. Proceedings of the Survey Research Methods Section, American Statistical Association, 913–918.

[3] Hansen, S. E. (2007). CATI Sample Management Systems. In Advances in Telephone Survey Methodology (pp. 340-358). Hoboken, NJ,

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 Biljana Bogicevic, Navishti Das, Emma Davies, Andrew Dillon, Steve Glazerman, and Michael Rosenbaum with helpful input from 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.

When determining the optimal call attempt cutoff, researchers should not only consider costs but also the implications for data quality and bias. There isn't clear evidence on whether additional call attempts reduce coverage bias. Some research suggests that respondents who answer after more attempts have different observable characteristics, such as age, gender, educational attainment, among others.[4] Other evidence suggests that this is not the case.[5]

There is scarce literature on the optimal number of call attempts in LMICs, in settings with fixed size sampling frames. A recent study in Turkey reports that increasing the number of calls from three to over fifteen led to a statistically significant increase in response rate compared to a previous follow-up.[6] On average, it took ten calls to successfully interview 70 percent of non-beneficiaries and four calls to successfully interview 70 percent of beneficiaries of a conditional cash transfer program.

The efficacy of rescheduling appointments has been studied far less than optimal contact rates. Most literature does not differentiate between additional calls and appointments, although having some appointment and scheduling protocols are considered standard. One recent paper notes that scheduled callbacks have higher completion rates.[7] Some strategies for scheduling include scheduling a call before the survey begins or requesting call schedules from respondents if they cannot complete the survey when first contacted.[8] This brief investigates the latter form of scheduling.

Data and Methods

IPA conducted nine RDD surveys between April and September 2020, resulting in 12,145 complete surveys from 64,635 attempted phone numbers. Surveys were conducted in Africa (Burkina Faso, Ghana, Rwanda, Sierra Leone, Uganda, and Zambia), Asia (Philippines), and Latin America (Colombia and Mexico).

Many elements of these surveys are common across countries.[9] All but one were variants of the same survey about the impacts of the COVID-19 pandemic on households. All surveys dialed numbers from lists of verified active numbers provided by Sample Solutions, surveyed the person who responded to the call, and required respondents to be over the age of 18.

---

[4] Legleye, S et. al. (2013). Improving Survey Participation: Cost Effectiveness of Callbacks to Refusals and Increased Call Attempts in a National Telephone Survey in France, *Public Opinion Quarterly*, Volume 77, Issue 3, 2013, Pages 666–695.; Vicente, P. & Marques, C. (2017). Do Initial Respondents Differ From Callback Respondents? Lessons From a Mobile CATI Survey. *Social Science Computer Review*, 35(5), 606-618.

[5] Montaquila, J. et. al. (2007). Aspects of Nonresponse Bias in RDD Telephone Surveys. In Advances in Telephone Survey Methodology (pp. 561-586). Hoboken, NJ, USA: John Wiley & Sons.

[6] Özler, B., & Cuevas, P. (2019). Reducing attrition in phone surveys. World Bank Development Impact Blog. Retrieved from: https://blogs.worldbank.org/impactevaluations/reducing-attrition-phone-surveys

[7] AAPOR Cell Phone Task Force (2010). New considerations for survey researchers when planning and conducting RDD telephone surveys in the US with respondents reached via cell phone numbers. *American Association for Public Opinion Research.*; Shino, E., & McCarty, C. (2020). Telephone Survey Calling Patterns, Productivity, Survey Responses, and Their Effect on Measuring Public Opinion. *Field Methods,* 32(3), 291-308.

[8] Amankwah, A. et. al.(2020). High Frequency Mobile Phone Surveys of Households to Assess the Impacts of COVID-19 (Vol. 3) : Guidelines on CATI Implementation. Washington, D.C.: World Bank Group; Kasy, M., & Sautmann, A. (2021). Adaptive treatment assignment in experiments for policy choice. *Econometrica*, *89*(1), 113-132.; Dabalen, A. et. al.(2016). Mobile Phone Panel Surveys in Developing Countries: A Practical Guide for Microdata Collection. Directions in Development--Poverty. Washington, D.C.: World Bank.

[9] The Uganda survey required respondents to have used mobile money and quota sampling, where respondents were sampled until a minimum number of respondents with combinations of educational attainment and geographical area had completed surveys.

Each survey used different protocols for the maximum number of attempts, spacing of attempts, and call rescheduling. The maximum attempts ranged from 2 in Ghana and Rwanda to 10 in Uganda. Calls in most countries started at 8 AM and ended between 5-9 PM, based on the country. Repeated call attempts were typically spaced 3-4 hours apart.

We examine the effect of repeated call attempts on answered calls, completed calls, and sample composition on a set of common demographic characteristics: age, gender identity, educational attainment, employment status, household size, and predicted probability of being below the national poverty line.

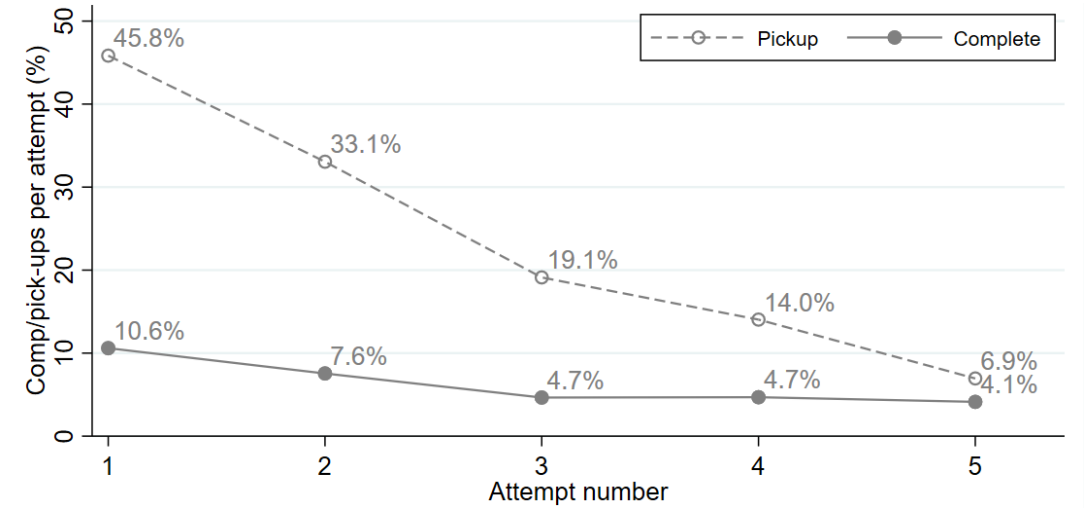
We analyze the rescheduling data similarly: estimating differences in the number of completed surveys between potential respondents that reschedule in any attempt and those who do not. We also estimate differences in compositions between the rescheduled and non-rescheduled completions. This comparison of respondent characteristics presents the sample that corresponds to the survey design choice where rescheduling is allowed to the most likely sample without rescheduling.

Findings

Number of attempts

When averaging across all nine countries, we find that the first call attempt results in the highest pick-up and completion rates, with decreasing pickup and completion rates for later call attempts. While subsequent attempts result in a steeper fall in pickup rates, completion rates fall off at a slower pace. This suggests that some respondents are screening out survey calls, but complete the survey when they eventually pick up. Figure 1 displays the pick-up and completion rates at different call attempts.

Figure 1: Marginal Completion and Pick-up Rates by Attempt



The most efficient protocol may be to attempt individual cases once because completion rates tend to be highest on the first dial. However, this protocol might result in a different sample of completed surveys. To test whether this is the case, we test whether respondents' characteristics vary between the first attempt and all subsequent attempts.

Across all countries, we find statistically significant differences jointly across all sample characteristics. This is driven by differences in employment status and predicted poverty level. Respondents who completed the

survey at a later attempt (2+) were 5.8 percentage points more likely to be employed (1.2 pp SE) and 1.0 percentage point less likely to be predicted to be below the national poverty line (0.5 pp SE).

Average sample characteristics of respondents who complete the survey on first and later attempts also differ in individual countries. Trends in employment and predicted poverty probability remain consistent across most countries. However, statistically significant differences in educational attainment are less consistent. In Ghana and Uganda fewer respondents of later attempts have attained secondary education by 6.2 percentage points (1.6 pp SE) and 7 percentage points (3.5 pp SE), respectively. Meanwhile, respondents of later attempts in Rwanda are 10.1 percentage points (4.3 pp SE) more likely to report having attained secondary education.

Table 2: Differences in Respondent Characteristics by Attempt Number

Country	Age	Female	Secondary education	Employed	Household size	Poverty probability	Joint equivalence of differences	N
Panel 1: Attempt 1, Average Values								
All sites	**32.7**	**0.45**	**0.71**	**0.41**	**5.2**	**0.20**	**-**	**9231**
Rwanda	30.4	0.36	0.59	0.43	4.9	0.16	-	1344
Sierra Leone	32.8	0.35	0.60	0.45	6.1	0.24	-	1117
Philippines	32.0	0.70	0.86	0.39	4.8	0.11	-	1190
Ghana	31.9	0.38	0.92	0.40	5.3	0.11	-	1122
Mexico City	37.0	0.54	0.68	0.50	4.1	0.36	-	741
Colombia	38.3	0.64	0.72	0.32	4.1	0.27	-	1187
Uganda	30.4	0.36	0.50	n.a.	n.a.	n.a.	-	395
Zambia	32.0	0.44	0.81	0.42	5.5	0.13	-	825
Panel 2: Difference Between Attempts 2+ and Attempt 1								
All sites	**-0.1**	**0.00**	**0.00**	**0.06*** **	**0.0**	**-0.01***	**0.000*** **	**2907**
Rwanda	-0.2	0.04	0.10**	-0.04	0.2	-0.03*	0.102	145
Sierra Leone	-0.7	0.01	0.04	0.09**	0.2	-0.04**	0.013**	167
Philippines	-0.9	-0.07*	0.04	0.07*	0.1	0.01	0.019**	199
Ghana	-0.5	0.01	-0.06***	0.09***	0.0	0.01	0.000***	515
Mexico City	1.3*	-0.02	0.03	0.06**	-0.1	-0.03**	0.077**	596
Colombia	-0.2	-0.03	0.05	0.07**	0.1	-0.03**	0.001***	317
Uganda	0.1	0.05	-0.07**	n.a.	n.a.	n.a.	0.180	436
Zambia	-0.4	-0.02	0.00	0.04	-0.3*	0.01	0.447	471

*Note: The sample is restricted to complete surveys. Each row presents results from a single OLS regression of sample characteristics on an indicator that the case was attempted more than once. Sites where less than 5 percent of completions were on later attempts are omitted (Burkina Faso). The all sites row includes project fixed effects and data from all countries. Robust standard errors were used to calculate statistical significance. The joint equivalence column reports the p-value of F-test of joint significance of all explanatory variables. 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. *** p<0.01, ** p<0.05, * p<0.10*

Rescheduling

The percentage of cases that reschedule over any call varies by country, from 1.6 percent of the sample of all attempted calls in Rwanda to 17.5 percent of the sample in Zambia. Surveying a rescheduled case decreases survey efficiency, measured through completion rates, compared to surveying a new case by 5.9 percentage points (0.5 pp SE). These findings are presented in Table 3. Seven out of nine countries show a decrease in completion rates for rescheduled cases, with large and statistically significant decreases in six countries, ranging from 23.0 percentage points (7.5 pp SE) in Burkina Faso to 9.8 percentage points in the Philippines (1.2 pp SE). Burkina Faso, Rwanda, and Sierra Leone show a slightly higher drop in efficiency with rescheduling

but all with a very low demand for rescheduling. For countries with a larger proportion of rescheduled cases - Colombia, Ghana, the Philippines - the reduction in completion rate is smaller, but still remains substantive and statistically significant.

Rescheduling may improve response rates and data quality in certain situations when surveys have (1) low response rates and/or (2) strict eligibility criteria. If the overall response rate and, thus, the response rate of new cases is low, working a new case is less productive. The stricter the eligibility criteria are, the more effective rescheduling might be if cases are rescheduled after cases are screened for eligibility. This can be shown in Mexico City, where rescheduling increases completion rate by a statistically significant margin. In Mexico City, the response rate was low (6.9% of attempted cases or 13.8% of contacted cases) and rescheduling almost doubled the completion rate (increase of 5.3 pp in attempted cases, 0.4 pp SE). Rescheduling also has a positive effect on the completion rate in Uganda, where the eligibility criteria were stringent, but this effect is not statistically significant.

Table 3: Rescheduling Success

	Respondents			Conditional completion rate			Total completion rate
	Attempted	Rescheduled		(1)	(2)	(1) - (2)	
	N	N	%	Rescheduled	Not rescheduled	Rescheduled - Not rescheduled	
All sites	**64635**	**7867**	**12.2%**	**13.6%**	**19.5%**	**-0.059*****	**18.8%**
Rwanda	3862	62	1.6%	19.4%	38.9%	-0.195**	38.6%
Burkina Faso	2328	44	1.9%	36.4%	59.3%	-0.230**	58.9%
Sierra Leone	3410	271	8.0%	21.0%	39.1%	-0.181***	37.7%
Philippines	8378	1069	12.8%	8.0%	17.8%	-0.098***	16.6%
Ghana	7806	1004	12.9%	11.9%	22.3%	-0.105***	21.0%
Mexico City	21391	2792	13.1%	10.9%	5.6%	0.053***	6.3%
Colombia	6184	889	14.4%	15.0%	25.9%	-0.109***	24.3%
Uganda	8024	1167	14.5%	11.5%	10.2%	0.013	10.4%
Zambia	3252	569	17.5%	37.1%	40.4%	-0.034	39.9%

Note: Sample is all respondents attempted in each site including non-working numbers. Respondents are reported as rescheduled if they request to reschedule across any attempt. The difference column is a t-test of completion rates between observations that rescheduled any calls and those that did not reschedule any call. A positive number means that attempting a respondent who rescheduled a call was more productive than a respondent that did not choose to reschedule a call.

The overall effect of rescheduling on survey completions is limited in magnitude due to the small number of cases that request rescheduling in most countries. Closing a case may be more valuable unless rescheduling helps increase the representativeness of the sample. As with increasing the number of attempts, we estimate whether respondents' characteristics vary between the rescheduled and new cases.

We find statistically significant differences across the countries in five of six measured characteristics as well as a statistically significant joint test of equivalence across all characteristics (Table 4). Respondents who reschedule are 0.69 years (0.37 years SE) older, 6.4 percentage points (1.6 pp SE) less likely to be female, 4.7 percentage points (1.4 pp SE) more likely to have attained secondary education, 9.5 percentage points more likely to be employed (1.7 pp SE), and are predicted to be 3.8 percentage points (0.7 pp SE) less likely to be below the national poverty line.

There is some variation between countries, although most countries follow a similar pattern for age, gender identity, employment status, and predicted poverty status. Rescheduling results in a statistically significant increase in male respondents in Columbia, Ghana, and the Philippines, a statistically significant increase in

employed respondents in Columbia, Ghana, Zambia, and Mexico City, and a statistically significant decrease in poor respondents in Columbia, Zambia, and Mexico City. Other characteristics, age and educational attainment, do not show a consistent pattern. In Mexico City, respondents who request to reschedule are older by 1.7 years (0.9 years SE) with smaller households by 0.4 members (0.1 member SE) on average. In Uganda, rescheduling yields an increase of 10 percentage points (5 pp SE) in respondents with secondary education in the sample of completed surveys.

Table 4: Differences in Sample Characteristics by Rescheduling Status

Country	Age	Female	Secondary education	Employed	Household size	Poverty probability	Joint equivalence	N
Panel 1: Not Rescheduled, Average Values								
All sites	**32.7**	**0.46**	**0.71**	**0.42**	**5.2**	**0.20**	**-**	**11045**
Philippines	31.9	0.70	0.87	0.39	4.8	0.11	-	1303
Ghana	31.7	0.39	0.90	0.41	5.3	0.12	-	1514
Mexico City	37.2	0.53	0.68	0.51	4.2	0.36	-	1012
Colombia	38.2	0.64	0.72	0.32	4.1	0.28	-	1371
Uganda	30.4	0.40	0.45				-	687
Zambia	31.7	0.44	0.81	0.42	5.4	0.13	-	1091
Panel 2: Difference Between Rescheduled and Not Rescheduled Attempts								
All sites	**0.7***	**-0.06*****	**0.05*****	**0.10*****	**-0.2**	**-0.04*****	**0.000*****	**1130**
Philippines	-1.4	-0.12**	0.05	0.07	0.1	-0.02	0.017**	86
Ghana	0.0	-0.14***	0.03	0.12***	-0.3	-0.01	0.000***	130
Mexico City	1.7*	-0.02	0.04	0.06**	-0.4***	-0.05***	0.087*	333
Colombia	1.3	-0.08*	0.06	0.15***	-0.0	-0.07***	0.001***	133
Uganda	-0.1	-0.07	0.10**				0.052*	144
Zambia	1.0	-0.05	0.01	0.08**	-0.1	-0.03*	0.308	219

*Note: The sample is restricted to complete surveys. Each row presents results from a single OLS regression of sample characteristics on an indicator that the case was rescheduled. Sites where less than 5 percent of completions were rescheduled are omitted (Burkina Faso, Rwanda, and Sierra Leone). The all sites rows include project fixed effects and data from all countries. 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. *** p<0.01, ** p<0.05, * p<0.10*

Implications

This brief provides additional insights on two interview protocol choices: the maximum number of call attempts and call rescheduling. These findings have practical implications for defining future survey protocols. Beyond implications for maximizing the sample size within budget constraints, these results suggest a tradeoff between sample composition and more intensive call protocols. We find decreasing efficiency in both rescheduling and additional attempts on survey completion, but meaningful differences in sample composition among respondents who complete surveys in later attempts or after rescheduling.

Moving to a new case after each attempt is the most-cost-effective decision for increasing response rates. The highest success in completing surveys is achieved at the first call attempt. We find that later attempts result in a sample that is on average more educated and less likely to be below the national poverty line. It is not clear how many attempts future researchers should specify by default. A project can decide to increase call attempts based on response rates in piloting as well as compositional differences by attempt tracking during data collection.

Similarly, rescheduling decreases survey efficiency but results in completions for substantively different respondents: more male, more likely to have attained secondary education, more likely to be employed, and less likely to be poor. Rescheduling survey protocols achieve representation of populations that may be harder to interview with just a single attempt.

We recommend that survey protocols allow rescheduling by request as a default to convert soft refusals. Although respondents who request rescheduling are less likely to complete the survey than respondents who do not reschedule, they differ from respondents who do not reschedule on observable characteristics - age, gender, education, employment status, and predicted poverty level.

When researchers are uncertain about survey protocols, one practical recommendation would be to adjust repeated attempts and rescheduling protocols based on pilot data. In contexts with low response rates or high eligibility criteria, reaching a new eligible and cooperative respondent may require more interviewer time than reaching a respondent who requested rescheduling. As both survey protocol choices lead to differences in sample composition relative to completed initial interviews, an important area of future research is studying representativity of repeated attempts and rescheduling.