Using Mobile Phones to Collect Panel Data in Developing Countries*

Brian Dillon[1]
Cornell University
October 2010

Abstract. The rapid spread of mobile telephony throughout the developing world offers researchers a new and exciting means of data collection. This paper describes and analyzes the experience of a research project that used mobile phones to collect high frequency, quantitative economic data from households in rural Tanzania. I discuss the research design, highlight some of the mistakes made and lessons learned, and speculate on the applicability of this method in other settings. (*JEL* O13, Q11, B41)

* I am grateful to Kathleen Beegle, Chris Barrett, Joachim de Weerdt, Hans Hoogeveen, Alessandro Romeo, Diane Steele, Mark Davies and Travis Lybbert for their helpful comments. Staff at Economic Development Initiatives in Bukoba, Tanzania provided invaluable support. The research described in this paper was supported with a Doctoral Dissertation Improvement Grant from the National Science Foundation (SES-0921833), as well as collaborative funding from the World Bank LSMS-ISA program in the Research Group, a Chester O. McCorkle Fellowship from the Agricultural and Applied Economics Association, and grants from the Mario Einaudi Center for International Studies, the Institute for the Social Sciences, and the College of Arts and Sciences at Cornell University. References to "we" in the paper refer to the author and members of the research team, particularly Diego Shirima, Geofrey Mwemezi and Msafiri Msedi. All errors are the sole responsibility of the author.
[1] Cornell University, Department of Economics, Ithaca, NY, 14853, USA; bmd28@cornell.edu; +1607.255.4254

I. Introduction

This paper describes the experience of a study entitled Research on Expectations in Agricultural Production (REAP), a survey conducted in rural areas of western Tanzania from July 2009-September 2010. The primary aim of REAP was to gather quantitative data on the evolution and effect of the subjective expectations that farmers hold over uncertain future outcomes, such as weather, pest intensity and crop yields. Such a project called for high frequency data collection. Instead of embedding enumerators in survey villages for an extended period of time, the REAP team used mobile phones to collect detailed agricultural, economic and demographic data from rural households on a high frequency basis.

The aim of this paper is to describe the mobile phone-based research design and highlight the lessons learned from REAP. Insights presented here can hardly be called "best practices", as they are based on the experience of only one project. Nevertheless, it is hoped that this paper will help others avoid some of the challenges that the REAP team has encountered during the planning and execution of a phone-based survey in a remote setting. The paper is organized as follows: in Section II I describe the project, in Section III I analyze the method's strengths and weaknesses, and in Section IV I conclude.

II. REAP Project Description

During preliminary visits to the study area in 2008 and July 2009, the research team carried phones from each major mobile network in Tanzania, and carefully noted signal availability. We found that one of the network signals was widely available throughout the study area. This did not guarantee network access in every sample village, but it gave us reason for optimism. When the survey began, we were fortunate to find a signal in at least part of every sample village. The network did not reach some respondents' homes, but all households were within a few minutes' walk from a signal.

Our sample consisted of 300 cotton farmers in 15 villages. During initial village meetings, we explained the project and provided phone-related training. REAP team members emphasized that the phones were not gifts, but were research tools that would be left in respondents' safekeeping. Participants were told that they could use the phones for personal use, and that they could keep the phones once the project was complete. From among the 20 sample farmers in each village, 13 were chosen to participate in the phone survey, for a total of 195 phone survey participants.[2] To reinforce the notion of random selection, we fully involved the farmers in this stage of selection, by inviting them to draw names from a hat. Prior ownership of a mobile phone did not exclude participants from receiving a project phone.

Phones were distributed on a later day, in the households, after completion of the baseline interview. Respondents also received laminated sheets with the village-specific call schedule

[2] The seven non-phone households constitute a control group for another study.

and contact numbers for the research team. With 15 villages receiving calls on a Monday-Friday schedule, each village had a calling day once every three weeks.

While none of our sample villages was on the electric grid, some source of power was available everywhere, be it a generator at the school, a house with a solar panel or an individual with a small collection of car batteries. The owners of these power sources operated them as businesses, collecting a small fee to charge a phone. We signed a contract with a "charging station" in each village, paying for survey participants to receive one free charge during the two days prior to each scheduled call.

From September 2009-July 2010, enumerators called respondents on the prearranged days. We made use of a special block price on within-network calls, paying about $1.50 per phone for four hours of calls. On most days, one calling block per enumerator was sufficient to complete all interviews. Interview time ranged from 10 minutes to just over an hour, depending on the length of the round-specific questionnaires and the answers given. Average interview time across the 14 rounds of the survey was 27 minutes. Questionnaires included pre-coded, quantitative questions on subjective expectations, labor on- and off-farm, crop sales, livestock sales and purchases, cultivation of cotton and other crops, changes to household composition, health shocks, expenditure on school fees, land holdings, weather, pest intensity, availability of inputs, phone usage, prices, and sources of information. Some of these data were gathered every 3-6 weeks, others less frequently.

Most phone companies will cancel a SIM card if no pre-paid credit is assigned to it for a period of months. Many REAP respondents were unlikely to purchase phone credit on a regular basis, if at all. Both to prevent the cancellation of project SIM cards and to compensate respondents for participating, we transferred 1,000 shillings (about $0.76) of credit to each phone after each completed interview. The ability to make such transfers is standard in most countries.

We reached an average of eight respondents on the scheduled day. A host of small obstacles prevented interviews from taking place as scheduled, such as illness, family events, network outages, and phone problems. Despite these challenges, virtually all respondents who were not interviewed on schedule were interviewed in the ensuing few days. Village leaders, charging station owners and other participants sometimes assisted us by contacting missing respondents and arranging interviews. Respondents who lost their phones, or whose phones were not working properly, were usually able to participate by borrowing the phone of a friend or neighbor.

A few months after completion of the baseline survey, we re-visited the survey villages and held short meetings with respondents and village leaders. We replaced broken phones, faulty batteries and malfunctioning SIM cards, and "topped up" our charging station contracts. These visits allowed us to receive additional feedback from respondents, and to demonstrate our commitment to the project.

III. Challenges, Solutions and Lessons Learned

In this section I analyze the REAP experience, and speculate more generally on the feasibility of phone surveys in developing countries. I divide the discussion into five subsections: Costs; Infrastructure Issues; Selection and Participation; Data Quality; and Replacement of Materials.

1. Costs

Relative to a traditional survey, the cost savings from a phone survey are most substantial if the project calls for the collection of panel data over relatively short time horizons. Some field costs cannot be avoided, as researchers must conduct baseline interviews and distribute phones. This involves many of the same budget items as a traditional survey, with the addition of phone-related costs. However, researchers can reduce the time of the initial visit by enumerating sections of the questionnaire that are not time-sensitive at a later time, over the phone. Such an arrangement can reduce field time by days, weeks or even months.

The phones used in REAP cost about $20 each, and each SIM card cost $0.38. The average cost of each of the 2,677 phone survey interviews was $6.98, including office rental, phone and SIM card purchases, phone charging expenses, air-time, respondent compensation and staff costs. By contrast, the average cost of each of the 195 baseline interviews was approximately $97, including staff costs, vehicle rental, food and accommodation, printing and other supplies. [3] While these estimates are not indicative of the overall cost difference between methods, because they do not assign to the phone survey the costs of the requisite baseline visit, they highlight the key point: once a survey is operational, the marginal cost of gathering additional rounds of data by phone is only a small fraction of what it would be to gather the data face-to-face.

For the discipline as a whole, phone-based enumeration has the potential to make new types of high frequency data collection feasible in a wide variety of settings, without requiring a substantial increase in funding for field surveys. Individual or household data that may be subject to substantial recall bias in a traditional survey can be gathered more accurately from a high frequency survey. The timing of particular events, such as the employment of inputs or the sale of assets, can be elicited with greater precision. And time-varying data on perceptions and expectations can be gathered in a high frequency panel setting. Such data cannot be reliably gathered with a recall survey. Furthermore, phone surveys are extremely cost effective for research questions that require data at levels of aggregation above the household or individual, such as market price data, quantities available at trading lots or auctions, road or weather conditions.

2. Infrastructure Issues

Charging the Project Phones

[3] These cost estimates exclude training, overtime, bonuses, fieldwork permits and some other extra costs, either because such costs were very REAP-specific, or because they applied to both phone and non-phone aspects of the research.

We were fortunate that although none of the REAP villages was connected to the electrical grid, some power source was available in each. Anecdotal evidence suggests that these independent sources of electric power have proliferated alongside mobile phones, in response to the demand for electricity by phone users in rural villages. If true, this bodes well for the feasibility of phone surveys elsewhere. However, a pre-existing source of electricity is not required for participation in a phone survey. If necessary, researchers can establish charging stations specifically to support the research. Large solar panels cost on the order of $200-$500, including installation costs. Alternatively, a number of companies produce small solar chargers for prices as low as $10, which could be distributed to each participating household.

REAP participants reported a substantial number of faulty batteries. We replaced about 10% of the original batteries during follow-up visits. Some battery problems, such as those caused by irregular voltage from the power source, were unavoidable given the available infrastructure. Other problems, however, were due to a lack of proper training. During the baseline visit, we did not advise participants to turn their phones off when battery power is very low, rather than letting the phones die completely. Nor did we instruct respondents to turn their phones off when outside the network. We also found that some charging station owners took advantage of respondent ignorance by unplugging phones once they display "full bars" on screen, even though the battery was only 75-80% charged at this point. During follow-up visits we tried to remedy these shortcomings by providing additional training.

Network Access

The limitations of the mobile network may present the most definitive challenge to the feasibility of phone-based data collection. It is impossible to provide a network signal to villages not covered by the existing mobile infrastructure.[4] Inconsistent mobile network coverage effectively creates a sampling problem by introducing bias at the village selection stage. This is important for many research questions, since network access is likely to be correlated with other important characteristics, such as distance from major towns, road quality, water supply and average wealth.

Researchers who find that network shortcomings preclude sampling from the original population of interest face tough choices about their project. One possibility is to scrap the phone idea altogether and gather data in the traditional fashion. Another is to draw the sample for the baseline survey from all areas of interest, regardless of network coverage, and then continue the phone survey in those villages with network availability, using characteristics observed during the baseline to construct sample weights. Unfortunately, such weights are only useful if the observables used to construct them are not substantially correlated with network access, and such correlations cannot be measured without first committing to this method of data collection. A third possibility is to establish a calling station as close as possible to a sample village. Unfortunately, this will replace one form of selection bias with another, if the capacity to travel to the calling point is correlated with age, disability, gender, domestic responsibilities, employment status, or other variables of interest.

[4] The only alternative would be to provide respondents with satellite phones, which cost $500-$1500 each.

3. Selection and Participation

Sampling

We sampled from a list of cotton farmers that we constructed from the official village registry. Village leaders assisted us by removing individuals who had died or moved away, and adding individuals who had moved into the village or formed new households since the most recent registry update. We found that it was best not to mention the phones until after the sample was drawn, to prevent village leaders from tampering with the list in order to increase the likelihood that they or their friends would receive a phone.

For questions related to poverty and household agriculture production, there are few situations in which researchers could justifiably sample from an available list of mobile phone users. SIM cards are inexpensive and widely available, and many phone users own multiple lines. More importantly, phone ownership is highly non-random, and rarely observed among the very poor. To prevent the introduction of substantial sampling bias, REAP was designed with the intention of providing phones to respondents. Although some respondents owned a mobile phone prior to their selection for the study, we did not make use of these phones for REAP, because we did not want to engender ill-will among those who were asked to use their personal phone. Also, although the wealth effect is small, we wanted to endow all participating households with goods of equal liquidity and market value.

In other situations it may be possible to rely on respondents' personal phones for enumeration. In Tanzania, phone ownership is nearly ubiquitous among traders, transporters, merchants, university students, government workers and urban formal sector workers. Studies that require sampling from these populations may find that phone distribution is unnecessary. However, attrition rates may be higher under such a design, both because phone endowment appears to engender a deep sense of commitment to the project, and because the opportunity cost of frequent survey participation will be higher among members of wealthier, phone-owning subpopulations.

Attrition and Participation

Potential rates of attrition and periodic non-response[5] are particularly high when researchers are out of sight for much of the survey period. We anticipated substantial attrition from REAP, due to lack of interest, network problems, or sales of project phones. However, on this point we were pleasantly surprised. Across all rounds of the survey, an average of 191.2 of the 195 respondents were interviewed each round. Missed interviews were due primarily to temporary circumstances, such as severe illness. Only one respondent completely abandoned the survey.

By chance, certain features of the study helped maintain high participation rates. The REAP sampling frame was explicitly restricted to cotton farmers, and we introduced the project as a study of cotton production. Farmers were excited to see interest in their cultivation of the crop

[5] "Periodic non-response" denotes failure to complete one or more rounds of the survey, while still participating in later rounds.

they call "white gold" in their language. Also, all of the respondents in the REAP sample lived in relatively small, culturally homogenous villages. Villagers were accustomed to cooperation and neighborliness, and thus were very willing to help find missing respondents. Such a high degree of cooperation among survey participants was made possible by clustered sampling. It seems very unlikely that response rates would have been so high if respondents had been selected from a higher level of geographic aggregation.

Compensation

Low rates of attrition and non-response were also due to the direct benefits of participation. Many respondents looked forward to the 1,000 Tanzanian shilling (about $0.76) credit transfer[6] that they received as compensation for each completed interview.[7] Some respondents viewed the free battery charge as an additional form of compensation, rather than a practical means of ensuring participation. There was nothing wrong with this perception, however, we were concerned at the outset that some respondents might try to rush through the interview in order to preserve battery life. Fortunately we saw no evidence of such behavior.

Timing

The phone survey did not begin until the 2-month baseline survey was near completion. This introduced a potentially harmful asymmetry into the experiences of respondents, as some waited many weeks for their first phone call, while others waited only a few days. To mitigate the effect of the delay, while we were still in the field, enumerators called respondents from the first villages, to greet them and to remind them of their first scheduled calling dates. However, we did not prearrange these calls, and many respondents were unreachable. We learned later that this was largely because their phones did not have any power. This method would have been more successful if we had formally scheduled these calling days, and provided free battery charges. Another way to avoid this problem would be to have a team of phone-based enumerators already in place when the baseline visits begin. Such an arrangement, however, requires enough resources to simultaneously manage data collection in the field and over the phone.

While designing REAP, we also had to decide how often to call respondents. Calling very often over an extended period of time is not only annoying, it is also costly. However, calling too infrequently could raise attrition rates, if respondents lose touch with the project. The optimal lag between calls is clearly related to the research content, the length of the project and the length of each interview. Calling many times a week for two or three weeks is not likely to be such an annoyance as calling many times a week for an entire year. Likewise, interviews that last only a few moments will be tolerated more frequently than those that last close to an hour. The

[6] These transfers were in the form of air-time credit, which can be used to make calls or send SMS messages, but may also be sold or transferred to other phones.

[7] We learned that it was best to use different phones for calling and for credit transfers. Otherwise an enumerator had to top up her phone with "transfer" credit, exceeding that needed to purchase a daily call bundle. But should she over-run the bundle, the credit intended for transfers was consumed very rapidly at the out-of-bundle rate.

decision to call once every three weeks was made after considering budget constraints, the expected length of each interview, and the nature of the data.

4. Data Quality

Multiple Languages

It is not uncommon in East Africa for research teams to interview a proportion of respondents in their tribal language, rather than the national language. When necessary, a translator for a face-to-face interview is often selected by the respondent from among his friends and family. In a phone survey, translation can be avoided if an appropriate proportion of the enumerators are fluent in local languages. If a phone enumerator who does not speak a tribal language doubts his ability to communicate effectively with a particular respondent, he or she can transfer the interview to an enumerator who speaks the local language. In practice this was only necessary at the beginning of the REAP phone survey, because during the baseline interview and the first round of REAP calls we identified respondents who were not fluent in Swahili. An enumerator who spoke Kisukuma, the local tribal language, always called these households.

Supervision

The responsibilities of a traditional field survey supervisor generally involve logistics, training and the maintenance of survey quality. In a phone survey, these tasks can usually be accomplished more quickly and at lower expense than in a traditional survey. If interviewers directly enter data into a computer while gathering it over the phone, which is advised for reasons discussed below, questionnaire checking can be automated and performed almost instantly. Supervisors can directly evaluate enumerator performance by listening to the interviews. In our experience, most community leaders have phones, so the supervisor can remain in contact with local leaders throughout the survey period. All of this can be done from one office, rather than throughout the research areas. The end result is that one supervisor in a phone survey can do the work of many in a traditional survey, without incurring *per diem* expenses in the field.

Confidentiality and the Interview Environment

Experienced face-to-face enumerators read the body language and facial expressions of the respondent, to see if he is tired, frustrated, confused or intentionally deceptive. Also, traditional interviews are conducted in private, to protect the confidentiality of the data. Unfortunately, phone enumerators cannot observe the respondent during the interview, and they cannot directly ensure confidentiality. This may introduce willful error by a respondent, if the questionnaire content is sensitive. For this reason, researchers studying issues of gender, domestic violence, corruption or other sensitive matters may have difficulty gathering reliable data via phone.

However, to some degree the very nature of a phone interview actually enhances confidentiality. If no one other than the respondent is able to hear the interviewer, and questions require a "yes", "no", or otherwise innocuous response, respondents can participate in the survey without revealing questionnaire content. The one-sided privacy of a phone conversation is likely not sufficient protection for truly sensitive personal data, but for other topics it may be enough.

Additionally, although a phone interviewer surrenders some control of the interview environment, he surrenders it to the respondent. This can raise response rates and improve data quality, since it allows respondents to easily reschedule interviews. Traditional enumerators often spend substantial time walking to respondents' homes. If a respondent is not at home or not in the mood to talk, then a costly re-visit must be scheduled, or the interview must be conducted with an anxious, hurried respondent. These problems are avoidable over the phone.

Data Entry

With regard to data entry, phone surveys seem unambiguously superior to face-to-face paper surveys. Data gathered on paper is transcribed twice: once by the enumerator during the interview, and again by the data entry technician. This creates additional costs, and introduces a delay between data collection and analysis. More importantly, this two-stage transcription of data increases the expected number of errors in the raw data. REAP phone enumerators entered the data directly into a computer during the interview, eliminating the time, expense and potential errors from entry of paper data.[8]

Clarification and Additional Questions

After completion of a traditional field survey, researchers often discover that despite their best efforts, some of the questions were misunderstood by respondents, enumerators or both. Even more frustrating is the realization that the inclusion of one or two additional questions would have allowed researchers to test unanticipated, yet interesting, hypotheses. Both of these setbacks can be avoided in a phone survey, provided that researchers remain actively engaged with the incoming data. Instantaneous data entry allows identification of potential problems, and interesting new questions, in real time. If REAP enumerators discovered mid-interview that they were unsure of the meaning of a question, or did not know how to handle a particular response, they asked for immediate guidance. Sometimes, enumerators called back respondents to clarify a response. If necessary, a clarification question was inserted into the next round.

5. Replacement of Materials

It was inevitable that over time, some of the phones and batteries provided by the project would be damaged or lost. Over the nine and half months of the REAP phone survey, eight percent of respondents reported a lost, damaged or malfunctioning phone. These respondents continued to

[8] The REAP questionnaires were almost exclusively quantitative. Data entry called for input of numeric responses, usually from a menu of pre-coded options, rather than extensive typing of qualitative answers.

participate in the survey, using the phones of their friends or neighbors. During follow-up visits, research team members replaced most lost or damaged materials.

Replacement of survey materials introduces an element of moral hazard, as respondents are more likely to be careless or to sell the phone and claim that it was lost if they believe it will be replaced. Minimizing the expenses induced by this moral hazard, while still maintaining a spirit of good faith between researchers and respondents, was one of the key challenges of REAP. To deter sales of the project phones, we told respondent from the outset that we could exchange malfunctioning phones and batteries for new ones, but we could not replace items that were lost. If a respondent lost the project phone but had another phone, we asked him to continue participating in the survey using his personal phone. If a respondent lost the project phone and did not have another phone, we made a determination on a case-by-case basis. We asked about the availability of other phones in the household, and assessed the likelihood of the respondent's ongoing participation. If he lived very near to other participants, we usually asked him to continue working with the project using his neighbors' phones. However, in a few of these cases we violated our strict policy on replacements, and provided a second phone. We were more likely to replace the phones of those who lived in more isolated areas.

It was clear that some of the "lost" phones were actually sold. From a research perspective this was not problematic, as long as respondents continued to participate in the survey. It is not clear *a priori* whether individuals who own another phone are more or less likely to sell the project phone before the survey is complete. The marginal value of a second phone is very low, suggesting that owners of personal phones would be more likely to sell their project phone. However, phone owners are also wealthier in expectation than those who do not own phones, and thus likely to benefit less from a quick sale of the phone for cash. The net effect of these opposing forces is ambiguous.

IV. Conclusion

On balance, the experience of the REAP study suggests that phone-based enumeration of complex economic surveys in low income countries is not only feasible, but also, under some circumstances, superior to traditional data collection methods. Relative to a traditional survey, the cost savings of a phone survey are substantial, as long as the questions of interest call for high frequency panel data. In addition, the centralized nature of phone-based data collection allows for rapid detection and correction of errors, interactive participation by the primary researchers in real time, and streamlined data entry.

There are situations in which a phone survey is infeasible. Network coverage throughout the study area should be investigated prior to committing to the phone method, so as to prevent the introduction of substantial sampling bias. Elicitation of sensitive data over the phone is unlikely to be successful, as it is impossible for phone enumerators to completely ensure confidentiality. Lastly, it's unlikely that the phone survey method will be cost effective for studies that do not require relatively high frequency enumeration of a single set of respondents.

Perhaps the most exciting aspect of mobile phone-based research is the potential it offers for collecting entirely new types of data sets. Current best practices in questionnaire design and data collection methodology are based on the traditional field survey. With the proliferation of mobile telephony comes the possibility of collecting high frequency panel data at reasonable costs. This should greatly expand the range and number of high frequency panel data sets gathered by development economists, without requiring a massive inflow of new research funding to the discipline.