

# Large Scale Data Collection in Challenging Environments

## *A Practical Guide*



# Overview

1. Pre-Launch Planning
2. Pre-Launch Preparation
3. Mid-Project Management
4. Finishing Up

# PRE-LAUNCH PLANNING

---

# Defining Project Goals



- How many respondents do you need? Households or individuals?
- Do you need 10 respondents in 100 villages or 50 respondents in 20 villages?
- Higher  $\frac{\text{Respondents}}{\text{Clusters}}$  cheaper and faster but lower power
- Time and money requirements  
=  $f(\text{team size, transport options, survey length, } N, \text{ approvals, IRB, sampling strategy, random shocks})$

# A Rough Project Timeline

W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14	W15	W16
Recruitment of team, begin approvals process.				Training team and piloting survey. Organizing transport and accommodation				Sampling and data collection.							
Arrival		IRB approval of team													

# A Rough Project Timeline

W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14	W15	W16
Recruitment of team, begin approvals process.				Training team and piloting survey. Organizing transport and accommodation				Sampling and data collection.							
Arrival			IRB approval of team												

...assuming training manual printed and written, instrument design complete, first stage of IRB approved, PAP registered, coding of survey into tablets finished, sample selection and purchase of study materials complete.

# Team Structure



- *Manager*: responsible for key decision-making (you)
- *Coordinator*: keeps team running, books accommodation, organizes transport, contacts local counterparts
- *Enumerator*: interviews respondents to collect data, gender balance often necessary
- *Mobilizer*: conducts random sampling and local-level consultation ahead of enumerators
- *Auditor*: randomly selects subset of respondents to re-interview, monitors discrepancies
- *Tracker*: follows up and interviews respondents unavailable on day of surveying

# Team Size



- *Manager*: 1
- *Coordinator*: 1-2
- *Enumerator*: 8-15
- *Mobilizer*: 2-8 (can be enumerators)
- *Auditor*: 2-4 (can be enumerators)
- *Tracker*: 2-4 (can be enumerators)

Example: if each village takes one day and two mobilizers to sample, and each enumerator can interview 4 people in one day, then in a 5 day week a team with 10 enumerator/mobilizers could collect  $\approx 160$  interviews.



# Choosing a Coordinator



- Having a local is essential
- Coordinator can be a good intermediary between you and rest of team
- Need someone who has good contacts throughout project area
- Experience with local authorities (chiefs, etc.) and organizing transport and accommodation very good
- Should be someone you can trust and who will take your side

# Transport and Accommodation



- What time do you need to be in field (do people go farming at 7AM?)?
- How long does it take to get from accommodation to field? This impacts daily rate of data collection
- Can public transport get you there? If not, how much does private hire cost?
- Capacity constraints can impose max. team size
- Are contracts enforceable
- Contingencies: having second options, stockpiling fuel.
- Safety concerns (more below)

# Budgeting Time and Money

## Example

Say we want to survey 1000 people in 20 villages with a team of 10 enumerators. Each village takes one day and 2 people to mobilize and enumerators can move independently.

# Budgeting Time and Money

- A Surveys-per-person-per-DC-day = function of transport time, survey length, resp. availability  $\approx 5$
- B Persons-per-DC-day = function of team size (determined by transport and accommodation constraints)  $\approx 10$
- C Mobilization-days-per-week = function of sampling strategy  $\approx 1$
- D Villages-mobilized-per-week  $\approx 5$
- E Data-collection-days-per-week = function of sampling strategy  $\approx 4$
- F Data-per-week =  $A \times B \times E = 200$
- G Data-collection-weeks =  $1000/F = 5$
- H Working-days =  $G \times 5 = 25$
- I Weekend-and-training-days =  $G \times 2 + 14 = 24$
- J Accomodation-person-days =  $(I + J + 14) \times 10 = 530$
- K Transport-person-days =  $H \times 10 = 250$

# Budgeting Time and Money

## Total time

1. This project requires 53 days for training, data collection and sampling
2. But we must also add on at least 30 days for recruiting enumerators, organizing transport and accommodation, and consulting with authorities to get the necessary approvals
3. So all up the project is likely to take around 3 months (assuming you do all of the instrument design and IRB prior to departure).
4. This does not account for any piloting!

# Budgeting Time and Money

## Total cost

1. Other costs (more on these below): team safety (mosquito nets, umbrellas, motorcycle helmets, painkillers, etc.), tablets and batteries, printing (consent forms, sampling sheets), food (breakfast, lunch, dinner), gifts for local consultation, communication (phone credit, walkie talkies), etc.
2. Total budget = accommodation + transport + team safety + tablets and batteries + printing + food + gifts + communication + your accommodation + international travel + FEX and other banking fees
3. Total budget =  $J \times \text{nightly acc. rate} + K \times 10 \times \text{return-travel-per-person} + \text{team safety costs} \times (10 + 5) + \text{tablets and batteries cost per person} \times (10 + 5) + \text{printing} \times 5 \times 1000 + \text{food} \times 10 \times J + \text{gift} \times (20 + 5) + \text{one person-day communication} \times H + \text{your accommodation} + \text{international travel} + \text{FEX and other banking fees} \times \text{number of transfers}$

# PRE-LAUNCH PREPARATION

---

# Recruiting the Team



- To find applicants: ask previous researchers and NGOs that have commissioned studies
- During interviews: explain project, small test, pay for transport
- Often you will want gender-balanced teams
- Contract: short-term employment or purchase of services?
- Contract should contain: performance expectations (e.g. min number of surveys but careful of incentives), services being provided, whether transport days and weekends are paid for, accommodation and food arrangement, period of tenure, ownership of work, certification, termination procedures



# Paying the Team



- Careful about paying team different amounts
- Can be good to pay in advance, every week or two weeks
- Some team members will not have bank accounts, be prepared to have a lot of cash on hand
- Mobile money can be a good solution
- Bonuses can be a good way to structure incentives

# Training the Team



- Write a manual! (sampling strategy, interview tips, HS protection, project aims and tentative schedule, etc.)
- Training should be paid
- One day to go over plan for project and principles
- One day to go over human subjects protection (IRB will require certification via CITI)
- 1-3 days to go over survey and survey techniques
- 1-3 days for the sampling strategy
- 1-7 days for piloting in non-sample areas

# Managing the Team



- (Paid) days off can be important
- Be careful about putting young women alone with men on missions (social perception)
- Look out for wellbeing: keep team well-fed, buy mosq. nets, make sure they have food options on weekend,
- Spell out at beginning: zero tolerance for drunkenness, lateness and uncivil behavior
- Have alternative accommodation options in case team members argue
- Careful with confrontations

# Local Approvals



- Send formal letters of introduction to national government, district government, and local government at least a month before beginning
- Some places may require documented approval (a signed letter of introduction). Make inquiries.
- Make sure to go there on a research visa
- Mobilization days can be used to consult with chiefs / village-level authorities
- Inquire about what gifts are appropriate and avoid giving money at all costs
- Ask chiefs to do some awareness-raising



- If you are recruiting enumerators, the IRB needs them to complete HS training and sign Individual Investigator Agreements
- This process can only take place after recruitment and can take 2-4 weeks
- Local IRB approval is often necessary and should be initiated early on, prior to CU approval

# MID-PROJECT MANAGEMENT



# Sampling

Household Member Grid #0.

1	2	3	4	5	6	7	8	9	10
1	2	3	1	3	5	3	6	3	4
1	2	3	4	2	1	1	2	3	5
1	1	3	4	2	1	3	2	8	5
1	2	3	2	3	5	7	4	8	4
1	1	3	1	1	4	5	5	9	7
1	2	1	4	1	5	5	4	1	10
1	2	1	4	1	1	3	2	3	8
1	1	1	3	4	2	1	4	9	3
1	2	2	4	1	3	4	8	4	
1	2	3	1	2	3	2	7	2	8
1	1	3	4	4	2	4	5	2	4
1	1	3	1	1	4	7	3	6	10
1	1	1	1	2	6	3	2	3	1
1	1	1	3	5	4	7	8	5	5
1	1	3	3	4	1	2	9	4	5
1	1	1	4	4	2	4	3	1	6
1	1	2	1	2	5	1	4	1	9
1	1	1	4	4	4	1	8	4	6
1	1	3	3	5	5	2	2	8	10
1	1	3	4	2	3	4	2	4	1
1	2	2	2	4	4	2	4	5	10
1	1	3	3	1	3	1	4	3	5
1	1	3	2	2	5	6	4	2	10
1	2	3	2	1	4	4	8	2	7
1	1	1	4	4	4	1	8	4	6
1	2	2	3	3	5	4	6	5	9
1	2	3	4	4	4	2	2	4	1
1	2	3	2	4	6	3	1	9	10
1	2	2	3	1	2	3	1	6	8
1	1	2	3	4	4	4	2	5	1
1	2	2	3	5	5	2	3	1	5
1	2	2	1	4	6	1	4	8	8
1	2	2	1	5	6	3	8	9	6
1	1	1	4	3	3	8	2	3	8
1	1	3	4	5	4	5	6	6	10
1	1	1	3	3	1	2	8	9	10
1	1	2	3	2	6	5	6	5	5
1	1	2	4	4	1	7	2	1	5
1	1	3	2	1	6	3	6	7	2
1	2	3	2	2	1	6	9	3	8
1	1	2	2	4	2	6	3	1	2
1	2	3	2	3	1	2	7	1	9
1	2	3	2	3	5	7	9	6	1
1	1	1	2	5	2	3	5	2	1
1	2	2	1	5	2	1	2	3	5
1	2	2	4	5	4	4	4	6	5
1	1	2	3	1	3	6	7	7	7
1	1	2	3	2	3	3	2	3	2
1	1	1	1	4	2	5	4	9	1
1	1	3	3	1	1	6	2	7	5
1	2	2	4	2	1	4	5	4	1
1	1	2	4	5	6	3	8	4	6
1	1	3	1	1	2	1	6	9	2
1	2	2	4	1	1	4	2	2	8
1	2	4	3	6	1	2	2	4	1
1	1	1	3	5	4	3	3	1	10
1	2	1	1	1	3	3	3	6	2
1	1	3	4	4	2	5	8	4	7
1	1	2	1	1	5	5	4	4	1
1	2	1	2	5	5	4	5	9	8

- Sampling can be done on paper or using tablets
- 2-stage random sampling requires you have a scheme for randomly sampling households, then randomly sampling members of those households
- One HH sampling technique: write down all names of all household heads in village with chief, reorder randomly and select first 40
- One R | HH sampling technique: if x number of eligible respondents (age and gender), choose  $z \in 1, \dots, x$ 'th eldest at random

# Sampling



- Sampling can be done on paper or using tablets
- 2-stage random sampling requires you have a scheme for randomly sampling households, then randomly sampling members of those households
- One HH sampling technique: write down all names of all household heads in village with chief, reorder randomly and select first 40
- One  $R|HH$  sampling technique: if  $x$  number of eligible respondents (age and gender), choose  $z \in 1, \dots, x$ 'th eldest at random



# Data Collection



- Tablets are far better than paper - can be backed up to cloud, not prone to transcription error, lightweight, dynamic programming
- Open Data Kit (ODK) is a great solution
- Try to have survey programmed before field
- Redundancies (extra tablets, power packs)
- Print consent forms at beginning

# Spending and Monitoring Money

ORIGINAL RECEIPT 661226

DATE 26/09/2016

RECEIVED FROM JASPER LOOPER

THE SUM OF ONE HUNDRED AND SEVENTY KINA ONLY

BEING FOR TAXI AROUND AREVA FOR 3 DAYS

PER [signature]

K 170 : 00

DUTY [signature]

- Have a spreadsheet backed up to Dropbox with all expenditures
- Make sure to re-estimate budget as you go along, using counts of what is left and projections for what you need
- Have receipt books in the field to write down expenses
- Inquire about transfers, include FE in budget
- Have police escort and safe for cash, ask to use safe room

# Quality Assurance



- Enumerators shirk: you can have one enumerator randomly select a respondent from the sampling list, and go to re-interview them on some small subset of questions
- Circular translation of instruments can be important
- Tradeoffs of being personally present: disturbance vs. monitoring and helping
- Piloting is crucial

# Risk Management



- Consultation is very important to ensure good reception by community
- Never be without comms: walkie talkies, cellphone credit, chargers, etc.
- During mobilization ask about events (funerals, religious ceremonies, etc.)
- Spend more to travel safely
- Don't risk others' lives for data

# Risk Management



- Consultation is very important to ensure good reception by community
- Never be without comms: walkie talkies, cellphone credit, chargers, etc.
- During mobilization ask about events (funerals, religious ceremonies, etc.)
- Spend more to travel safely
- Don't risk others' lives for data

## FINISHING UP

---

# Closing Ceremony



- Ending well is important: you may want to do future projects with same enumerators or in the same villages
- Giving certificates at a closing ceremony goes a long way (laminate!)
- Letters of reference
- Reports for local authorities



THE  
END