



Using Mobile Phones to Collect Panel Data

By **Brian Dillon**

In 2009 and 2010, a small research team conducted a study in rural areas of western Tanzania entitled Research on Expectations in Agricultural Production (REAP). The primary aim of REAP was to gather high frequency quantitative data on agricultural inputs and farmers' subjective expectations for uncertain outcomes such as weather, prices, and crop yields. Instead of embedding enumerators in survey villages for an extended period of time, the REAP team used mobile phones to conduct surveys on a high frequency basis.

The REAP Survey

The REAP study involved a baseline in-person survey in the summer of 2009, 14 phone survey rounds over the next 10 months, and a follow-up in-person survey in the summer of 2010. During scoping visits, the team confirmed that a network signal from one of the major Tanzanian carriers was available throughout the study area. The sample consisted of 195 randomly chosen cotton farmers in 15 villages in the Mwanza and Shinyanga regions. We provided a phone to all respondents, regardless of whether they already owned one. Phones were distributed in the households, after completion of the baseline interview. Respondents also received laminated sheets with the village-specific call schedule 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.

In rural areas of developing countries it is unlikely that researchers could justifiably sample from an available list of mobile phone users.

Phone ownership is highly non-random and is rarely observed among the very poor. Therefore, we sampled from a list of cotton farmers in the official village registry. While none of our sample villages was on the electric grid, some source of power was available everywhere. 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. We also transferred credit to each phone after each interview, both to prevent the cancellation of project SIM cards, and to compensate respondents for their time.

Average interview time across the 14 rounds of the survey was 27 minutes. Questionnaires included pre-coded, quantitative questions on subjective expectations, labor, sales and purchases, agricultural inputs, demographic changes, weather, and numerous other topics. Some data were gathered every 3-6 weeks, others less frequently.

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. Community members sometimes assisted us by contacting missing respondents and arranging interviews.

Infrastructure Matters

The limitations of the mobile network may present the most definitive challenge to the feasibility of phone-based data collection. Inconsistent mobile network coverage introduces bias at the village selection stage. This is important for many research questions, since network access is likely to be correlated with distance from major towns, road quality, water

supply, average wealth, and other relevant characteristics. Researchers who find that network shortcomings preclude sampling from the original population of interest may need to reconsider the idea of conducting a phone survey. We were fortunate that while none of the REAP villages were connected to the electrical grid, each village possessed some source of power. However, a pre-existing source of electricity is not requisite for participation in a phone survey; if necessary, researchers could establish charging stations specifically to support the research or distribute small solar chargers (which cost less than \$10) to each participating household.

Reduced Costs

Table 1 shows some of the basic elements of the REAP budget. Once the survey was operational, the marginal cost of gathering data by phone was only a small fraction of what it would have been to conduct every round face-to-face. That said, the cost savings from a phone survey are only substantial if the project calls for the collection of panel data over relatively short time horizons. Furthermore, some field costs cannot be avoided, due to the necessity of conducting baseline interviews and distributing phones.

REAP Survey Costs	
Budget item	Cost (USD)
Price per phone	\$20
SIM card	\$0.38
Credit transfer per survey round	\$0.76
Battery charge per respondent per round	\$0.15
Average cost of phone interview	\$6.98
Average cost of face-to-face interview	\$97

Avoiding Attrition

Attrition is clearly affected by the frequency of interviewing. The optimal lag between calls depends on the research content, the length of the project and the length of each interview, as well as any constraints to the budget. REAP enjoyed surprisingly low rates of attrition: an average of 191.2 out of 195 respondents were interviewed in each round. It helped that participants were very willing to help find missing respondents. Low rates of attrition and non-response were also due to the direct benefits of participation, as many respondents looked forward to the free battery charge and the credit transfer that they received for each interview.

Data Quality

In a phone survey, multiple languages can easily be accommodated by staffing the project with at least one enumerator who speaks each local language. Likewise, if enumerators directly enter data into a computer during the interview, supervisors and researchers can check questionnaires in real time. This allows for both the detection of problems as well as the identification of new questions of interest, while the survey is still in the field. Finally, if the respondent is not able to talk when called, a phone survey can be rescheduled more easily than a face-to-face survey, reducing the necessity of surveying of anxious, hurried respondents.

Some issues such as gender discrimination, domestic violence, or corruption may be difficult to study via phone, because the interviewer cannot ensure confidentiality. On the contrary, the nature of a phone interview may enhance confidentiality, if only the respondent is able to hear the interviewer, and questions require a "yes", "no", or otherwise innocuous response.

Replacing Phones

Over 10 months, about 8% of sample households reported a lost or broken phone. Replacement of survey materials introduced an element of moral hazard, which required careful management. To deter sales of the project phones, we told respondents at baseline that we could exchange malfunctioning phones and batteries for new ones, but we could not replace items that were lost. In a few cases we violated this policy, usually for someone living in an isolated area. Nonetheless, it was clear that some of the "lost" phones were in fact sold. From a research perspective, this was not problematic as long as respondents continued to participate in the survey by using a different phone.

This brief is based on: Dillon, Brian (2010). Using Mobile Phones to Collect Panel Data in Developing Countries. *LSMS-ISA Working Paper*, Washington, D.C.: The World Bank. The full paper is available for download at www.worldbank.org/lsms-isa.

For more information, please visit:

www.worldbank.org/lsms-isa

Or contact:

Brian Dillon, Cornell University

bmd28@cornell.edu



THE WORLD BANK

