

Insights into the Utility and Management of Mobile Phone Panel Surveys: Evidence from Surveys of Household Resilience in Myanmar

Lindsey Jones ^Φ

Johannes von Engelhardt [§]

September 2020

Abstract

Interest in the use of mobile phone surveys has skyrocketed in recent years. The need for quick, reliable and low-cost information in the wake of the COVID-19 pandemic has only added to the momentum. Phone surveys offer numerous advantages when considered alongside traditional face-to-face surveys, including: remote access to respondents; near-real-time data collection; and lower administration costs. Many of these benefits are especially relevant to evaluators seeking to track resilience in post-disaster environments – contexts where it is often unsafe or difficult to access communities of interest. While rapid proliferation of phone networks and use in developing countries has opened new opportunities for researchers and development practitioners, the strengths and limitations of mobile phone surveys are still poorly understood.

In this paper we provide insights into the utility and management of mobile phone surveys using quantitative and qualitative results from a unique high-frequency mobile phone panel survey on household resilience in Myanmar. In particular, we draw on a programme of work launched by Building Resilience and Adaptation to Climate Extremes and Disasters (BRACED), an initiative that collected seventeen waves of survey data across two sites in eastern Myanmar between June 2017 and March 2019. By comparing phone surveys with a parallel face-to-face survey we reveal modest mode effects in perceived levels of resilience. For the most part, however, mode effects do not differ significantly across key socio-economic groups. The same is also true for non-response rates across panel waves – with the exception of household size. We also highlight a series of lessons learned in promoting low attrition rates in high-frequency phone surveys, including discussion of intrinsic and extrinsic motivations drawing on key informant interviews with enumerators and respondents. Finally, we offer suggestions and tips on how to effectively manage a mobile phone panel survey drawing on the experiences of the BRACED programme – ranging from call centre set-up to quality control and enumerator.

^Φ *London School of Economics and World Bank*; [§] *University of Amsterdam*

1. INTRODUCTION

Face-to-face interviews and surveys have long been the mainstay of social science research: the primary means of collecting quantitative socio-economic information on household development outcomes. While their predominance is unlikely to change in the near future, mobile phone surveys are increasingly touted as offering solutions to many of the drawbacks of traditional surveying (Dabalen et al. 2016). The trend has recently been catalysed by a surge in mobile surveying in response to the COVID-19 pandemic (World Bank 2020).

The advantages of mobile phone surveys are manifold. For a start, they allow information to be collected in a timely manner. Surveys can be designed such that respondents can answer at a time of their convenience, with information feeding back to evaluators in near-real-time. The remote nature of mobile surveying also bypasses the immense logistical and resource costs of sending large survey teams into the field. These factors are particularly relevant when collecting high-frequency panel data, or in post-disaster contexts where it may be too dangerous or difficult to access respondents. Most importantly, by taking advantage of the rapid proliferation of phone networks across the globe, and availability of cheap headsets, mobile phone surveys can be carried out in developed and developing countries alike. Indeed, as 96% of the world population now lives within range of a mobile-cellular network signal (ITU 2018), mobile surveys offer considerable promise in complementing, or in some cases replacing, face-to-face surveys.

Despite their immense potential, mobile phones surveys are not a silver bullet. They come in many different formats, each with their own strengths and weaknesses: from text-based formats such as short message service (SMS) and unstructured supplementary service data (USSD) services, to voice-based alternatives such as interactive voice response (IVR) and live interviews using computer-assisted telephone interviewing (CATI). As such, while SMS surveys are cheap and can be answered at the respondent's convenience, they are of limited use in areas with high levels of illiteracy. Live interviews using CATI can somewhat overcome this drawback but are comparatively expensive.

Perhaps the main difficulty in running a robust mobile survey is generating a representative sample. While most of the world has seen a staggering growth in mobile phone adoption rates, phone access in low and lower-middle income countries is still neither universal nor equally distributed across regions and social strata. Indeed, while adoption processes differ across countries, a number of universal patterns can be observed: those without access to a mobile phone tend to be more likely to live in rural communities, to be older, to be less wealthy, and to be female (Pew, 2015; Leo et al., 2015). Unlike face-to-face surveys, where information on household listings is often publicly available, no central list of phone ownership exists. Phone networks that may have such information are often unwilling to share it for obvious privacy and commercial reasons. These drawbacks often mean that evaluators wishing to run a panel survey opt for a

traditional face-to-face survey and sampling strategy, before then collecting mobile phone numbers (or handing out phones) in order to ensure representative samples (see Dabalen, et al., 2016).

Another drawback of mobile phone surveys is the need for questionnaires to be short and questions to be simple. Even a skilled interviewer will have difficulties keeping the attention and goodwill of a respondent for much longer than 20-30 minutes during a phone interview. While in face-to-face interviews, a smile, friendly gesture or handshake can help to establish rapport and put the respondent at ease, building a trust relationship with the respondent can be much more difficult over the phone. This is particularly true if the phone interviews have not been preceded by a face-to-face encounter with an interviewer as part of a baseline.

Lastly, there are concerns that responses to surveys delivered over the phone may differ when compared to face-to-face administration. Known as ‘mode effects’, these can affect data quality, especially if data across different modes of administration are combined. As we will see in the Sections that follow, the issue of data validity in phone surveys often relates to social desirability bias and priming.

Despite a growing body of literature on the uses and challenges of phone surveys in developing countries, it is clear that a number of methodological and practical questions remain. Given the recent surge in interest regard mobile phone surveying in developing country contexts, largely spurred by the COVID-19 pandemic, understanding their strengths and limitations is key to improving the effectiveness of remote data collection for development actors. It is in this context that this paper aims to add value. Specifically, we draw on insights from a high-frequency mobile phone panel survey on household resilience in Myanmar. The Rapid Response Research project (RRR), run by the Building Resilience and Adaptation to Climate Extremes and Disasters (BRACED) programme, collected 17 waves of survey data across two sites in eastern Myanmar between June 2017 and March 2019. It provides a wealth of information and learning that builds on and extends the emerging body of literature on mobile panel surveys.

In particular, we use insights from the RRR survey to understand two important aspects of mobile phone panel data collection and management. Namely: i) do responses differ between face-to-face and phone modes of survey administration? and ii) do rates of attrition evolve over the course of a mobile panel survey, and how can respondents be encouraged to remain actively engaged?

To answer these questions, we combine quantitative analysis of RRR survey responses (and associated meta-data) with key informant interviews of survey respondents, enumerators and administrators of the RRR survey. Finally, we summarise lessons learned from delivery of the RRR survey, highlighting important gaps in knowledge and providing insight into the strengths and limitations of delivering mobile phone panels surveys in developing countries

2. BACKGROUND AND METHODOLOGY

Below we describe details of the BRACED programme setup, steps take in conducting the mobile phone surveys and survey design methods used to explore the research questions in detail.

a. The BRACED panel surveys

To gain insight into differences between face-to-face and mobile phone modes of survey administration we make use of a set of mixed-mode panel surveys conducted in Eastern Myanmar between 2017-2018. The surveys were designed to track levels of self-assessed household resilience over an 18-month period. Surveys took place as two separate exercises conducted in adjacent townships – one in Hpa-An, the other in Mudon 60 km South. Access to the sites was facilitated through the Building Resilience and Adaptation to Climate Extremes and Disasters (BRACED) programme in Myanmar. The BRACED Myanmar Alliance ran between March 2015 to January 2018, delivering a range of resilience-related activities in eight townships across the country (see <http://www.braced.org> for further details). It consisted of a consortium of NGOs led by Plan International and consisting of five partner agencies: ActionAid, World Vision, BBC Media Action, the Myanmar Environment Institute and the UN Human Settlements Programme (UN-Habitat).

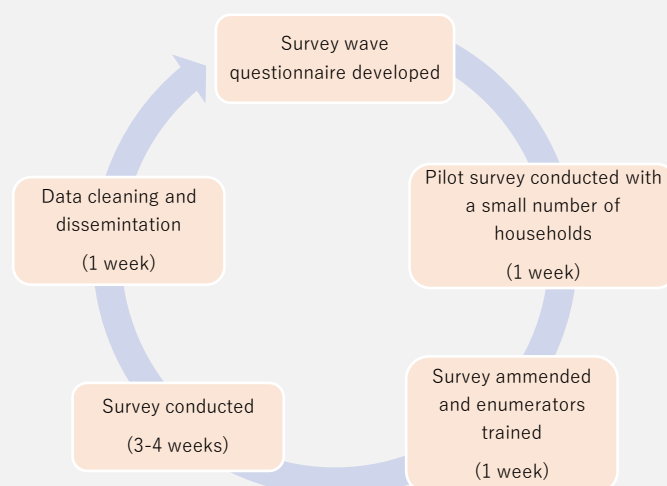
BRACED's panel surveys were delivered in two parts. The first step involved a face-to-face household survey (1,203 households in Hpa An and 767 in Mudon). Surveys were carried out with all households in the Township, effectively constituting a census of the immediate area. In each household, one member of the bread-winning couple was selected to answer questions, and continued as the contact-point for the remainder of the panel. To ensure diversity between males and females, selection of the respondents within each bread-winning couple was randomised. Respondents were then asked an extensive set of livelihood and socio-economic questions in seeking to gain a full understanding of household-dynamics.

After completion of the face-to-face survey, all respondents were handed a mobile phone (Singtech G9) and a small solar array. Phone numbers of other household members (as well as immediate neighbours) were also collected to ensure that respondents were easily contactable. These actions facilitated the second step in the panel: a series of mobile phone surveys. Immediately after completion of the baseline survey a call centre was set up in the city of Yangon. The set-up was used to collect remote oral surveys delivered via a team of 8-10 enumerators. Questions covering a range of resilience-related topics and were limited to 10-12 minutes in order to avoid drop-out and fatigue. Respondents received a small financial incentive to take part (\$0.5 of airtime credit) after successful completion of the survey.

In total, eleven waves took place as part of the panel survey in Hpa An (1 face-to-face and 10 mobile phone), with four conducted in Mudon. Each wave took roughly 6-8 weeks to complete, from questionnaire

design to delivery of data. In each round, a set of successive actions were carried out starting off with the development of a draft questionnaire. The questionnaire was comprised of a set-list of resilience-related questions, as well as a series of questions based on alternating themes that differed across waves. Questions were piloted with a small number of households (typically 15-20) to test for comprehensibility and relevance. Enumerators subsequently underwent a short period of training, before respondents were contacted as part of the panel – a process that typically took between 3-4 weeks to complete. Enumerators made use of an automated dialling system to carry out the surveys, with respondents randomly dialled. In instances where respondents were too busy, or unable to answer, arrangements were made with the enumerator to complete the survey at a more convenient time. Respondents were dropped from the panel if they failed to respond in three successive waves.

Figure 1: Cycle of actions carried under successive waves of the BRACED panel survey



Notes: Timelines are approximate, and represent the average length of time taken in each step of wave preparation and delivery. Exact times varied for each wave.

Given that surveys were carried out using tablets (in the case of the face-to-face baseline), and computers for mobile phones surveys thereafter, a wealth of meta-data are available. These include interview timestamps, number of attempted calls, measures of call duration as well as enumerator logs. These data form much of the basis of the analyses below – combined with the rich socio-economic profiling of households gathered during the baseline surveys. In exploring these data together, we seek to answer the main research questions using a combination of basic descriptive analyses, alongside a series of regression

analyses to further account for confounding influences (further methodological details are described in the sections below).

Research design and procedures in the Hpa-an and Mudon exercises were largely identical. The main difference between the two exercises was the length of time taken between the face-to-face baseline and first wave of the mobile phone panel survey. In the case of Hpa-an, the average difference in time between the two surveys was 45 days. For Mudon the gap was far shorter, with an average of just 8 days. In addition, a period of heavy seasonal flooding was experienced in the Hpa An between the baseline and the first two phone surveys. It is highly likely that these events impacted heavily on self-reported resilience scores between the two modes of survey administration. Though Mudon was similarly affected by flooding, the majority of these events took place prior to the baseline survey, and are therefore less likely to act as a confounder in explaining differences between the two modes of survey administration. It is for these reasons that data from the Mudon site are largely preferred in exploring possible mode effects in the analyses that follow.

Alongside quantitative analysis of survey responses and meta-data, we also sought to gather qualitative insights into opportunities and challenges faced in running the BRACED mobile phone panel. To do so, we conducted two key-informant interviews with managers of the survey. Namely, the director and team leader of Third Eye Ltd., the survey company contracted to carry out the surveys in Hpa An. The interviews were semi-structured in nature, probing successes and barriers to survey administration, as well as potential causes of key trends from the quantitative analysis. Insights from the interviews are used to supplement the main statistical analyses throughout.

b. Measuring household resilience

In order to track household resilience over time, the BRACED survey used a self-reported measure. Namely, the Subjective self-Evaluated Resilience Score (SERS) devised by Jones et al. (2018). The module is designed to illicit individual's perceptions of their household's ability to deal hypothetical future risk. SERS is comprised of a series statements each linked to a resilience-related capacity. Respondents are asked to rate levels of agreement with respective statements using a five-item bipolar response scale, ranging from Strongly Agree to Strongly Disagree. Scores are then converted numerically¹ and the average across all questions is computed. This results in an overall resilience score that ranges from 0 (not at all resilient) to 1 (fully resilient).

¹ This assumes cardinal comparability, an aspect that is relatively widespread in related fields such as subjective well-being (OECD, 2013), with some evidence to suggest the practice is robust (Kristoffersen, 2017).

Table 1: List of nine resilience-related capacity questions used in the SERS model of overall resilience

Resilience-related capacity	Question
<i>Preamble: 'I am going to read out a series of statements. Please tell me the extent to which you agree or disagree with them.' [Read out each statement and ask] 'Would you say that you strongly agree, agree, disagree, strongly disagree or neither agree nor disagree?'</i>	
Absorptive capacity²	Your household can bounce back from any challenge that life throws at it
Transformative capacity	During times of hardship, your household can change its primary income or source of livelihood if needed
Adaptive capacity	If threats to your household became more frequent and intense, you would still find a way to get by
Financial capital	During times of hardship, your household can access the financial support you need
Social capital	Your household can rely on the support of family and friends when you need help
Political capital	Your household can rely on the support of politicians and government when you need help
Learning	Your household has learnt important lessons from past hardships that will help you better prepare for future threats
Anticipatory capacity	Your household is fully prepared for any future natural disasters that may occur in your area
Early warning	Your household receives useful information warning you about future risks in advance

In the context of this study, we use two variants of the SERS module – one comprising all nine capacities (herein referred to as SERS-9A) and one comprising of a smaller subset of three capacities (referred to as SERS-3A). Importantly, SERS measures a household's overall resilience (its ability to deal with a range of overlapping risks) rather than one specific hazard.³ This recognises the multifaceted nature of resilience – people rarely respond to a single threat, but rather to multiple overlapping threats that co-evolve over time. It also helps prevent framing bias: drawing attention to a specific hazard may bias people's recollections and perceived impacts of the threat (OECD, 2013). Moreover, the questions explicitly make no reference

² This definition of transformation used here is based largely on the ability of a household to modify livelihood activities when and if required – see Béné et al. (2012) and Kates et al. for more (2012).

³ Hazard-specific variants of the SERS model exist and are detailed in Jones et al. (2018) and Jones (2018).

to resilience in the question structure, in order to limit confusion with many different interpretations of resilience. For full details on the approach see Jones (2018).

3. RESULTS AND DISCUSSION

a. Response rates

Attrition and non-response are often seen as major threats to mobile phone panels in developing countries (Dabalen et al., 2016).⁴ For example, the World Bank's 'Listening to Latin America' project saw response rates dropping to 72% in Honduras after calling the panel for four and half months of calling, and to 39% in Peru after six months (Ballivian et al., 2013)⁵. Similarly, during the four rounds and four months of the South Soudan Experimental Phone Survey, no more than 48.5% of the original panel continued responding (Demombynes, Gubbens, Rome, 2013).

Other mobile phone panels have been more successful in maintaining high response rates even after many months of surveying. In one of the first documented efforts to gather panel data through mobile phones in low-income countries, Dillon (2012) observed an average response rate of 98% across the 14 rounds and ten months of surveying a panel of cotton farmers in Tanzania. Similar high response rates were achieved by the World Bank's Listening to Madagascar panel. Two years and 18 mobile phone rounds after the baseline, the survey team was able to still achieve a response rate of 92% (Dabalen et al., 2016). Importantly, however, in order to achieve such high figures, the project team had to conduct field revisits to boost engagement and to bring dropped-out respondents back into the panel (Dabalen et al., 2016; p.37).

What were response rates in BRACED's Hpa An surveys?

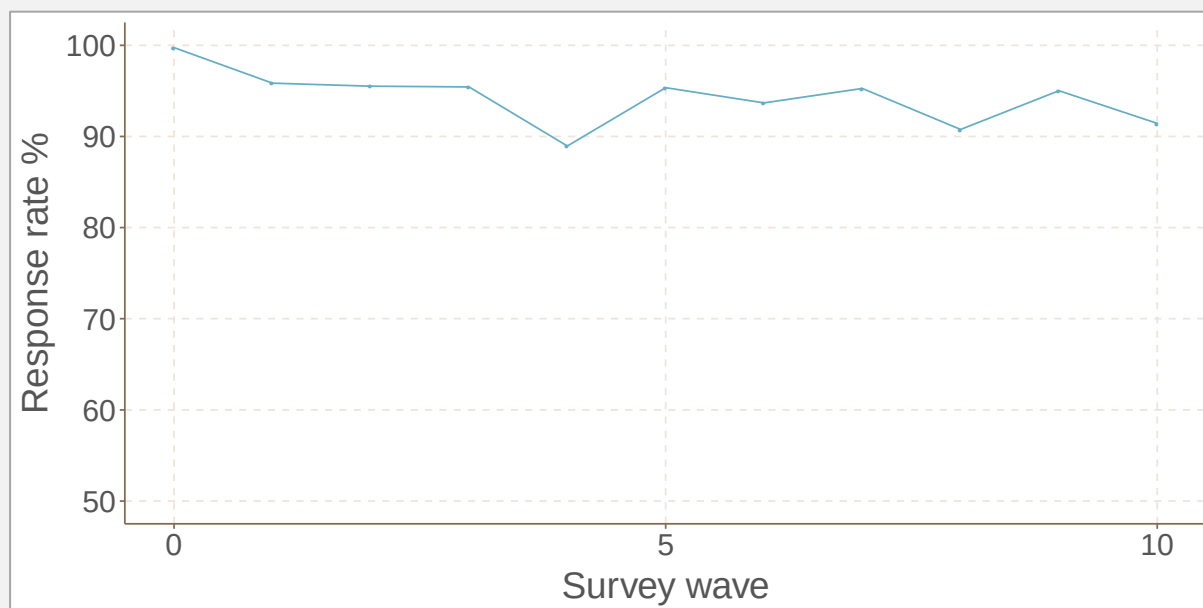
Ten BRACED mobile phone survey rounds had been conducted in Hpa An from June 2017 to March 2019. Throughout this period, response rates remained consistently high and showed no sign of significant decline. No less than 91% of the original panel still participated in round 10. In other words, 91% of those

⁴ It should be noted that, in the context of panel surveys, the term attrition is reserved for respondents completely dropping out of the panel, while non-response refers to respondents not available for interviewing in one or several survey rounds.

⁵ Given the experimental setup of the project, these figures have to be treated with some caution, as they also include respondents in a no-incentive condition and thus likely underestimate response rates in the incentivized part of the panel. It should also be noted that only those respondents who did not yet own a phone were provided one for free.

households⁶ who had agreed to participate in the mobile phone panel during the baseline in June 2017, were still answering questions through the phone some 21 months later. In the nine intermediate rounds, response rates were similarly high, varying between 89% in the fourth round (January 2018) and 96% in the first wave right after the baseline (August 2017).⁷

Figure 2: Response rates over the course of BRACED's Hpa An panel survey



Note: Wave 0 relates to the face-to-face baseline survey. Subsequent phone surveys are listed as 1-10

In the following, we aim to explore in some depth factors that contributed to these consistently high response rates in the BRACED panel.

⁶ As we will detail below, the BRACED panel used the household, rather than the individual, as its primary unit of analysis. In practice, this meant that if a respondent could not be reached but another adult member of his or her household could, that other person was considered eligible to complete the survey.

⁷ The slight dip in the fourth round can be explained by the fact that up until and including that round, respondents who had missed a survey were considered as dropped-out and not contacted again in following rounds. From round 5 onwards, the entire pool of respondents was approached again each time, irrespective of prior response behaviour.

Interviewers' persistence & tracing strategies

The analysis of meta-data on contact attempts highlights the crucial role of a dedicated and persistent interviewing staff. Interviewers had to place an average of 2.9 calls in order to complete one interview. This average varied across rounds, with a minimum of 2.5 (wave 1) and a maximum of 3.5 (wave 6). As expected, however, some respondents proved to be much more difficult to reach than others. For example, in wave 8, conducted right after our target communities had been hit by widespread flooding, 13% of the interviews could be completed only after seven or more attempts to reach them. During wave 7, 17% of respondents were interviewed only after more than five days of trying to reach them (despite the fact that 96% of respondents were eventually reached during the wave).

Such high levels of perseverance on the side of the interviewer team is crucial for reaching the 'high-hanging fruit' respondents. Successfully interviewing as many of these respondents as possible is important not merely from a perspective of keeping the panel alive, but also as this group, as we will see below, tend to differ systematically from their more easily reachable co-panellists.

An additional factor that contributed to the high response rates that were achieved was the fact that when respondents were not reachable on their own phones, the survey team could rely on other channels to get a hold of them. For this purpose, enumerators had been instructed during the face-to-face baseline visits to the communities to record phone numbers not only of existing phones the respondents might have, but also of other household members, neighbours and community leaders. Following best practice for setting up mobile phone survey panels (see Hogeveen et al., 2014; Dabalen et al., 2016), this was done to facilitate future tracing of respondents that are unreachable on their own phones.

This strategy clearly paid off. Already in round 1, no less than 22.3% of respondents were reached through someone else's phone. This proportion averaged at 26.4% across the first nine rounds, with the last two rounds seeing the highest figures at 35.3% and 35.0%, respectively. In other words, in both of the last two rounds, more than one third of respondents had to be traced through other people's phones, in almost all cases their neighbours' (34.2% and 32.5%).

In tracing respondents, the BRACED panel also benefitted from its specific sampling design. With its limited geographical focus – comprising only eight villages – and inclusion of almost all inhabitants of that area, the study was essentially set up as a de-facto census. This approach is similar to that successfully employed by Dillon (2012) in his study of cotton farmers communities in Tanzania. This design meant that literally everyone in the communities knew about the project and was aware when calls were being placed. This made it much easier for interviewers to trace down unreachable respondents through others living in those communities.

When asked about why they could not be reached on their own phones, the BRACED respondents most often pointed to phones being switched off or problems with the solar chargers, changed SIM cards and limited network coverage. These findings are in line with what other large-scale phone surveys have found to be the main obstacles to keeping non-response and attrition levels low (Dabalen et al., 2016, p. 65; Twaweza, 2013).

Using households as primary units of analysis

Given the primary purpose of the BRACED panel to trace households (post-shock) resilience, sampling happened on the level of the household, rather than the individual. This approach has also been applied in other mobile phone panel projects such as the World Bank's L2LAC project, for largely pragmatic reasons as phones are often shared among adult household member (Ballivian et al., 2013).

For the case of the BRACED panel, it is clear from an analysis of the meta-data that this research design decision also ended up helping to minimize non-response. As respondents were first called on the numbers collected during the face-to-face baseline survey, interviewers typically spoke to the respondent who was interviewed in the baseline. However, if that respondent was not available to be interviewed, another adult household member was interviewed. This approach was deemed practicable for the BRACED panel, as most concepts of interest that were tracked over time were household- rather than individual-level. Typically, about one in ten interviews per round were conducted with another household member. However, during round 8, conducted in the immediate aftermath of massive floods in the communities, more than half of the interviews were held with another household member (54%).

Exploring the predictors of non-response

In order to identify factors that might affect whether a respondent is more difficult to reach for a phone survey, or cannot be reached at all, two regression analyses were conducted. A review of literature on previous mobile phone panel studies helped to identify a set of potential predictors (Ballivian et al., 2013; Demombynes, Gubbins, and Romeo 2013; Hoogeveen et al. 2014; Schipper et al., 2014; for overview of these, see Dabalen et al., 2016, p. 67).

These predictor variables are:

- a household's vulnerability to poverty, expressed as its POP score⁸

⁸ The Progress out of Poverty (POP) approach matches household-level survey data with census data of the same country to estimate the likelihood of a household falling below or above a given poverty line. For a detailed discussion, see Desiere et al. (2015).

- whether a household's main income comes from agriculture
- age of the respondent
- size of a household
- whether the household already owned a mobile phone prior to the baseline
- gender of the respondent
- whether the head of household has had any formal education
- the village tract of the household⁹

The first regression model aimed to identify respondent characteristics that affected non-response, i.e. non-participation in one or more survey rounds. A logistic regression analysis was conducted, predicting whether a respondent had missed at least one of the eleven survey rounds. The overall model was statistically significant ($\chi^2(14)=178.6$, $p < 0.001$) and its results are summarized in Table 2.

⁹ A set of dummy variables was included in the regression models to estimate the effects of a household's residency in one of the eight village tracts covered by the BRACED panel. Among other possible differences between location, it was expected that variation in network coverage between village tracts will affect a respondent's likelihood to be reachable on their mobile phones.

Table 2. Logistic Regression results, predicting non-participation in at least one survey round

	B	SE	OR
POP score	-0.014†	0.009	0.99
agriculture main income source	0.212	0.198	1.24
age	-0.005	0.006	0.99
household size	-0.130*	0.054	0.88
gender (female=1)	-0.225	0.169	0.80
completed primary education	-0.194	0.205	0.82
household already owned phone	-0.066	0.147	0.94
Village Tract 1	-0.781	0.534	3.28
Village Tract 2	0.408	0.405	6.72
Village Tract 3	1.124*	0.504	5.60
Village Tract 4	0.943*	0.427	3.31
Village Tract 5	0.418	0.394	6.71
Village Tract 6	1.123*	0.510	2.31
Village Tract 7	0.058	0.439	2.18
Constant	-0.453	0.692	0.29
Model χ^2	178.6		
Pseudo R ²	0.16		
N	1,136		

Note: †p<0.1; *p<0.05; **p<0.01. Robust standard errors are shown in SE column

Results show that household size has a significant effect on non-response ($p<0.05$). Living in larger households decreased the likelihood of having missed any survey rounds. As described above, interviewers invested considerable efforts into tracing down respondents that could not be reached on their primary phone numbers. Using the phone numbers of other household members that were gathered during the baseline played an important part in that. It is therefore likely that someone living in a larger household was more likely to still be reachable, even if their own phone was switched off. Along the same line, even if the primary respondent could not be reached, all other adult household members were considered eligible to

answer the questions. Also, in that sense, larger households thus had a smaller chance of not being reachable at all.

The significant effects of the village tract dummy variables suggest that, once controlled for other demographics, respondents living in three of the eight village tracts were significantly less likely to have participated in all rounds compared to the omitted baseline village tract.

However, when interpreting the results of the logistic regression, it is important to keep in mind the high response rates discussed above. As a result, the outcome variable in this model had limited variance, with 22% of respondents having missed one or more rounds (218). Moving beyond the binary distinction between participation and non-response, a second regression model was therefore conducted, exploring factors that influenced how much effort the interviewers had needed to invest into reaching the respondent. After all, if the call centre staff had invested less effort into calling back and tracing respondents, many of the ‘high-hanging-fruit’ would have ended up in the group of non-responses. The dependent variable was operationalized as the average number of call attempts it took to reach and interview a respondent across all rounds in which they participated¹⁰.

The independent variables in this OLS regression were the same as in the previous model. The overall model was statistically significant ($F(14, 1121) = 19.253, p < 0.01$) and its results are summarized in table 3.

¹⁰ While this approach makes it possible to distinguish well between those who were easily reached and those that required various call attempts to be interviewed, it does not take into account non-participation in a given round. The average call attempts value that was used as the dependent variable is calculated over only those rounds in which a respondent actually participated.

Table 3. OLS Regression results, predicting average number of call attempts needed to reach respondent

	B	SE
POP score	-0.006*	0.003
agriculture main income source	0.100	0.080
age	0.0001	0.003
household size	-0.022	0.017
gender (female=1)	-0.098	0.069
completed primary education	-0.254**	0.081
household already owned phone	-0.301**	0.056
Village Tract 1	-0.676**	0.158
Village Tract 2	0.020	0.137
Village Tract 3	0.002	0.190
Village Tract 4	0.911**	0.174
Village Tract 5	-0.392**	0.133
Village Tract 6	-0.541**	0.210
Village Tract 7	0.381*	0.168
Constant	3.781**	0.233
F	19.253	
adjusted R ²	0.184	
N	1,136	

Note: †p<0.1; *p<0.05; **p<0.01 Robust standard errors are shown in SE column

In comparison to the previous model, a somewhat different picture emerges. The level of formal education of the household shows to have a significant effect on the average number of calls it took to reach a respondent. Those living in a household headed by someone with at least primary education were in general more easily reached. Furthermore, it took less call attempts to speak to respondents in households at lower risk of poverty – reflected in a higher POP-score. Maybe unsurprisingly, reaching respondents who had provided at least one additional secondary phone number during the baseline also proved less difficult than reaching those who had not. In addition, where a respondent lived also mattered, with four village tract

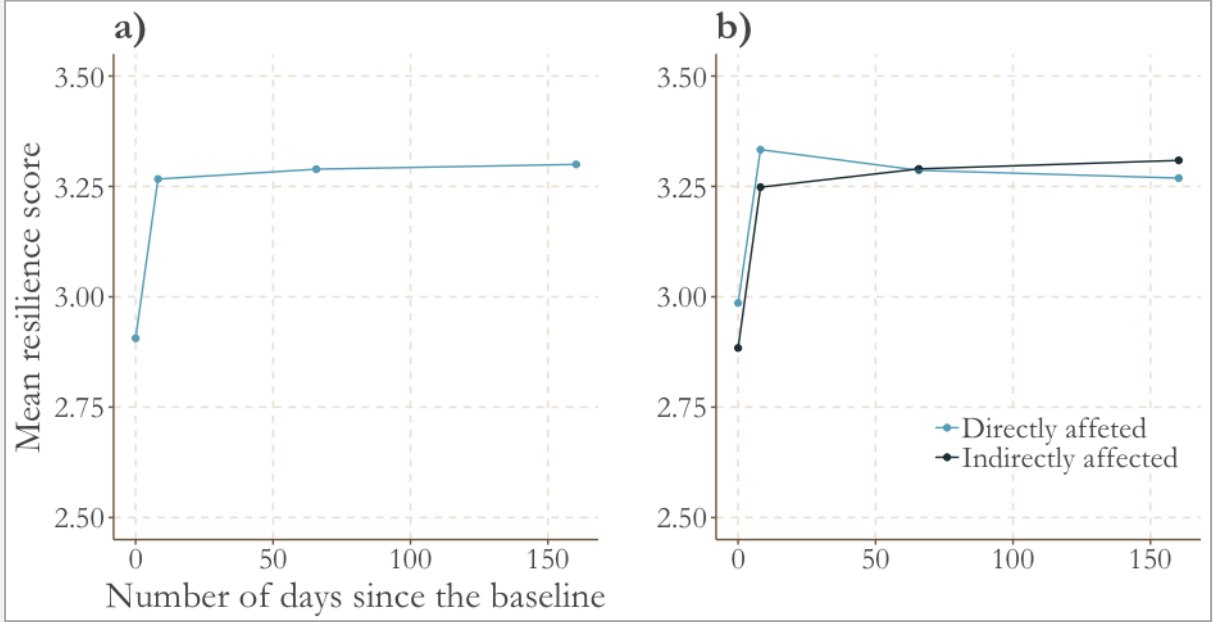
dummy variables showing a significant effect on average call attempts, compared to the omitted baseline village tract.

b. Mode effects

One important aspect of running a multi-mode panel survey is understanding whether different forms of administration will affect survey responses. Indeed, mode-effects are known to play a significant influencing role in surveys across a number of fields and disciplines (Greenleaf et al., 2017; Mahfoud et al., 2015; Heath et al., 2016; Ballivian, Azevedo & Durbin, 2015).

In the case of the RRR, the primary variable of interest is subjectively-evaluated resilience (measured via the SERS-3A module). Owing to the close proximity between F2F and phone waves, the Mudon dataset can provide an invaluable opportunity to shed light on the presence of mode-effects in the RRR survey. Importantly, as part of the survey design, call centre enumerators were tasked with contacting respondents via mobile phone as early as possible in the days that followed the F2F baseline survey. In practice, the mean waiting time between baseline and W1 was 8 days across the sample. Yet, despite the small interval, Figure 3a) reveals a large uptick in mean resilience scores between baseline and Wave 1. The difference is equivalent to a +0.36 rise in mean SERS score, and is statistically significant¹¹ ($t = 8.8$). Interestingly, scores appear to level off in subsequent waves, with collective gains of only +0.02 and +0.01 for Waves 2 and 3 respectively (each of which conducted via mobile phone).

¹¹ Differences in means are calculated using a paired t-test comparing SERS-3A scores from Baseline and Wave 1 of the RRR survey

Figure 3: Changes in mean subjectively-evaluated resilience (SERS-3A) over time for Mudon

While it is likely that responses to the RRR survey will have been affected by the wide-spread seasonal flooding that was experienced in the weeks prior to the baseline, Figure 3b) shows that the same uptick is clearly visible in both directly and indirectly affected households within the sample.

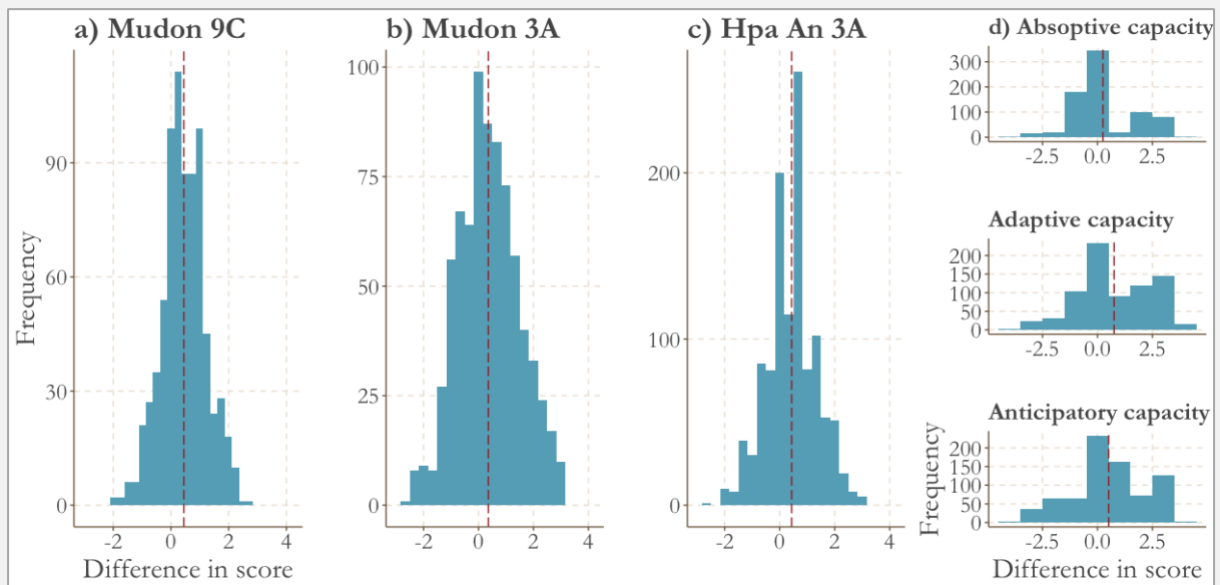
In many ways, the existence of mode effects is not a problem for comparisons of resilience over time, as long as any effects can be shown to be normally distributed and systematic across socio-economic groups. To examine these assumptions further, we calculate a new variable $Wavediff_h$, equivalent to the difference in resilience score for household h between the baseline ($Resilscore_{ht0}$) and Wave 1 ($Resilscore_{ht1}$)

$$Wavediff_h = Resilscore_{ht1} - Resilscore_{ht0} \quad 1)$$

By plotting the difference in resilience scores for both SERS-3A and SERS-9C variants (Figures 3a and b), we see that they are somewhat normally distributed with positive means. This implies that there is a positive change in resilience scores between F2F and mobile phone modes of administration. Indeed, the same general pattern is visible when looking at differences in resilience scores for the Hpa An sample (Figure 4c),

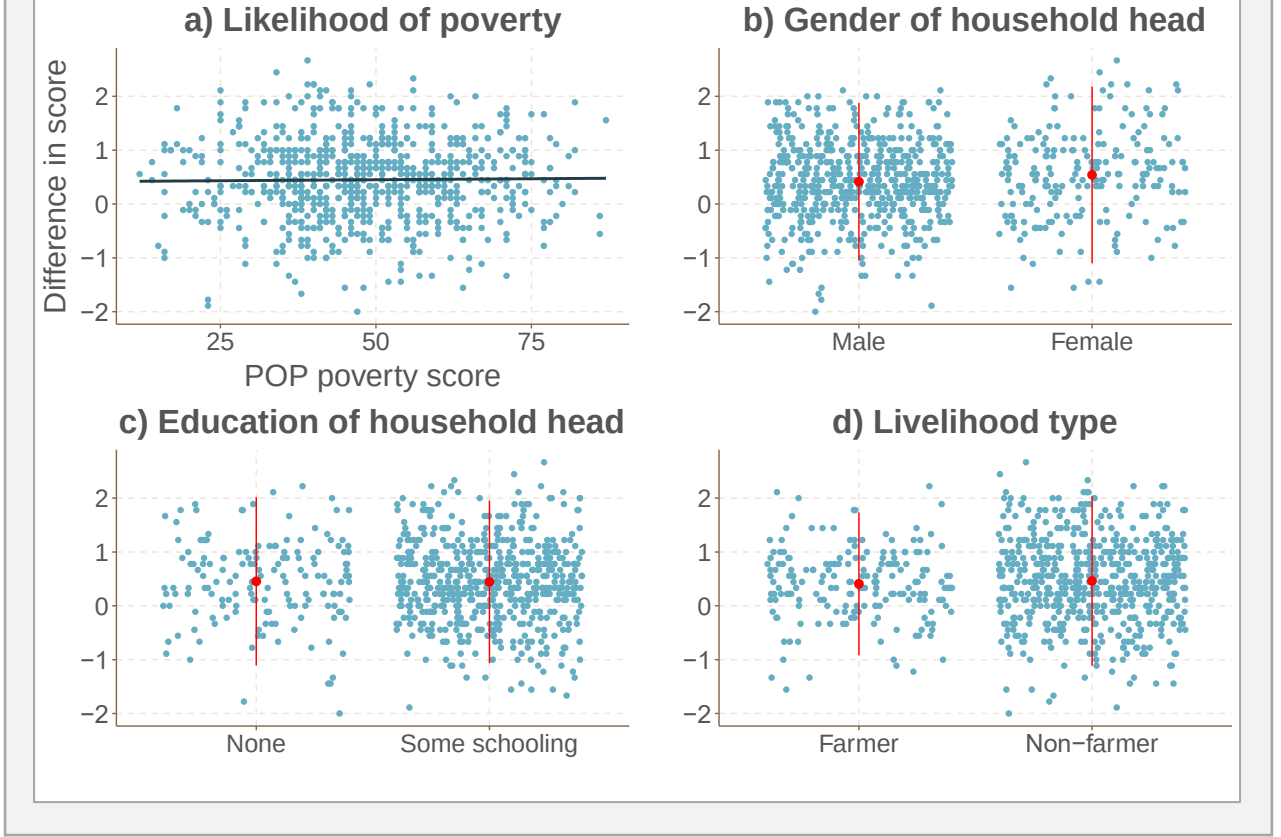
suggesting that mode effects play a similar role in this site too¹². In breaking differences down by individual resilience capacities, we see slight distributional differences – with right skews for adaptive and anticipatory capacities. However, all three of the capacities that make up the SERS-3A module showcase positive mean values (again hinting at a positive difference between F2F and mobile phone surveys).

Figure 4: Distributions of differences in resilience scores between F2F and mobile phone surveys for a number of properties of interest



¹² Note that we choose to focus on the Mudon sample when investigating potential mode-effects within the RRR survey. This is due to the fact that a) widespread flooding took place between the two time periods in Hpa An (rather than prior to surveying as in Mudon); and b) a far greater length of time passed in carrying out the follow up mobile phone surveys (an average of 44 days). Differences in resilience scores are therefore likely to be heavily influenced by a number of confounding factors that are difficult to account for. It is, however, reassuring to see that trends are broadly consistent across the two study sites.

Figure 5: Differences in resilience score between F2F and mobile phone waves by key socio-economic groups



To examine the second assumption (no differences between groups), we can plot differences in resilience scores across a number of important socio-economic characteristics. Figure 5 suggests that while mode-effects may exist, there are few visible differences between factors such as likelihood of poverty, education of the household head and whether the household is primarily reliant on farming as a source of livelihood (the most common vocation in Mudon). There may, however, be slight differences between the gender of the headed household, with differences slightly larger for females than for males (+0.48 and +0.32 respectively).

In order to test these differences more formally, we run a series of OLS regressions. Firstly, we model the difference in resilience scores between F2F and mobile phone waves for household b in village v ($Wavediff_{bv}$) as a function of key socio-economic characteristics ($Socioecon_{bv}$) and a dummy variable for whether the household was directly affected by flooding in the weeks prior to the baseline

(*Flooding_{hv}*). To account for any geographic differences, we also include a set of village-level fixed effects (Φ_v), with an error term represented by e_{hv} .

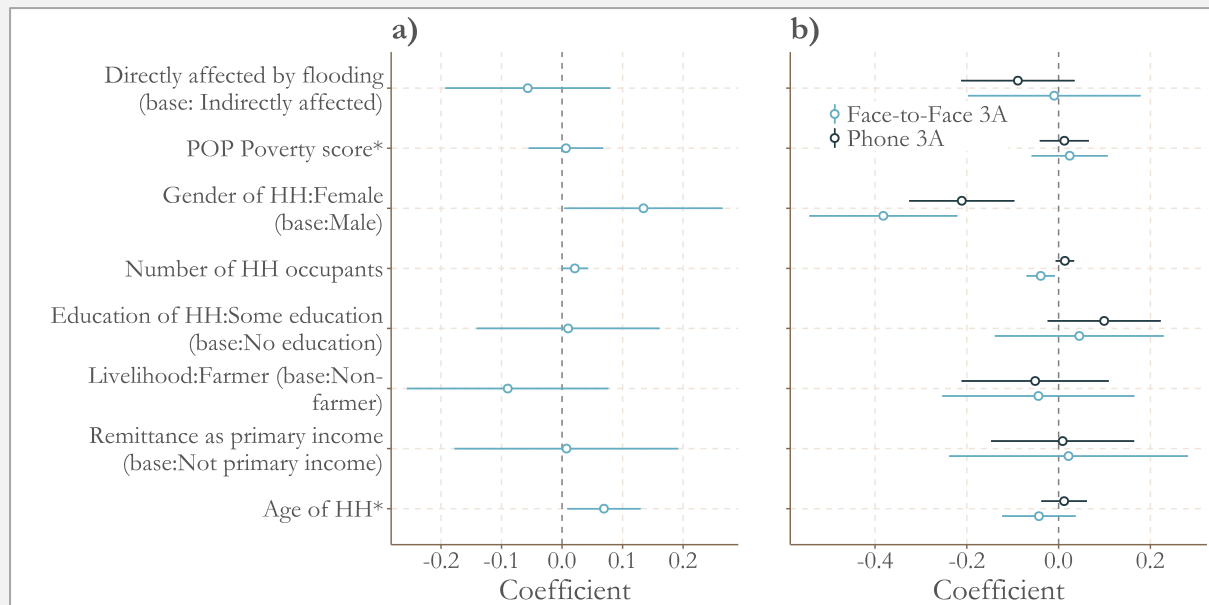
$$Wavediff_{hv} = \beta_1 Socioecon_{hv} + \beta_1 Flooding_{hv} + \Phi_v + e_{hv} \quad 2)$$

To provide further clarity, we replicate equation 2 replacing the outcome variable $Wavediff_{hv}$, with resilience scores for the F2F ($ResilscoreF2F_{hv}$) and mobile phone ($Resilscorephone_{hv}$).

$$ResilscoreF2F_{hv} = \beta_1 Socioecon_{hv} + \beta_1 Flooding_{hv} + \Phi_v + e_{hv} \quad 3)$$

$$Resilscorephone_{hv} = \beta_1 Socioecon_{hv} + \beta_1 Flooding_{hv} + \Phi_v + e_{hv} \quad 4)$$

Doing so allows side-by-side comparisons of the influence of various socio-economic factors on the two modes of survey administration. Robust standard errors are reported for all models.

Figure 6: Differences in resilience score (SERS-3A) and their association with key socio-economic characteristics

Notes: Age of HH and POP poverty scores are standardised (by dividing by the standard deviation). All models include Village-level fixed effects. Dots represent beta coefficients and thin lines are 95% confidence intervals using robust standard errors.

Figures 6a showcase associations between changes in resilience score for F2F/phone waves and a range of socio-economic factors. Firstly, we note the lack of statistically significant association between those directly and indirectly affected by flooding; a strong association would otherwise suggest that comparisons of the two groups across a multi-mode panel are biased. Interestingly, associations with most other socio-economic factors similarly exhibit no statistically significant associations (at the 95% CI level). This may tentatively suggest that mode effects are, by and large, systematic across groups. This claim is somewhat supported by Figure 6b, where effect sizes and confidence intervals can be seen as broadly in line for both F2F and mobile phone modes of administration.

However, there are two important exceptions: the age and gender of the household head. Relating to the former, larger score differences are associated with older household heads (with older respondents reporting lower resilience scores over the phone). Effect sizes are small with a one standard deviation increase in age resulting in a 0.07 change in reported resilience (Figure 6a). Relating to the latter, female-headed households appear to have a larger difference in scores compared with males. The effect is sizeable, with females exhibiting a 0.13 increase in score between F2F and phone modes. Indeed, the direction of this effect (a positive change between F2F and phone administration) is particularly interesting. In both the Hpa An and Mudon samples, female-headed households have significantly lower resilience scores than

male-headed households. If results from Equation 2 are robust, then they would suggest that female headed households fare even worse than reported once differential mode effects are accounted for.

4. DISCUSSION AND CONCLUSION

As mobile phone surveys continue to grow in popularity globally, it is essential that the effects of non-response and attrition are well understood and that survey outcomes can be readily compared with those from traditional face-to-face surveys. The latter is particularly the case for perceptions-based questions that are known to be affected by a number of biases, including social desirability and priming. In this paper, we have taken advantage of a large dataset to explore attrition and response behaviour and compare both modes of survey administration in short proximity. To provide structure to the various findings presented in the paper, we return back to the three research questions initially posed in the Introduction.

a. Do responses differ between face-to-face and phone modes of survey administration?

Evidence from the BRACED survey suggests that modest mode effects do indeed exist when measuring perceived levels of resilience between face-to-face and mobile phone surveys. The positive jump in scores (with an average difference of +0.36) while transitioning to mobile surveys is marked, and statistically significant. More strikingly, it contrasts with relatively little variation in resilience scores between subsequent waves of the phone survey in Mudon. These findings are similar to those reported by Dolan and Kavetsos (2016) who document how individuals consistently report higher subjective wellbeing scores over the phone when compared to face-to-face surveys.

As with the Dolan and Kavetsos study, it is difficult to say what may be driving the reported differences in scores between modes. One interesting point to consider is whether the gap in scores is an artefact of social desirability? Sadly, answering this depends on the direction with which desirability is likely be expressed in the context of resilience in Mudon. On the one hand, it's possible to reason that individuals may wish to express higher levels of household resilience. Doing so would likely reflect positively on the part of the BRACED programme, in the hope of providing continued future support – even though survey participants were explicitly notified that survey results are decoupled from future funding initiatives. Contrastingly, respondents may have an incentive to report lower scores, assuming that this comes with greater need for external assistance and funding. In both cases, theory posits that closer proximity to enumerators in face-to-face modes of administration is likely to result in answers that respondents perceived to be socially desirable.

What is clear from the from the survey responses, however, is that the former scenario is unlikely to be driving results in Mudon: resilience scores are lower for face-to-face interviews despite enumerators being

of closer proximity to respondents. While the latter scenario may therefore be a fit in explaining differential scores between modes in the Mudon survey it is important to note that a range of other biases may be at play. For example, the greater perceived ‘distance’ between enumerators and respondents under phone surveys may result in respondents devoting less cognitive attention and focus to the interview process than compared with the face-to-face equivalent – though it’s unclear whether this would systematically result in scores being affected in one direction (Holbrook et al. 2003; Dolan and Kavestos 2016). Moreover, wide empirical evidence of the influence of social desirability bias between survey modes is mixed. A recent study conducted by the World Bank in Madagascar (Ndip & Nwokike, under review) revealed a number of significant differences in data gathered during phone and face-to-face interviews that are also not easily brought in line with this social desirability thesis. Similarly, a study by the World Food Programme (2017) found that respondents in Kenya reported higher dietary diversity and food intake when interviewed face-to-face, seemingly refuting the idea that social-desirability bias is less of an issue in phone interviews. With the above in mind, it is likely that the influence of social-desirability in explaining mode effects is highly context specific. Disentangling it, together with other cognitive biases, is likely to require detailed understanding of local political, institutional and economic drivers, alongside careful survey design.

The implication of modest mode effects for subjectively-oriented resilience modules on policy is profound. As development funders increasingly turn to use of a diverse range of survey modes in evaluating the impact of resilience-building interventions, caution must be expressed when comparing results between different types of survey administration. Here the danger is that that evaluators may attribute gains/losses between baseline and end-line surveys to the impacts of a resilience-building intervention, when a portion of any differences may result from mode-effects. Indeed, this would have been the case in the context of Mudon – noting that aggregate resilience levels in the subsequent phone surveys remained relatively constant. Where possible, evaluators should be encouraged to be consistent in the mode of survey administration when collecting panel or repeated cross-sectional data. We also note that findings from a survey cannot establish whether one mode of administration is more robust than the other. While further research may want to compare results from face-to-face and phone surveys with alternative resilience metrics, the lack of a ‘gold standard’ for resilience measurement makes this a challenging task (Constas, Frankenberger & Hoddinott, 2014).

It is also important to recognise that mode effects may not be a problem for evaluators interested primarily in within-sample comparisons. However, the assumption here is that any mode effects are systematic across different social groups (and consistent over time). For example, so long as all individuals in a sample are affected in the same way, comparisons between a treatment and control group (or any other socio-economic groups) across two different survey types remain valid. This is indeed a large assumption. One that Figure 6 helps to shed considerable light on. Findings suggest that for the most part, mode effects are not

significantly different across key socio-economic groups in the context of Mudon. However, two exceptions exist in the form of gender and age of the household head. Reasons for this are likely to require further quantitative and qualitative insights, though may result from a combination of cognitive, cultural and institutional factors. Similar to conclusions for across-sample comparisons, these results suggest that evaluators should be cautious when comparing within-sample scores in instances where two different modes of administration have been used.

b. Why do respondents stay in the panel? Intrinsic and extrinsic motivations

After each completed interview, respondents in the BRACED panel received a mobile phone credit top-up equivalent to \$0.50. Following common practice for mobile phone panels (Dabalen, et al. 2016; Morell & Leo, 2016), this monetary incentive was meant to increase respondent's motivation to stay in the panel, and to thus minimize non-response and attrition. Conversations with the local research team on their interactions with the respondents did indeed confirm the expectation that the top-up was generally much appreciated as a token of appreciation. However, it also became apparent that this form of extrinsic motivation alone would not have sufficed to keep panellists engaged for such a prolonged period of time. As the local project manager in charge of overseeing the phone interviews explained, monetary incentives turned out to be only part of the story: "The credit top-up should not be seen as a remuneration [of respondents' time], but just as a Thank You. It helps of course, but a high level of buy-in and trust on the side of the respondents is at least equally important".

For the BRACED panel, the face-to-face baseline was essential for fostering these high levels of buy-in and trust. This face-to-face encounter with the interviewers who took the time to explain in detail the project to respondents, helped enormously to keep engagement high throughout the duration of the project.

Overall, it is clear that the BRACED respondents saw their participation as more than merely a financial transaction. In fact, when asked whether they would agree to be interviewed in the future, even if no credit top-up would be provided, an overwhelming majority said that they either might (55.1%) or would definitely (33.8%) still continue to participate in the panel. Most respondents also believed that the amount of \$0.50 they received after each interview was adequate (84.8%).

The idea that intrinsic motivation can play a key role for keeping attrition low in mobile phone panels, was also one of the findings that came out of the World Bank's Listening to Latin America project (Gallup, 2012). When respondents were asked whether and why they would be willing to participate in future research, the single most frequently mentioned factor was in fact the monetary incentive. However, as the authors continue to explain, "when adding up all the answers related in some way to the 'project's mission',

the motivation that this project will improve people's lives in the long-term seems to be significantly more important for panelists than one-time monetary incentives." (Gallup, 2012, p. 30).

In particular, the researchers found that a sense of "being heard" mattered for peoples' decisions to continue their participation. This certainly was also true for the BRACED respondents. When asked why – after more than 18 months of call rounds – she has always picked up the phone and readily answered all questions about her current living situation and hardships, one respondent explained: "I have been answering [your questions] because I wanted others to know what happened in our area. I [also] got knowledge from your questions. I shared information which happened in our village with you. You asked questions which you want to know from us. Like that, I wish that we can both benefit."

Drawing on these experiences, future mobile phones surveys should find creative methods of involving respondents in ways that foster this sense of purpose and of reminding them that they are contributing to something meaningful through their ongoing participation in the project. For example – depending on the purpose and context of the study – respondents can be informed about outcomes of prior survey rounds, or about concrete consequences that the findings might have had, if any. Such feedback mechanisms – also suggested by the local research manager as a way to boost respondent engagement even more in future studies – were not implemented in the BRACED panel.

REFLECTIONS ON SURVEY ADMINISTRATION AND MANAGEMENT

A number of wider lessons learned in running and operating the mobile phone panel were evident over the course of the 10 waves. Below we highlight key findings and reflection for future survey activities.

Setting up the call-centre

One of the most fundamental decision when setting up any mobile phone survey project, is whether calling operations will be organized in-house or outsourced to a call centre / research firm. While both options can have their pro's and con's depending on project aims, available resources and capacities¹, for the case of large-scale and long-running phone panels the advantages of hiring an professional survey organization will most likely outweigh its costs.

For the BRACED panel, a well-established research firm based in Yangon was hired for conducting both the face-to-face baseline survey and the follow-up phone survey rounds. While this organization had an extensive track

record conducting (survey) research in rural Myanmar, this was the first time for them to set up and manage a mobile phone panel.

A call centre was set up at the survey firm's offices, with interviewers placing calls wearing headsets and entering responses on a PC. In order to capitalize on available capacities, the firm used the same data entry software for the phone surveys as they would normally use for face-to-face surveys: Survey Solutions, developed by the World Bank. The sophisticated design and advanced customization affordances of this software means that, while initially developed for face-to-face surveys, an experienced user can adapt it very well for Computer Assisted Phone Interviewing (CAPI). It should be noted, however, that a whole range of possible software solutions are available, ranging from highly specialized call-centre software to simple data entry / survey systems.

The advantages of most specialized CAPI software is that they allow for more efficient process management and data quality checks. Such a centralized call center software will automatically assign respondents to interviewers, possibly based on the time preferences of the respondents and, possibly, other matching criteria such as language. The respondent's phone number will be loaded onto the screen for calling, and the interviewer also has access to other relevant information, such as the respondent's sociodemographic background characteristics or past nonresponse behaviour. Interviewers can therefore use also information such as name, age, and gender to confirm the identity of the respondent. In such a scenario, supervisors are able to monitor closely overall progress, as well as the performance of individual interviewers, regarding the number calls placed, contact and refusal rates and interview durations.

Recruitment and quality control

Besides the choice of software, the recruitment of qualified call centre staff evidently is essential for running a successful mobile phone panel. In setting up the BRACED panel, it soon became clear that even experienced face-to-face interviewers required training and close supervision during the first rounds, before becoming skilled phone interviewers. While, during face-to-face interviews, a smile, friendly gesture or handshake during introductions can help to build rapport and put the respondent at ease. In phone interviews, it is generally more difficult to quickly establish good rapport, and finding the right tone of voice can be a challenge.

Following common call-centre practice of quality control, researcher supervisors conducted spot-check call-backs to a sub-set of respondents. This helped to flag performance issues and identify interviewers who needed additional training.

In an effort to build a robust trust relationship across survey rounds, respondents were called by the same interviewer throughout whenever possible, as recommended in the literature (Dabalen, et al. 2016). According to the research team, this strategy did indeed contribute to keeping respondents engaged and willing to continue their participation.

Keeping the panel alive

Probably the biggest challenge in keeping the non-response and attrition low, was the fact that respondents frequently changed their phone numbers and that their phones are oftentimes not charged (despite the fact that solar chargers had been handed out). As discussed above, this meant that the interviewer team had to invest considerable efforts into tracking down respondents through other people in the village.

During the first few survey rounds, it also became clear that much flexibility was needed on the part of the interviewer when scheduling calls. While many male farmers who worked on their fields during the day tended to be best reachable during the evenings, women who worked at home could best be called in the late morning. Constantly adapting calling schedules to respondents' availability was key to ensure high response rates.

In conversations with the local research team on their successful efforts of keeping the panel alive for such a long time, the important role the face-to-face baseline played in fostering good-will and buy-in among respondents was a recurring theme. As the local research manager explained: "The key is that you know each other. This builds trust. Otherwise, this would not have worked like that."

REFERENCES

- Ballivian, A., & Azevedo, J.P. (2013). *Listening to LAC: Using Mobile Phones for High Frequency Data Collection. Final Report.*
- Ballivian, A., Azevedo, J.P., & Durbin, W. (2015). Using Mobile Phones for High-Frequency Data Collection. In: Toninelli, D, Pinter, R and de Pedraza, P (eds.) *Mobile Research Methods: Opportunities and Challenges of Mobile Research Methodologies*, pp. 21–39.
- Béné, C., Wood, R. G., Newsham, A., & Davies, M. (2012). Resilience: new utopia or new tyranny? Reflection about the potentials and limits of the concept of resilience in relation to vulnerability reduction programmes. *IDS Working Papers*, 2012(405), 1-61.
- Constas, M., Frankengerger, T., & Hoddinott, J. (2014). *Resilience Measurement Principles: Toward an Agenda for Measurement Design*. Resilience Measurement Technical Working Group.
- Dolan, P. & Kavetsos, G. (2016). Happy talk: Mode of administration effects on subjective well-being. *Journal of Happiness Studies*, 17(3), 1273-1291.
- Dabalen, A., Etang, A., Hoogeveen, J., Mushi, E., Schipper, Y., & von Engelhardt, J. (2016). *Mobile phone panel surveys in developing countries: a practical guide for microdata collection*. World Bank, Washington, DC.
- Desiere, S., Vellema, W., & D’Haese, M. (2015). A validity assessment of the Progress out of Poverty Index (PPI)TM. *Evaluation and program planning*, 49, 10-18.
- Demombynes, G., Gubbins, P., & Romeo, A. (2013). *Challenges and Opportunities of Mobile Phone-Based Data Collection. Evidence from South Sudan*. World Bank, Washington, DC.
- Dillon, B. (2012). Using Mobile Phones to Collect Panel Data in Developing Countries. *Journal of International Development*, 24(4), 518–27.
- Gallup (2012). *The World Bank Listening to LAC (L2L) Pilot. Final Report*. Available at <https://microdata.worldbank.org/index.php/catalog/2022/download/30816>
- Greenleaf, A. R., Gibson, D. G., Khattar, C., Labrique, A. B., & Pariyo, G. W. (2017). Building the evidence base for remote data collection in low-and middle-income countries: comparing reliability and accuracy across survey modalities. *Journal of medical Internet research*, 19(5).
- Heath, R., Mansuri, G., Sharma, D., Rijkers, B., & Seitz, W. (2016). *Measuring Employment in Developing Countries. Evidence from a Survey Experiment*. World Bank Working Paper.
- Hoogeveen, J., Croke, K., Dabalen, A., Demombynes, G., & Giugale, M. (2014). Collecting High Frequency Panel Data in Africa Using Mobile Phone Interviews. *Canadian Journal of Development Studies* 35(1), 186–207.

- Holbrook, A. L., Green, M. C., & Krosnick, J. A. (2003). Telephone versus face-to-face interviewing of national probability samples with long questionnaires: Comparisons of respondent satisficing and social desirability response bias. *Public opinion quarterly*, 67(1), 79-125.
- ITU (2018). Measuring the Information Society Report. *The United Nations International Telecommunication Union*. Geneva: Switzerland. Accessible at: <https://www.itu.int/en/ITU-D/Statistics/Documents/publications/misr2018/MISR-2018-Vol-1-E.pdf>
- Jones, L. (2018). New methods in resilience measurement: early insights from a mobile phone panel survey in Myanmar using subjective tools. *Overseas Development Institute*.
- Jones, L., Samman, E., & Vinck, P. (2018). Subjective measures of household resilience to climate variability and change. *Ecology and Society*, 23(1).
- Kates, R. W., Travis, W. R., & Wilbanks, T. J. (2012). Transformational adaptation when incremental adaptations to climate change are insufficient. *Proceedings of the National Academy of Sciences*, 109(19), 7156-7161.
- Kristoffersen, I. (2017). The metrics of subjective wellbeing data: An empirical evaluation of the ordinal and cardinal comparability of life satisfaction scores. *Social Indicators Research*, 130(2), 845-865.
- Leo, B., Morello, R., Mellon, J., Peixoto, T., & Davenport, S. (2015). *Do Mobile Phone Surveys Work in Poor Countries?* CGD Working Paper 398. Center for Global Development, Washington, DC.
- Mahfoud, Z., Ghandour, L., Ghandour, B., Mokdad, A., & Sabai, A. (2015) Cell Phone and Face-to-face Interview Responses in Population-based Surveys: How Do They Compare? *Field Methods*, 27(1), 39-54.
- OECD (2013). *OECD guidelines on measuring subjective well-being*. Organisation for Economic Cooperation and Development (OECD). Paris: France
- Pew Research Center (2015). *Cell Phones in Africa: Communication Lifeline*. Available at <https://www.pewresearch.org/global/2015/04/15/cell-phones-in-africa-communication-lifeline/>
- Schipper, Y., Mushi, E., Hoogeveen, J., & Etang, A. (2014). *Response Behavior and Attrition Management in Mobile Phone Panel Surveys*. Unpublished working paper, Twaweza, Dar es Salaam, Tanzania; World Bank, Washington, DC.
- Twaweza (2013). *Sauti za Wananchi. Collecting national data using mobile phones*. Available at <https://twaweza.org/uploads/files/Sauti-za-Wananchi-English.pdf>
- World Food Programme. *mVAM for Nutrition. Kenya Case Study*. Available at https://documents.wfp.org/stellent/groups/public/documents/manual_guide_proced/wfp291938.pdf

World Bank (2020). *High Frequency Mobile Phone Surveys of Households to Assess the Impacts of COVID-19 : Overview (English)*. World Bank, Washington, DC. Available at <http://documents.worldbank.org/curated/en/703571588695361920/Overview>