CATI and SMS Surveys in Sub-Saharan Africa

John Murunga

Charles Lau

Outline

1. Introduction to CATI and SMS (~5 minutes)

2. Use cases for CATI and SMS (~10 minutes)

3. Challenges and Solutions: CATI (~20 minutes)

4. Challenges and Solutions: SMS (~10 minutes)

5. Discussion (~45 minutes)

GeoPoll

GeoPoll

Overview

We are a technology-driven research organisation with a mission to collect high-quality, timely data at scale. We conduct research using CATI, face-to-face, web, and SMS using our purpose-built platform. GeoPoll is headquartered in the USA, but our largest office is Nairobi. Most staff work in low- and middle-income countries.

Key Clients

GeoPoll: By the Numbers

- 4.5 million interviews annually
- Projects in 144 countries
- 8,414 distinct CATI projects
- Network of global call centres
- Surveys in 113 languages
- Over 10,000 trained interviewers
- Database of 300+ million mobile subscribers

Illustrative Projects

- *People's Climate Poll (UNDP):* 74,000 interviews in 77 countries
- *Food Security Monitoring (WFP):* 1.9 million CATI interviews in 44 countries
- *Public Service Delivery Index (AFDB):* Face-to-face surveys in 27 African countries in 4 months

GEOPOLL CORE VALUES   Respect | Honesty | Integrity | Problem Solvers | Team Players
Customer Oriented | Continuous Improvement

1. Introduction to CATI and SMS

Overview of CATI

How it Works

- Interviewer-administered phone surveys
- Permanent and temporary call centers
- Similar QC procedures as in high-income countries
- Questionnaire length: Optimal length is unclear, guidance ranges from 15-30 minutes.
- Virtual and in-person call centers
- Mobile phones only (no landlines)

Advantages

- ✓ Doesn't require literacy (compared to SMS)
- ✓ Makes certain question types possible –e.g., collecting village name can be challenging in self-administered modes
- ✓ Interviewers can tailor survey request and set appointments
- ✓ Interviewers can probe and clarify
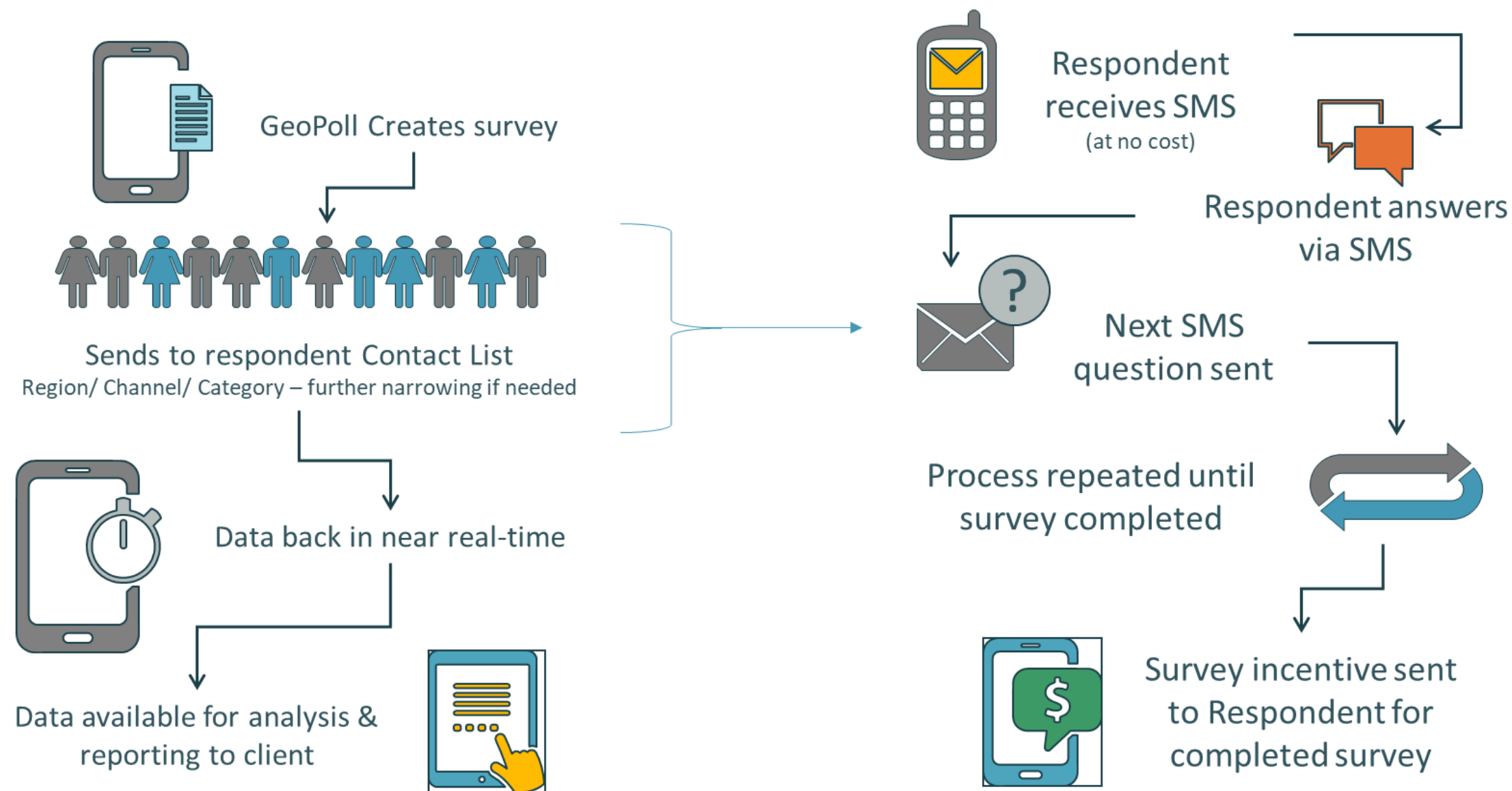- ✓ Longer questionnaires compared to other modes

Disadvantages

- ✗ Can be more expensive than IVR/SMS/web. BUT this depends on many factors: number of questions, number of interviews, whether professional or temporary call center used
- ✗ Interviewer effects
- ✗ Worse for sensitive topics (??? – more research needed)
- ✗ Synchronous mode: requires respondent to pick up
- ✗ No visual content

GeoPoll

HOW IT WORKS

GeoPoll Creates survey

Sends to respondent Contact List
Region/ Channel/ Category – further narrowing if needed

Data back in near real-time

Data available for analysis & reporting to client

Respondent receives SMS (at no cost)

Respondent answers via SMS

Next SMS question sent

Process repeated until survey completed

Survey incentive sent to Respondent for completed survey

- Short message service = SMS, or text messages.

- Respondents receive text messages, one per survey question.

- Texting available on all phone types.

- This is self-administered **automated** mode: there aren't interviewers on other end (except in rare cases – e.g., West, Ghimire, Axinn, 2015).

- Each question has 160-character limit for question and response options (including spaces).

- Multiple messages possible on some tools, but these can lead to errors.

- Character limit depends on country and mobile network operator.

- Sometimes special characters (é) can count double towards character limit (check the survey tool).

- ~20 questions.

- Language selector at beginning of survey for multi-lingual surveys.

GeoPoll

Advantages

✓ Extremely rapid deployment

✓ Extremely rapid data collection

✓ Asynchronous: people can respond at leisure

✓ Natural form of communication

✓ Less expensive than IVR

✓ Suitable for "momentary assessments" (real-time data)

✓ Potentially better for sensitive topics

Disadvantages

X Less familiar mode to respondents

X Requires literacy

X Questions must be short, problems splitting up messages

X Questionnaires must be short (~10-20 questions)

X No visual content

X Small screen size often requires scrolling

X Responses remain on phones (sensitive topics)

X Harder to verify who is responding

GeoPoll

Sampling Approaches

	Advantage	Disadvantage
Random digit dial (RDD)	✓ Large sample size ✓ Easy and inexpensive	✗ Inefficient and varies by mobile operator ✗ Numbering system typically lacks geography ✗ Designs typically limited to stratification by operator ✗ Lower response rates (vs. re-contact)

Sampling Approaches

	Advantage	Disadvantage
Random digit dial (RDD)	✓ Large sample size ✓ Easy and inexpensive	✗ Inefficient and varies by mobile operator ✗ Numbering system typically lacks geography ✗ Designs typically limited to stratification by operator ✗ Lower response rates (vs. re-contact)
Re-contact from previous survey	✓ Higher response rates than RDD ✓ Lots of data for non-response adjustments	✗ Limited sample size ✗ Many people switch phone numbers (after 6 months, 43% of numbers in Liberia switched off)

Sampling Approaches

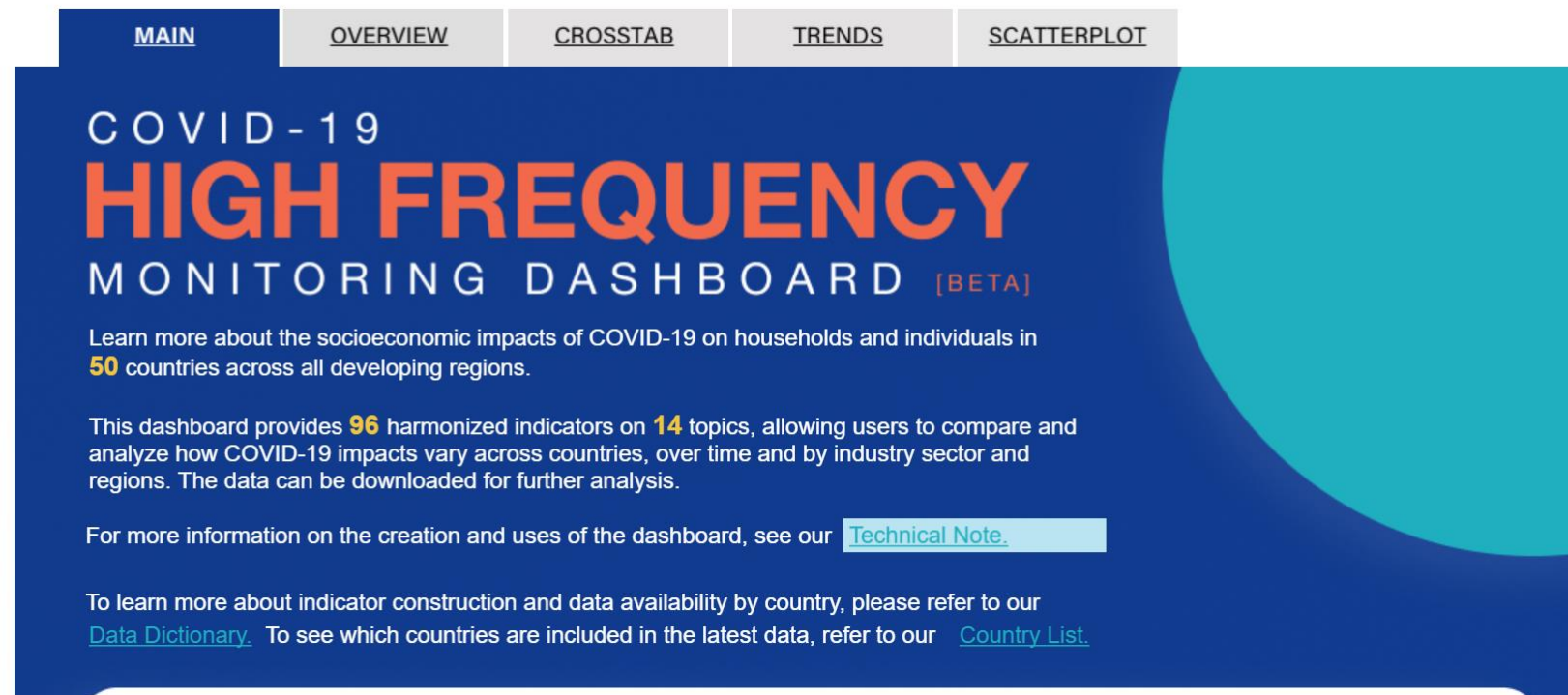|	Advantage	Disadvantage
Random digit dial (RDD)	✓ Large sample size ✓ Easy and inexpensive	✗ Inefficient and varies by mobile operator ✗ Numbering system typically lacks geography ✗ Designs typically limited to stratification by operator ✗ Lower response rates (vs. re-contact)
Re-contact from previous survey	✓ Higher response rates than RDD ✓ Lots of data for non-response adjustments	✗ Limited sample size ✗ Many people switch phone numbers (after 6 months, 43% of numbers in Liberia switched off)
List sample from vendor	✓ Rich auxiliary information for sample targeting and non-response adjustment	✗ More expensive ✗ Non-probability ✗ Panel conditioning effects ✗ Potential for coverage error when lists exclude certain mobile network operators

Sampling Approaches

	Advantage	Disadvantage
Random digit dial (RDD)	✓ Large sample size ✓ Easy and inexpensive	✗ Inefficient and varies by mobile operator ✗ Numbering system typically lacks geography ✗ Designs typically limited to stratification by operator ✗ Lower response rates (vs. re-contact)
Re-contact from previous survey	✓ Higher response rates than RDD ✓ Lots of data for non-response adjustments	✗ Limited sample size ✗ Many people switch phone numbers (after 6 months, 43% of numbers in Liberia switched off)
List sample from vendor	✓ Rich auxiliary information for sample targeting and non-response adjustment	✗ More expensive ✗ Non-probability ✗ Panel conditioning effects ✗ Potential for coverage error when lists exclude certain mobile network operators
Mobile operator sample	✓ Large sample size ✓ Potential for auxiliary information	✗ Hard to get for most researchers

2. Use Cases for CATI and SMS

World Bank High Frequency Phone Surveys

Relevant links
- [Dashboard](#)
- [Data](#)
- [Review paper](#) "Viewpoint: High-frequency phone surveys on COVID-19: Good practices, open questions"
- [Upcoming conference](#): December 2024

- During the COVID-19 pandemic, the World Bank conducted CATI surveys in 89 countries

- Surveys used two main sampling frames:
  - "Follow-back" surveys, where respondents from previous face-to-face surveys were recontacted
  - RDD

- Many countries use longitudinal CATI designs

- Core questionnaire and country-specific modules: harmonized data publicly available

World Food Programme: Food Security

- The Hunger Monitoring Unit of the UN World Food Programme uses CATI (among other data) to develop near real-time estimates of food security

- Sampling: RDD, partner databases, beneficiary lists

- Questions on food security (e.g., Food Consumption Score)

- Data visualized on the Hunger Map LIVE

Link to Approach and Methodology

UNICEF MICS Plus

- "Follow back" surveys to previous FTF survey respondents

- Belize, Georgia, Mongolia, Samoa, Nigeria, Jamaica

- Fills data gaps between time consuming and expensive FTF surveys (~5+ years between MICS waves)

- Implementation Guide

- UNICEF MICS Plus

Nigeria 2023-2024
WAVE 1 (February-May 2023)

MICS PLUS · nbs · unicef for every child

The Nigeria National Bureau of Statistics (NBS), in collaboration with UNICEF, began the implementation of 2023-2024 MICS Plus in 2023. MICS Plus generates data to understand the situation of households, families, and children to inform decision-makers and stakeholders accordingly. This statistical snapshot summarises the results of the first wave of calls from 10th February to 2nd May 2023.

Survey Implementation

Response Rates

Wave 1 (February 10 – May 2, 2023)	Before Substitution	After Substitution
Households sampled	3,699	3,699
Interviewed	1,581	2,226
Refused	78	11
No eligible respondent	1	0
Phone number(s) does not belong to sampled household	68	18
Phone number(s) inactive	43	21
Respondent busy/postponed	3	1
No response after repeated call attempts or phone(s) turned off	784	554
Phone number(s) not connecting	80	29
No phone number available for sampled household	1061	839
Response rate (percent)	42.7	60.2

Interview duration
Median duration of interviews by degree of urbanisation

	Total	City	Town & Semi-dense area	Rural area
Minutes	15	15	16	16
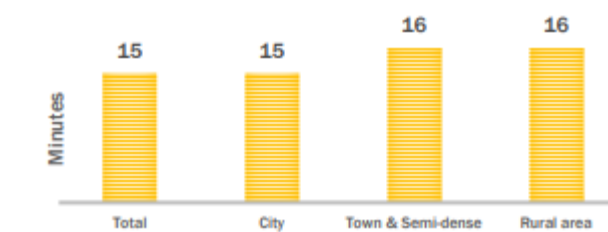
Average call attempts
Average number of call attempts for completed interviews, by degree of urbanisation

	Total	City	Town & Semi-dense area	Rural area
Number of calls	2.5	2.1	2.5	3.2

GeoPoll

# Public Opinions and Attitudes to Abortion in Kenya

Methods:

- **Study design**: Nationwide cross-sectional survey using CATI/Telephone
- **Target population**: all adults ≥18 years across all Kenyan counties **Sampling and sample size**: 8,942 people from a GeoPoll database of >8million telephone numbers
- **Data collection:** Telephone interviews using structured questionnaires having 4 components
- **Period:** September-November 2022

Objectives

1) Assess the overall opinions of abortion in Kenya
2) Assess the knowledge and attitudes towards abortion and abortion laws in Kenya
3) Measure the extent of abortion stigma in Kenya
4) Explore the circumstances under which the Kenya public would view abortion as acceptable and allowable by law
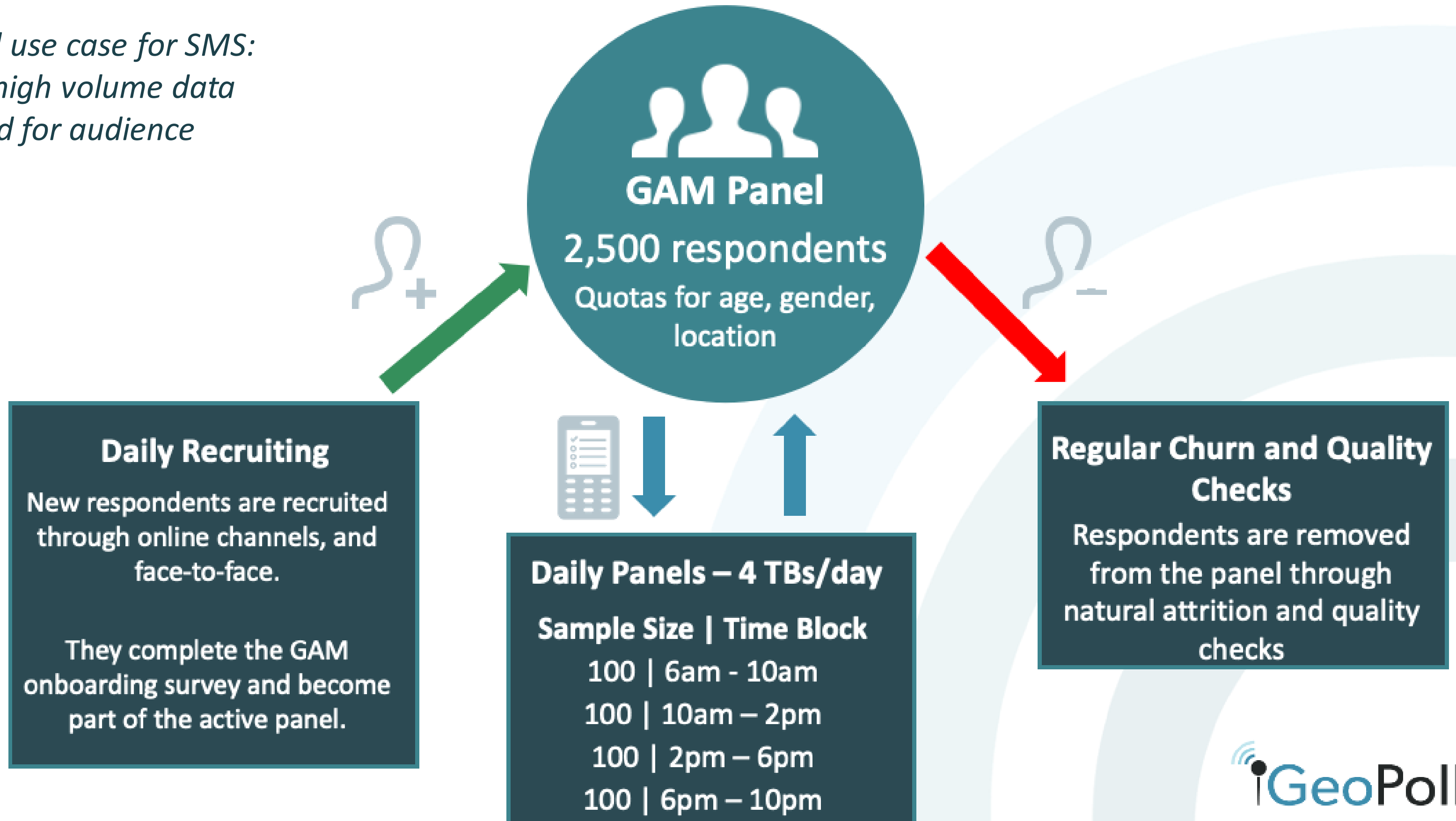
Tools/Questionnaires
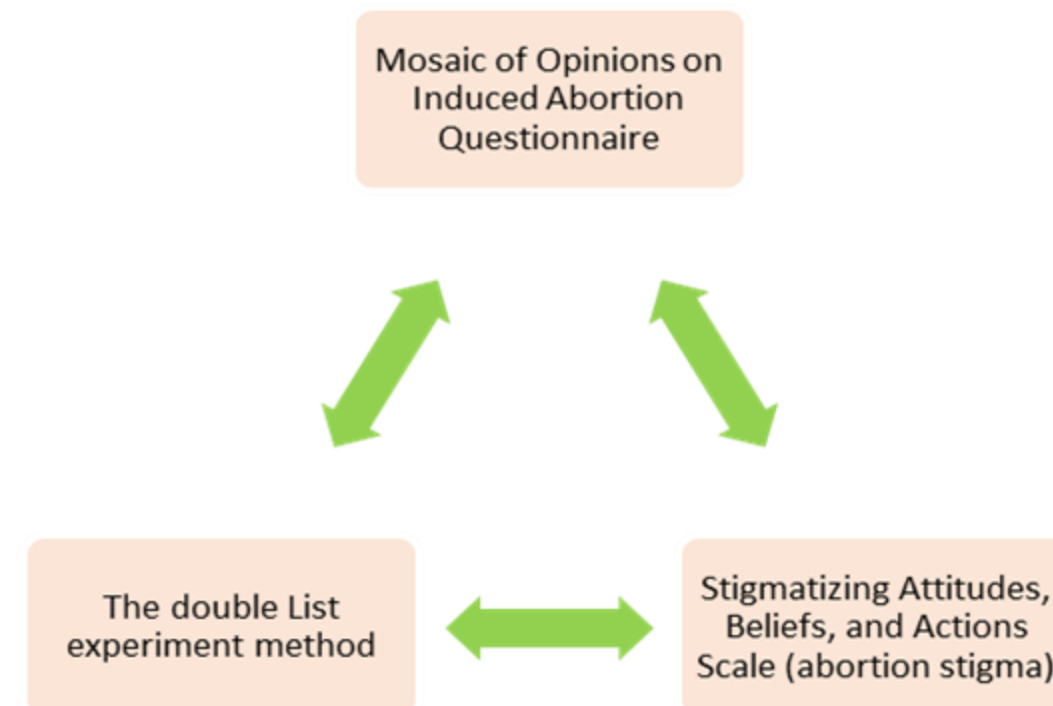
Respondent characteristics

	Characteristic	Frequency	Percent
Age group	18-24	2,219	24.8%
	25-34	2,640	29.5%
	35+	4,083	45.7%
Sex	Male	4,464	49.9%
	Female	4,478	50.1%
Highest education level	None/DK	228	2.6%
	Primary	1,481	16.6%
	Secondary	3,348	37.5%
	University	3,882	43.4%
Marital status	Partnered	4,790	53.6%
	Unpartnered	4,149	46.4%
Religion	Christian	8,303	92.9%
	Others	636	7.1%

GeoPoll

Media Audience Measurement Using SMS

Example of good use case for SMS: high frequency, high volume data collection needed for audience measurement

GAM Panel

2,500 respondents

Quotas for age, gender, location

Daily Recruiting

New respondents are recruited through online channels, and face-to-face.

They complete the GAM onboarding survey and become part of the active panel.

Daily Panels – 4 TBs/day

Sample Size | Time Block

100 | 6am - 10am

100 | 10am – 2pm

100 | 2pm – 6pm

100 | 6pm – 10pm

Regular Churn and Quality Checks

Respondents are removed from the panel through natural attrition and quality checks

iGeoPoll

SMS for Public Health Surveillance

- Example from the H1N1 pandemic in Mexico

- Using SMS for "syndromic surveillance" (monitoring symptoms of disease in a population)

- May 2009: 982,708 numbers received SMS from Ministry of Health

- Received responses 56,551 unique numbers

- 53% of responses received with 24 hours, 89% within 24 hours

[Source](#)

CDC Centers for Disease Control and Prevention
CDC 24/7: Saving Lives, Protecting People™

Search

EMERGING INFECTIOUS DISEASES®

ISSN: 10

EID Journal > Volume 16 > Number 9—September 2010 > Main Article

Volume 16, Number 9—September 2010

Letter

Mobile Messaging as Surveillance Tool during Pandemic (H1N1) 2009, Mexico

Cite This Article

GeoPoll

3. Challenges and Solutions: CATI

CATI: Challenges and Solutions

1. **Representation**: How well does CATI approximate the demographic distributions of target populations?

2. **Questionnaire Length:** How long can (or should) CATI questionnaires be?

3. **Measurement**: How reliable are respondent's self-reports in CATI surveys?
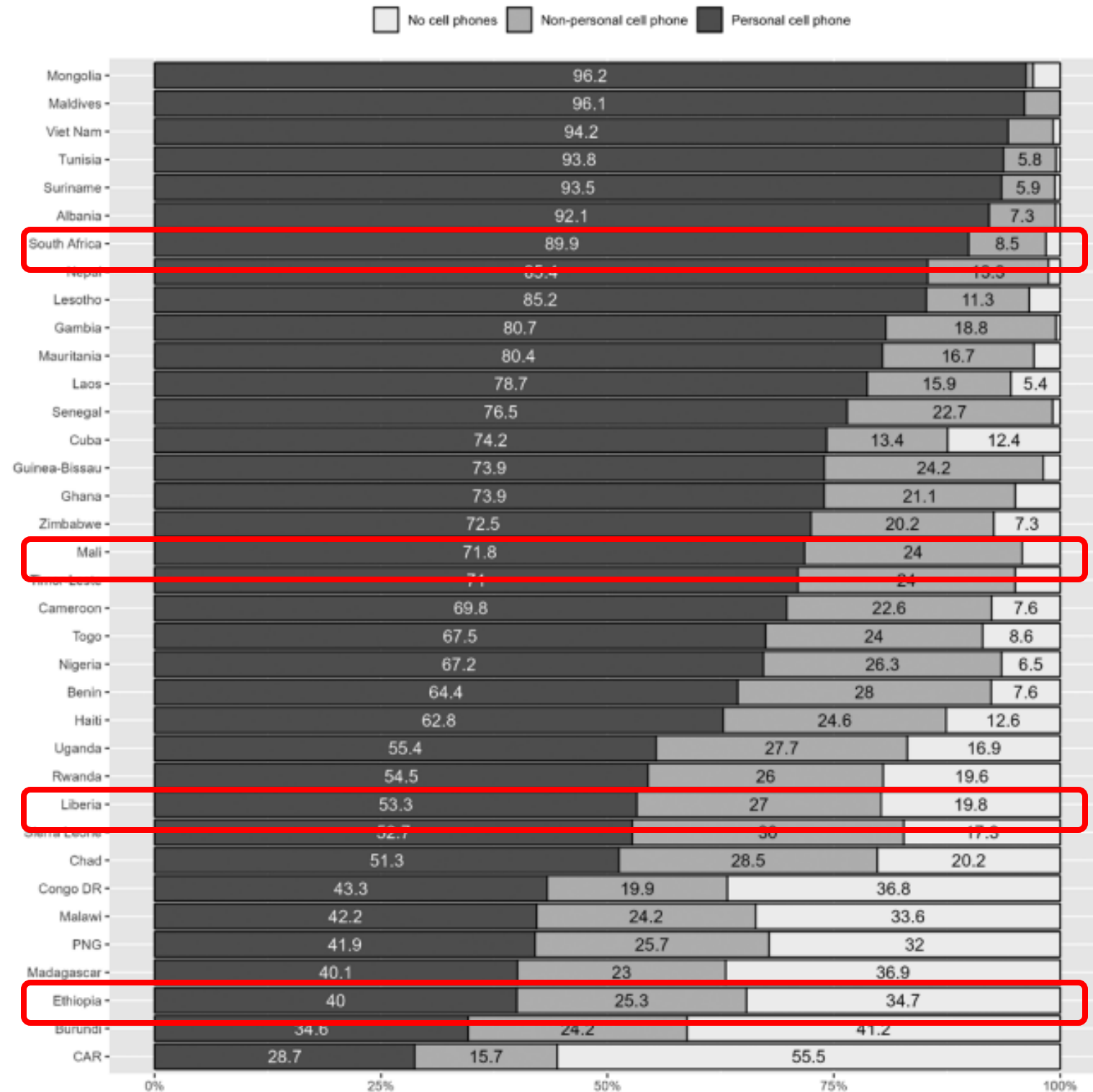
GeoPoll

Challenges in Representation: Undercoverage

Undercoverage

☐ No cell phones ☐ Non-personal cell phone ■ Personal cell phone

Country	Personal	Non-personal	No cell phones
Mongolia	96.2		
Maldives	96.1		
Viet Nam	94.2		
Tunisia	93.8	5.8	
Suriname	93.5	5.9	
Albania	92.1	7.3	
South Africa	89.9	8.5	
Nepal	85.4		
Lesotho	85.2	11.3	
Gambia	80.7	18.8	
Mauritania	80.4	16.7	
Laos	78.7	15.9	5.4
Senegal	76.5	22.7	
Cuba	74.2	13.4	12.4
Guinea-Bissau	73.9	24.2	
Ghana	73.9	21.1	
Zimbabwe	72.5	20.2	7.3
Mali	71.8	24	
Timor-Leste	71	24	
Cameroon	69.8	22.6	7.6
Togo	67.5	24	8.6
Nigeria	67.2	26.3	6.5
Benin	64.4	28	7.6
Haiti	62.8	24.6	12.6
Uganda	55.4	27.7	16.9
Rwanda	54.5	26	19.6
Liberia	53.3	27	19.8
Sierra Leone	52.7	30	17.3
Chad	51.3	28.5	20.2
Congo DR	43.3	19.9	36.8
Malawi	42.2	24.2	33.6
PNG	41.9	25.7	32
Madagascar	40.1	23	36.9
Ethiopia	40	25.3	34.7
Burundi	34.6	24.2	41.2
CAR	28.7	15.7	55.5

Figure 2. Mobile phone coverage of adults aged 15–49 years by country.

Source: Elkasabi & Khan (2023)

## Complexities in Measuring Coverage in Africa

1. Coverage statistics (number of subscriptions per 100 people) based on **ownership** may overstate coverage errors because of sharing
2. But family members don't share phones equally
3. Some people who share phones may not accept incoming calls

GeoPoll
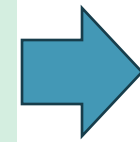
Challenges in Representation: Nonresponse

Noncontact

- Poor network connection
- People keep phones off
- Miss calls
- Differential access to phone for phone sharers
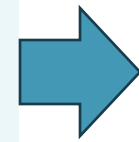
Language Barrier

- Language not offered in survey
- Difficulty handing off to another interviewer
- Language switching makes interview too challenging, especially after pleasantries

Refusal

- Distrust of surveys
- Lack of familiarity with surveys
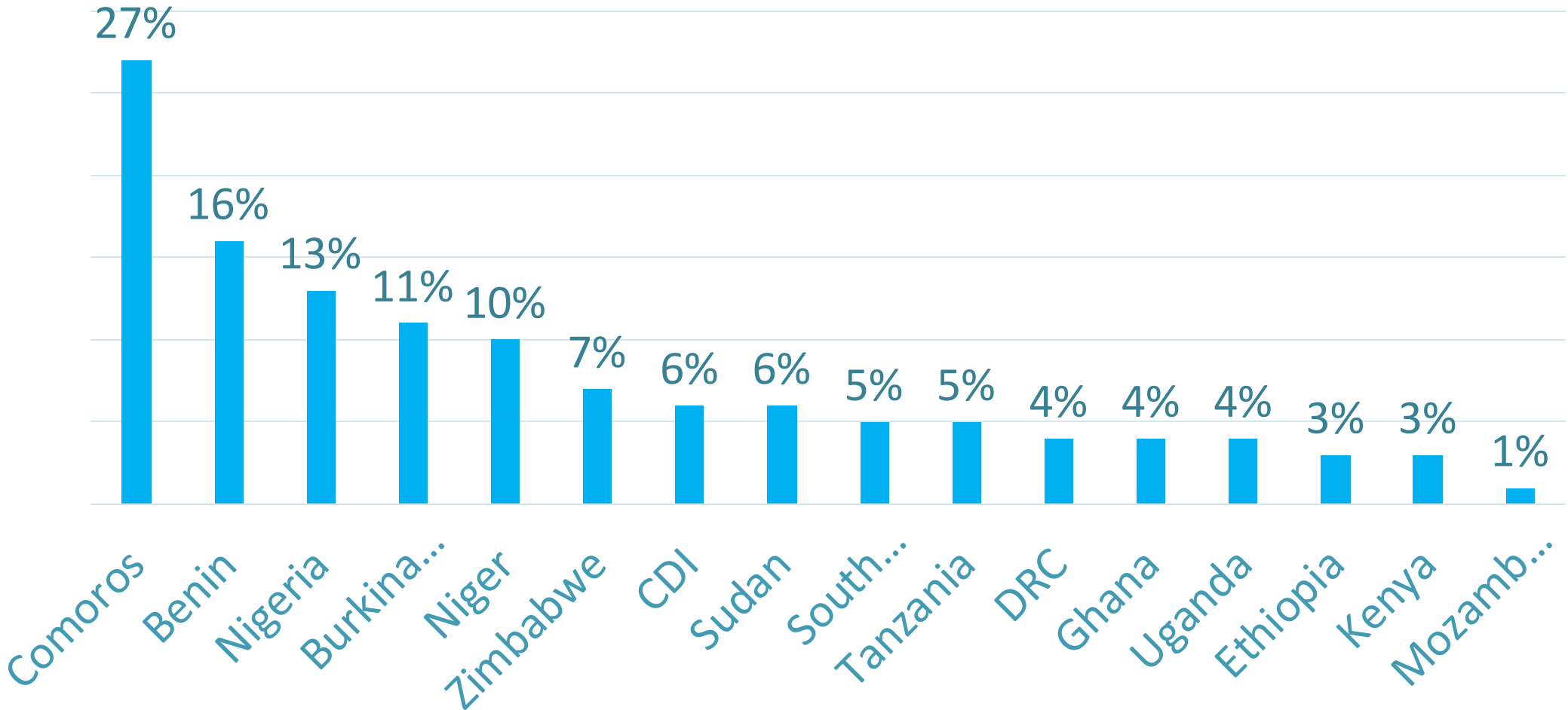- Competing priorities or responsibilities ("busy")
- Concerns about data charges
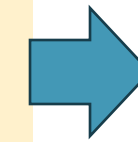
Breakoff

- Battery runs out
- Airtime or tech issue
- Phone network issue
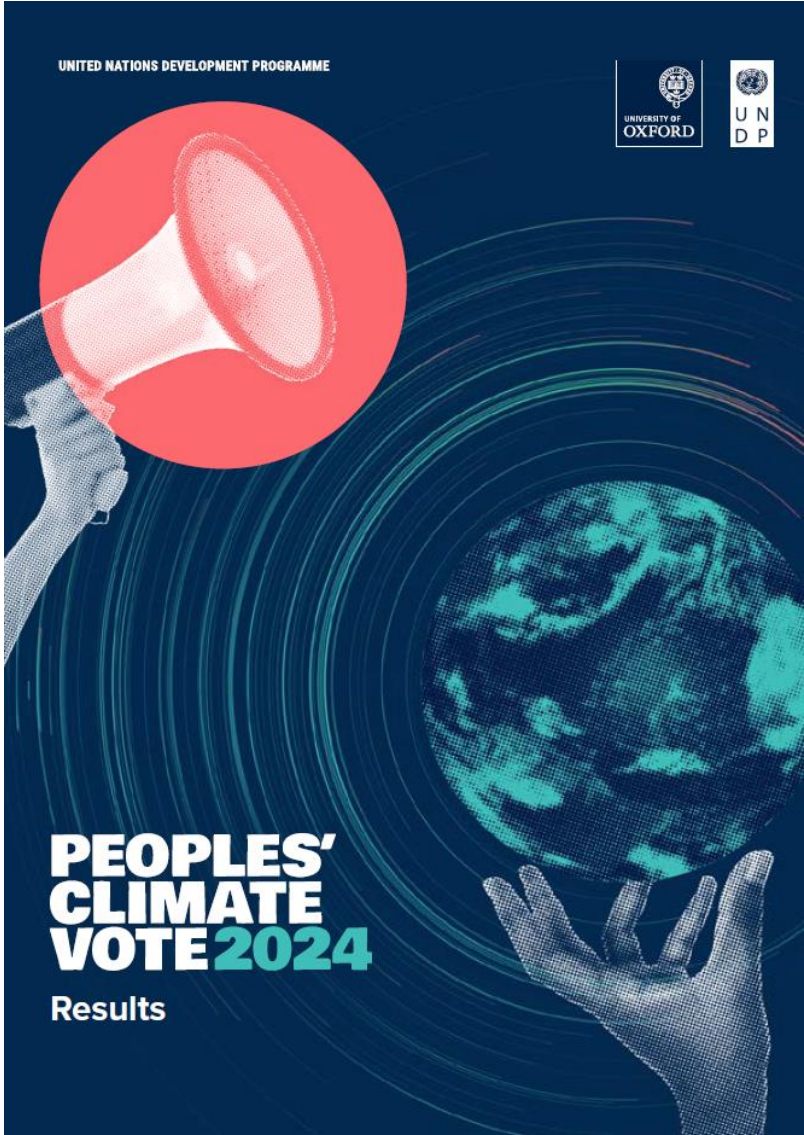- Long surveys
- Complex questions
- Sensitive questions

GeoPoll

Response Rates: Example

Response Rate (AAPOR RR1)

Comoros 27%, Benin 16%, Nigeria 13%, Burkina... 11%, Niger 10%, Zimbabwe 7%, CDI 6%, Sudan 6%, South... 5%, Tanzania 5%, DRC 4%, Ghana 4%, Uganda 4%, Ethiopia 3%, Kenya 3%, Mozamb... 1%

[Source](#)

People's Climate Vote: Cross-sectional CATI RDD and List-Assisted Sampling

GeoPoll

Design of Lau et al. (2019)

- Nigeria cross-sectional RDD CATI study in 2016

- 15% Response Rate (AAPOR RR3)

- 86 questions, focus on technology

- Simple design, no quotas or weighting

Representation

- Women underrepresented by 20 percentage points (pp)

- Age 50-64 underrepresented by 10 pp under. Age 40-49 underrepresented by 6 pp.

- People without formal schooling underrepresented by 34 pp

- Primary school 17 pp under

- Married underrepresented by 25 pp

- English speakers are overrepresented

- But CATI better than IVR/SMS

Survey Research Methods

Journal of the European Survey Research Association

SRM   Current   Archives   About ▾

Home / Archives / Vol. 13 No. 3 (2019) / Articles

In Search of the Optimal Mode for Mobile Phone Surveys in Developing Countries. A Comparison of IVR, SMS, and CATI in Nigeria

Charles Q Lau
RTI International

Alexandra Cronberg
Kantar Public

Leenisha Marks
RTI International

Ashley Amaya
RTI International

📄 PDF

Published
2019-12-10

How to Cite

Source

GeoPoll

Source: Lamarage et al. (2016)

Design of Lamarage et al. (2016)

- Among the first examples of CATI RDD in Africa

- Project in Côte d'Ivoire from 2013

- 7-minute survey

Representation

Unweighted data versus general population from the Demographic and Health Survey:

- Women underrepresented by 19 pp (31% versus 50%)

- Age 50+ underrepresented by 10 pp (12% versus 22%)

- No education underrepresented by 34 pp (28% versus 62%)

- Rural underrepresented by 30 pp (12% versus 52%)

Nationally, coverage is 85% (96% urban, 75% rural): most differences are due to nonresponse, not undercoverage.

- Survey overestimates knowledge and behaviours related to HIV testing

INED 026A16
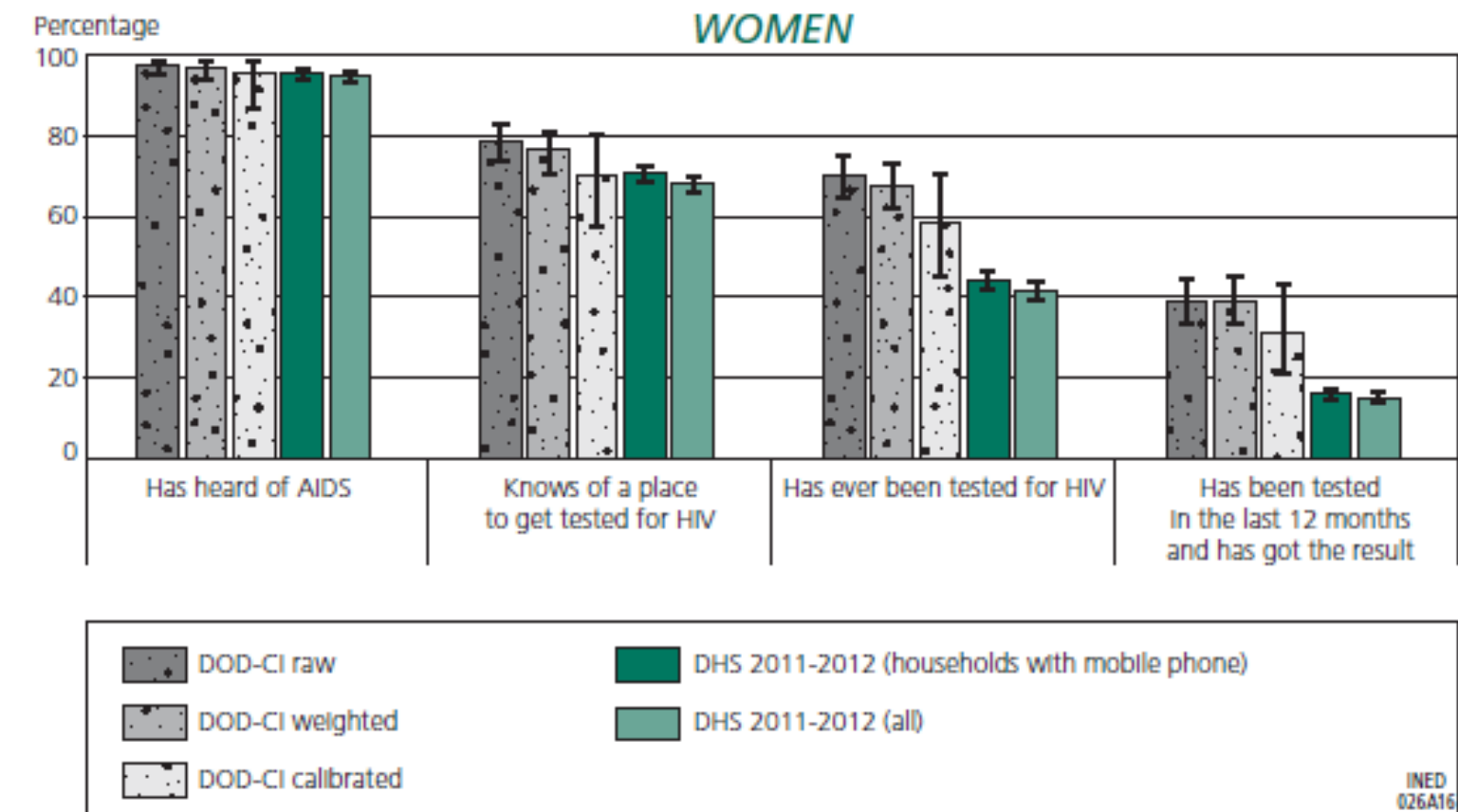
Example from Ethiopia, Malawi, Nigeria, Uganda

Design of Brubaker, Kilic, Wollburg (2021)

- Recontact of participants from household-level FTF surveys in each country

- First-round CATI response rates: Malawi: 74%, Ethiopia 60%, Nigeria 65%, Uganda 93%

- Project conducts multiple rounds of CATI follow-up

Representation

Groups less likely to participate in the first-round CATI:

- Not a household heads or spouse of HH head

- Not owning a mobile phone

- Less educated

- 15-24 year olds

- Not having a HH enterprise

- Lower household wealth

PUBLISH   ABOUT   BROWSE

PLOS ONE

🔒 OPEN ACCESS   📄 PEER-REVIEWED

RESEARCH ARTICLE

Representativeness of individual-level data in COVID-19 phone surveys: Findings from Sub-Saharan Africa

Joshua Brubaker, Talip Kilic, Philip Wollburg ✉

Published: November 17, 2021 • https://doi.org/10.1371/journal.pone.0258877

Source

GeoPoll

What Can We Do to Improve Representativeness?

1. Propensity Score Weighting (for follow-back surveys only)
2. Referrals
3. Quota-Based Design with weighting
4. Calling Protocols

5. Offering an Incentive
6. SMS Prenotification
7. Better Introductions
8. Refusal Conversation
9. Non-Response Follow up: Shorter surveys
10. GREAT INTERVIEWER RECRUITMENT AND TRAINING
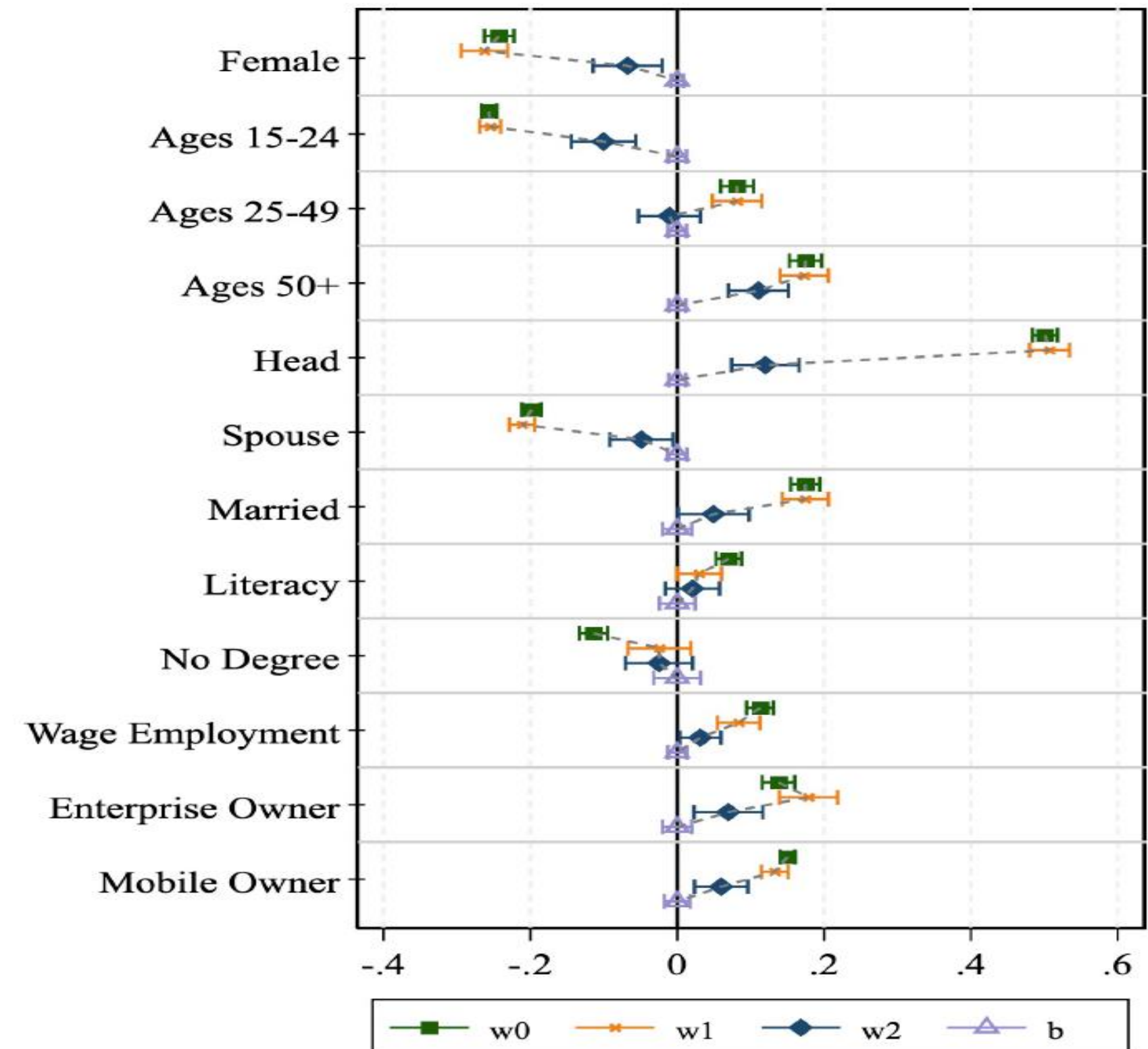
GeoPoll

Method only works when the CATI survey uses a prior FTF survey as a sampling frame ("follow back" design)

Nigeria (Figure 3) from Brubaker, Kilic, Wollburg, 2021)

Adjusting for Individual-Level Factors
(Brubaker, Kilic, Wollburg; 2021)

- Weighting generally helps reduce selection biases of CATI survey ☺

- Example: Weights address representation by gender and age in Malawi

- Example: Weights address representation by education in all countries except Ethiopia

- But effects are inconsistent, weighting doesn't solve issues, and variance increases ☹

Adjusting for Household-Level Factors

Ambel, McGee, Tsegay (2021)

- When the outcomes are at the **household level** (vs. individual level), weighting works much better compared to individual level outcomes.

- Bias due to undercoverage and nonresponse is reduced when weights are applied in most cases.

- But bias is not always eliminated completely.

- Remember: This method requires a "follow-back" survey design where the CATI survey interviews respondents from a prior household survey.

Figure 4.4: Bias Reduction Results – Consumption expenditure quintiles*

Consumption Expenditure Quintile*
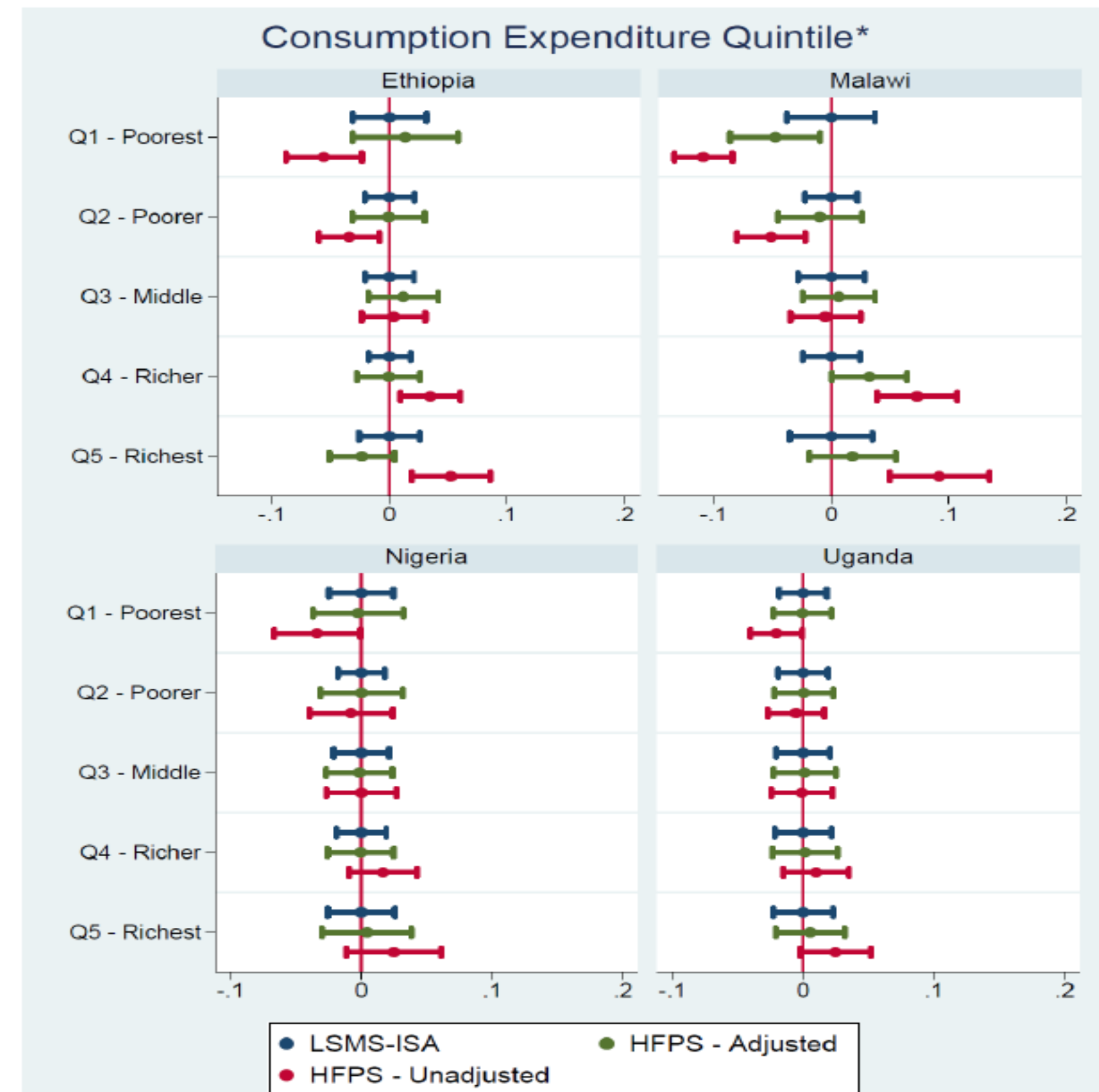
*For Malawi, the quintiles are based on a wealth index rather than consumption expenditures.

Referrals

Glazerman et al. (2023)

- 2020 CATI survey in Kenya using RDD
- If a man answered, 50% were randomized to "Pass the Phone" group
  - Most agreed, but many didn't have woman to refer to
- 263 referred woman complete
- Method improves representation by age, low assets, areas with low connectivity
- Method is more costly
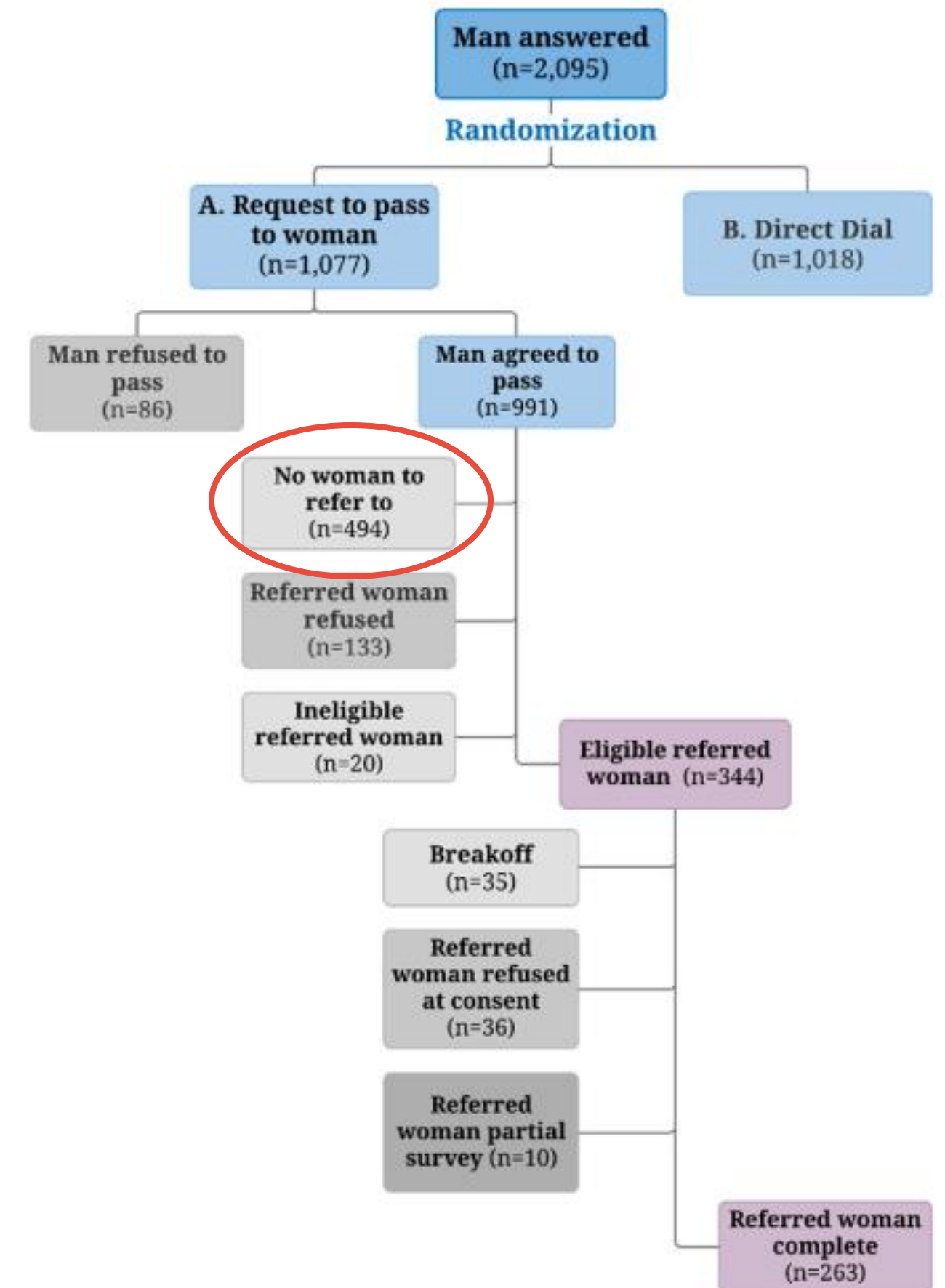
Quota Based Design

- Four-round CATI survey 2021-2022

- Sampling frame: Vendor list (non-probability)

- 36-minute survey on household welfare

- Quota design
  - Set quotas by location, gender, urban/rural, education, farming livelihood. Interlocked demographic x states.
  - Respondents were asked screening questions: if they belong to closed quota, survey ends.
  - Didn't complete to all quotas in 2 small states and low education in some states.

- After weights are applied, CATI matches demographics from prior FTF surveys

- CATI reaches 310 out of 330 townships: not possible in FTF

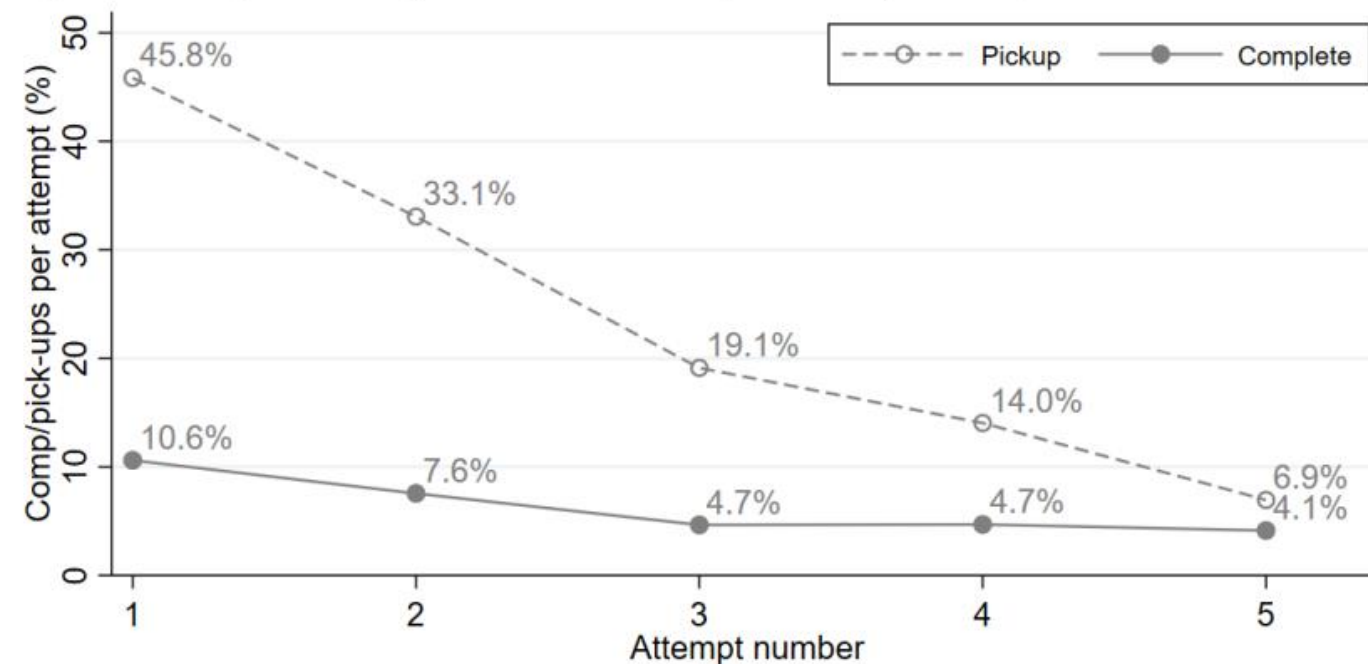- For statistical underpinnings of this approach, see Labrique et al. (2017)
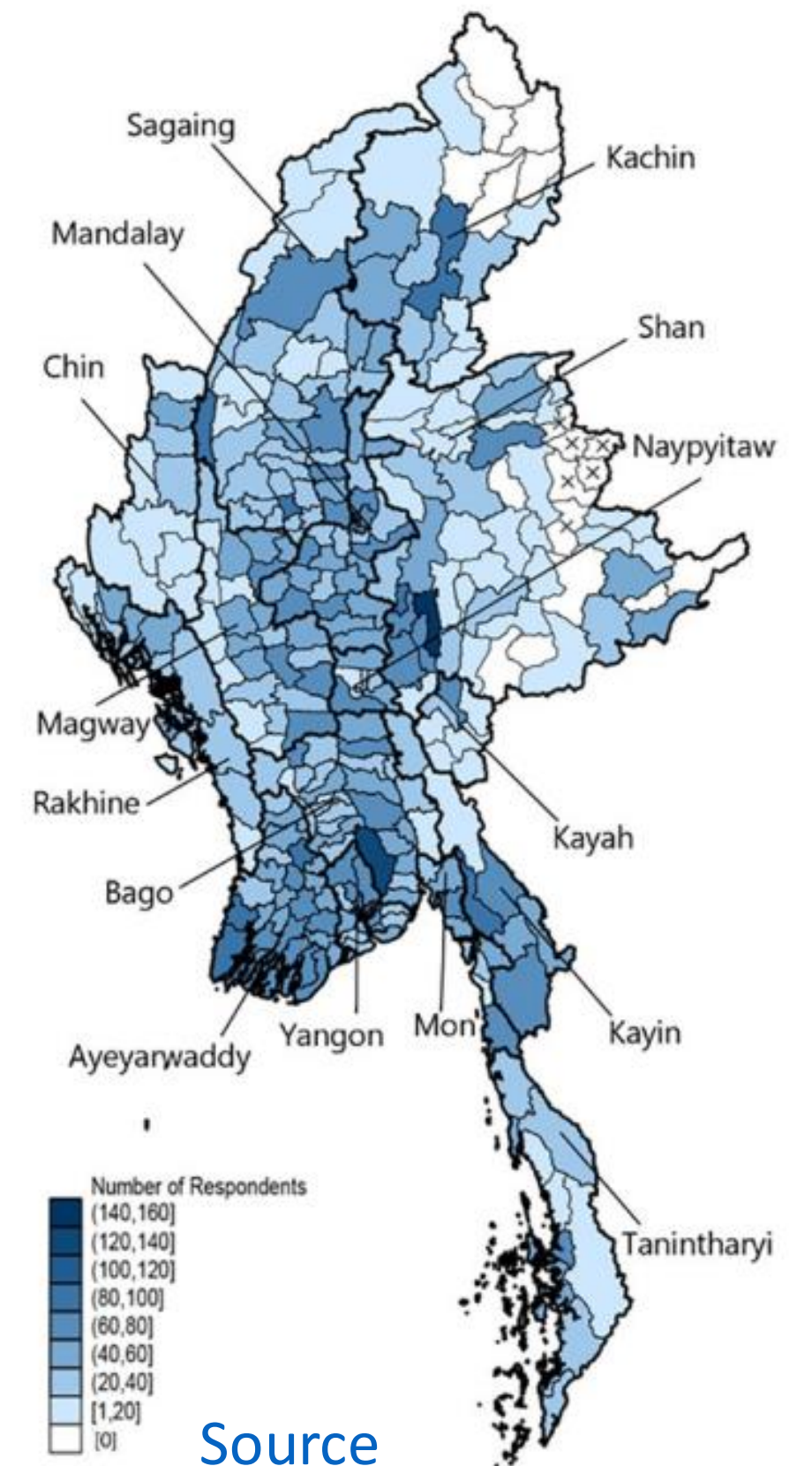
Source

CATI surveys in nine RDD surveys in 2020 (IPA) [Source](#)

More Call Attempts

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

Finding: Works, but inconsistently! Respondents who answer on 2+ call attempt are different from those who answer on attempt #1.

But very inconsistent: In Ghana and Uganda, those who answer after 2+ contact attempts are less likely to have a secondary education. But the reverse is true in Rwanda!

Rescheduling

Completion Rates

	Rescheduled	Not rescheduled
All sites	**13.6%**	**19.5%**
Rwanda	19.4%	38.9%
Burkina Faso	36.4%	59.3%
Sierra Leone	21.0%	39.1%
Philippines	8.0%	17.8%
Ghana	11.9%	22.3%
Mexico City	10.9%	5.6%
Colombia	15.0%	25.9%
Uganda	11.5%	10.2%
Zambia	37.1%	40.4%

Finding: Does not work. Respondents who complete after a reschedule are older and more likely to be male, have secondary education, employed, not in poverty. No evidence actively rescheduling calls reduces underrepresentation.

GeoPoll

Other Methods to Improve Representation

5. Offering an Incentive	6. SMS Prenotification	7. Better Introductions
8. Refusal Conversation	**9. Non-Response Follow up: Shorter survey**	**10. GREAT interviewer training**

GeoPoll

CATI Questionnaire

Two common questions:

1. How long **can** CATI questionnaires be?
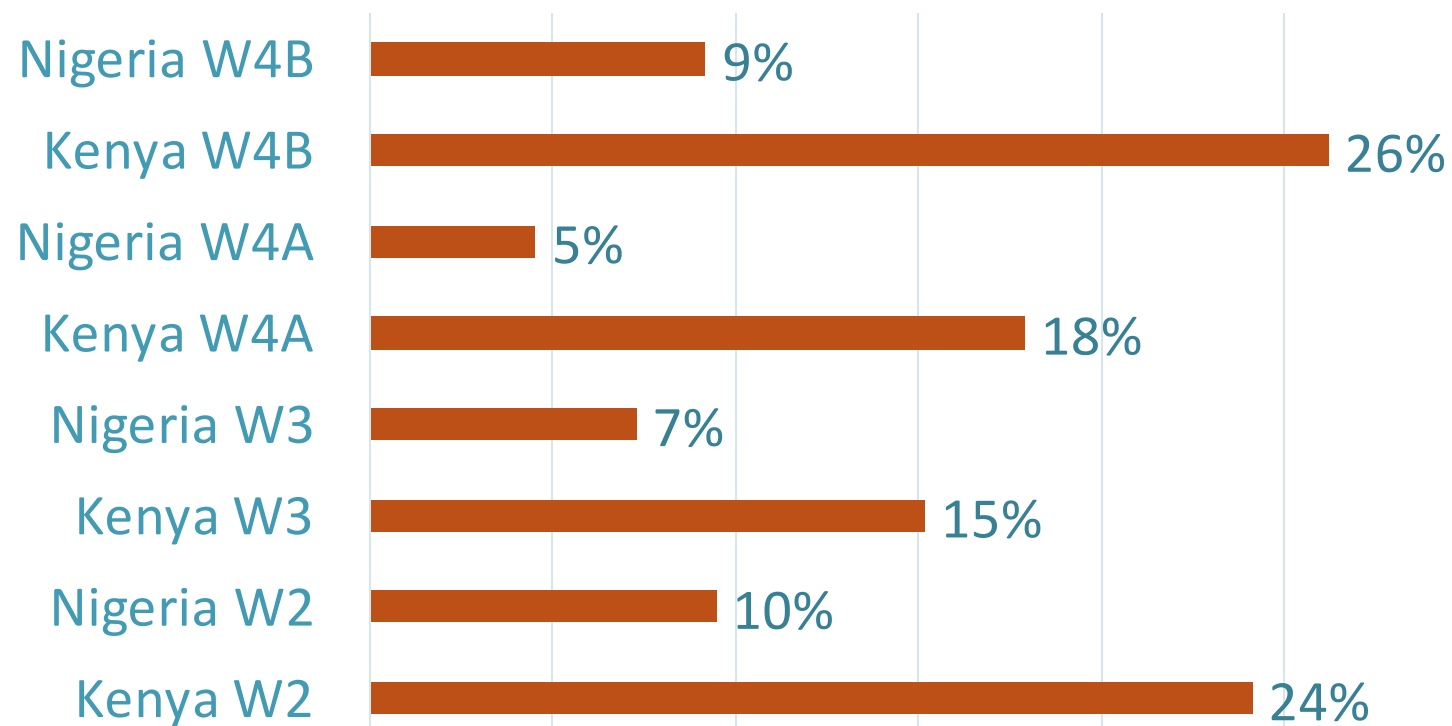2. How long **should** CATI questionnaires be?

Lots of opinions, but not a lot of empirical evidence

GeoPoll

How long **can** CATI questionnaires be?

GeoPoll Research on Breakoff in Kenya and Nigeria

- 8 CATI RDD surveys in Kenya and Nigeria: 2023-2024. Surveys are long (median >30 min) and ask about sensitive topics related to gender issues

- Breakoff rates vary, but most breakoff is early in survey

- Breakoff is associated with lower education, but not other respondent characteristics

Breakoff Rates

Survey	Rate
Nigeria W4B	9%
Kenya W4B	26%
Nigeria W4A	5%
Kenya W4A	18%
Nigeria W3	7%
Kenya W3	15%
Nigeria W2	10%
Kenya W2	24%

Experimental Evidence from Malawi

- 2022 CATI RDD survey in Malawi

- Respondents randomly assigned to 10-minute, 20-minute, or 30-minute questionnaire

- Breakoff (completion) rates were similar across experimental groups

Table I. Primary Outcomes, by Stated Interview Duration.

	10-minute Survey (%)	20-minute Survey (%)	30-minute Survey (%)	Total (%)
Cooperation rate	96.2	94.7	94.0	95.0
Completion rate	99.1	98.0	97.3	98.1

Data quality.

Torrisi et al. (2024)

How long **should** CATI questionnaires be?

- We don't *really* know …

  … and it probably depends on what we're measuring

- OK, but what **do** we know?

The Journal of Nutrition
Community and International Nutrition (ASN)

Respondent Fatigue Reduces Dietary Diversity Scores Reported from Mobile Phone Surveys in Ethiopia during the COVID-19 Pandemic

Kibrom A Abay,[1] Guush Berhane,[2] John Hoddinott,[3] and Kibrom Tafere[4]

[1]International Food Policy Research Institute, Cairo, Egypt; [2]International Food Policy Research Institute, Washington, DC, USA; [3]Division of Nutritional Sciences, Charles H. Dyson School of Applied Economics and Management, Department of Global Development, Cornell University, Ithaca, NY, USA; and [4]Development Research Group, The World Bank, Washington, DC, USA

Original Article

Field Methods
2024, Vol. 0(0) 1–17
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1525822X241248906
journals.sagepub.com/home/fmx
S Sage

Revisiting the Recommended Duration of Interviews Conducted by Mobile Phone in Low- and Middle-income Countries: A Randomized Trial in Malawi

Orsola Torrisi[1] ⊙, Jethro Banda[2] ⊙, Georges Reniers[3], and Stéphane Helleringer[1]

Abay et al. (2022) randomize the placement of a cognitively demanding food recall questions. When these questions are placed *later* in the questionnaire, respondents underreport food consumption (8-40%, depending on outcome)

Torrisi et al. (2024) compare data quality in a module on parental survival, which is at the end of a 10-, 20-, and 30-minute survey. Questionnaire stated length (10/20/30) didn't affect item-nonresponse, age heaping, or speed of answering parental survival module.

iGeoPoll

Reducing Breakoff: Impact of Topic and Placement

- Plutowski and Zechmeister (2024):
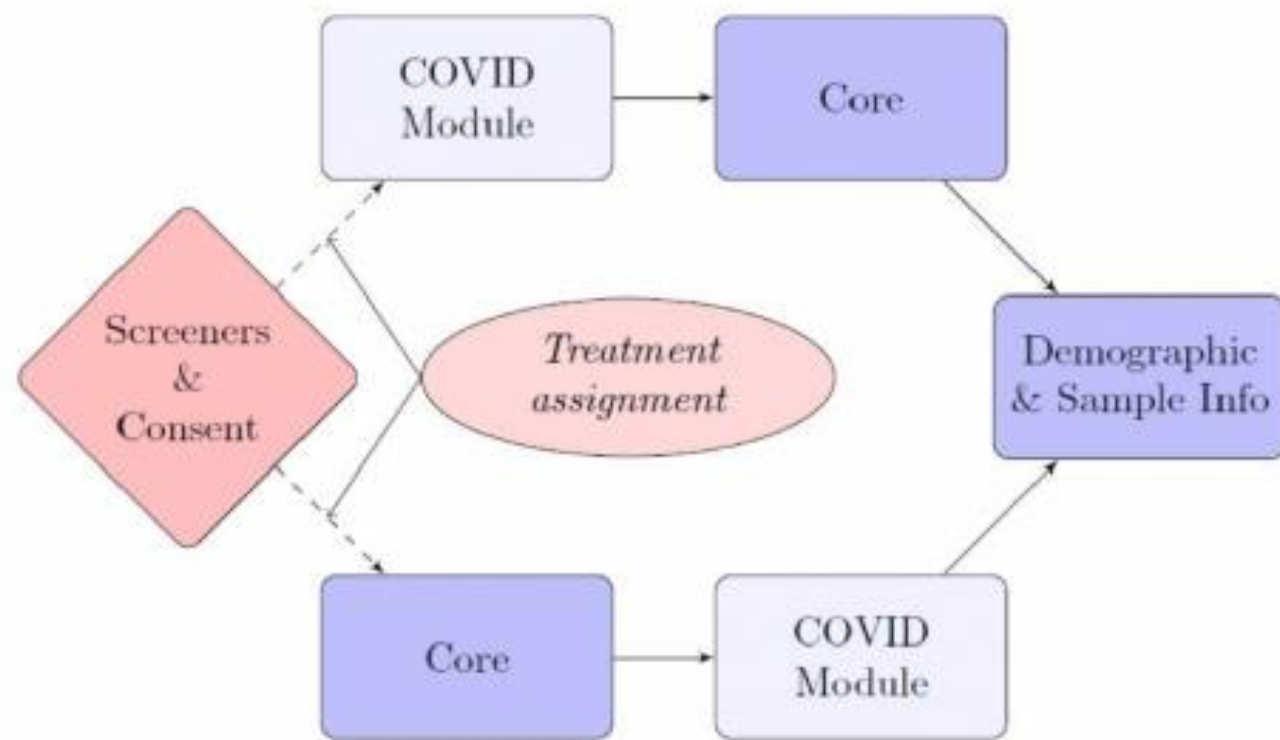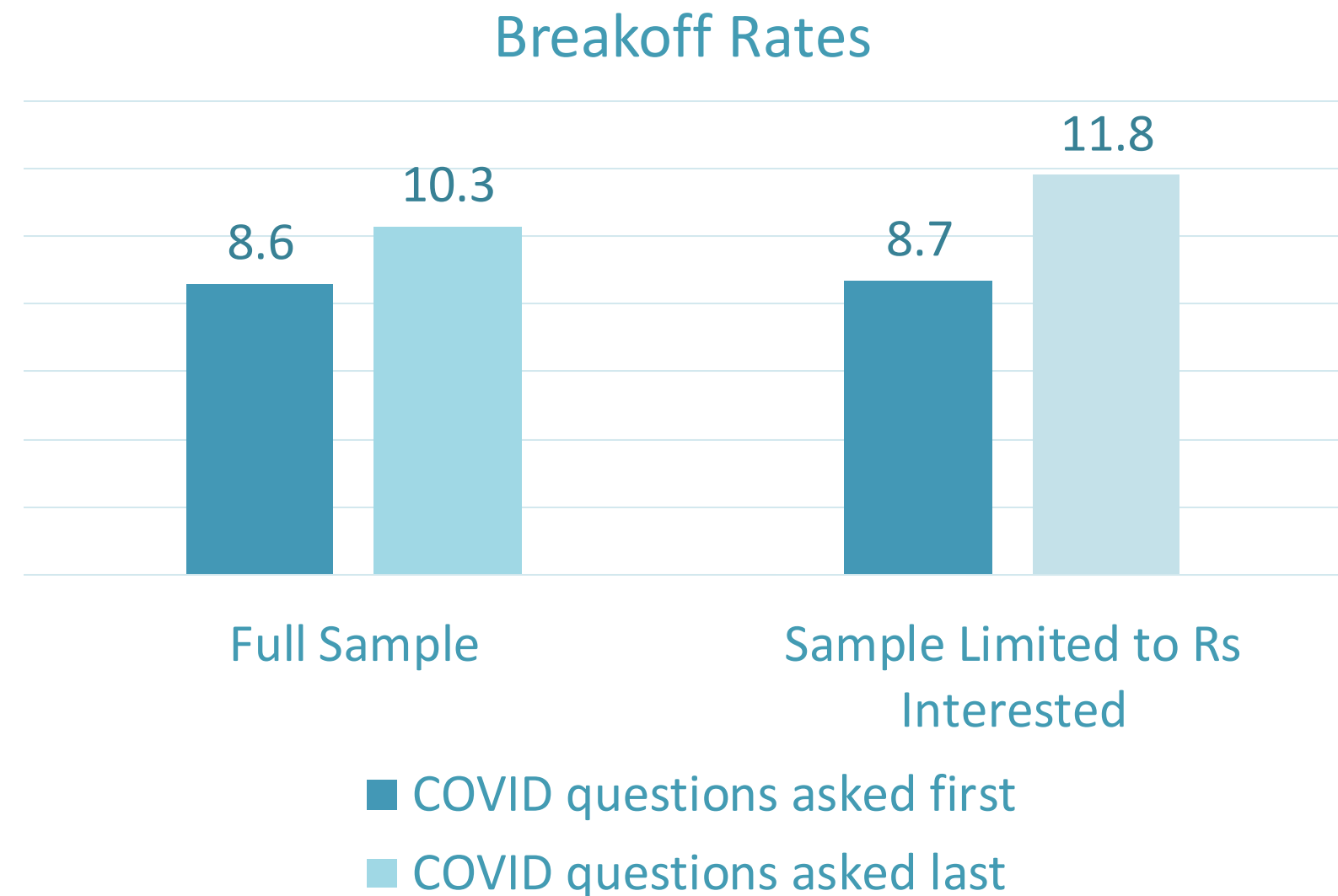 2020 CATI RDD survey in Haiti

Figure 1: Questionnaire Structure

Source

Breakoff Rates

|	COVID questions asked first	COVID questions asked last
Full Sample	8.6	10.3
Sample Limited to Rs Interested	8.7	11.8

CATI Questionnaires: Summary

- CATI questionnaires **can** be long: Many respondents will stay on phone and answer long (>30 min) CATI questionnaires

- CATI questionnaires that include cognitively demanding tasks (e.g., dietary recall) **should** probably not be too long

- There is some evidence that CATI questionnaires on less complex topics **could** be long, but much more research is needed
  - Please be careful, though! Pilot testing and careful monitoring is critical.
- Please embed experiments in your research when possible and share results!

GeoPoll

Reliability in CATI Surveys

Design of [Pariyo et al. (2019)](#)

- RDD CATI and IVR surveys in Tanzania and Bangladesh on noncommunicable diseases

Group 1

```
CATI Survey  → 7 days → IVR Survey
```

Group 2

```
IVR Survey  → 7 days → CATI Survey
```
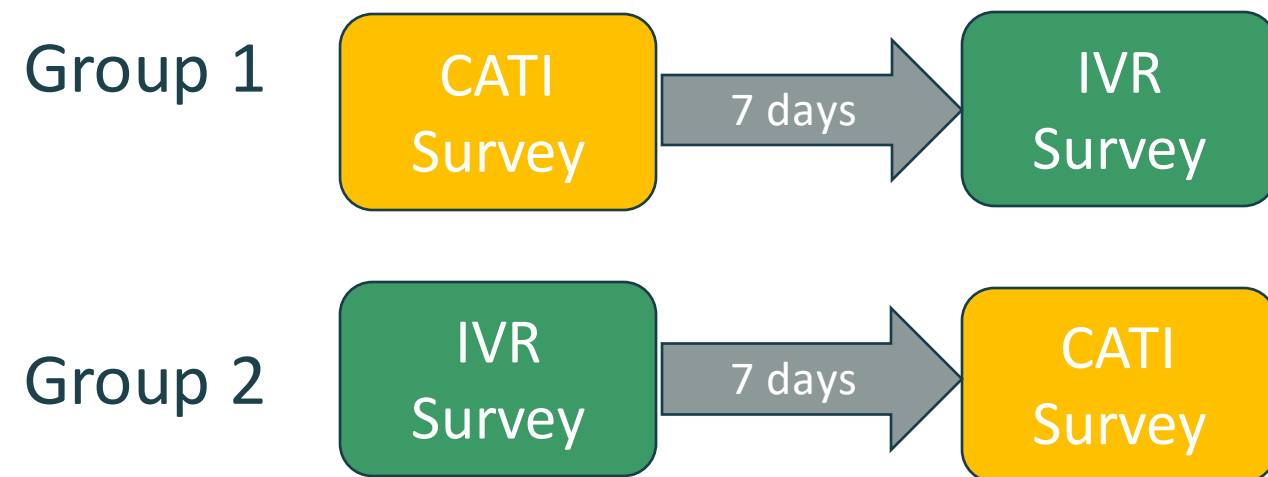
- Overall question: How well do reports about demographics and health agree between two surveys?

Key Findings

1. **High reliability** for demographics, alcohol, tobacco

2. **Moderate to high reliability** for hypertension

3. **Lower reliability** for physical activity, fruit/vegetable, salt intake

Now think about activities which take **MODERATE** physical effort that you do in a typical week. Moderate physical activities make you breathe a little harder than normal and may include brisk walking, dancing, digging or housework. Do not include walking. Again, think about only those physical activities that you did for at least 10 minutes at a time. **Sasa fikiria kuhusu shughuli zinazochukua kiwango cha nguvu cha wastani ambazo unafanya hasa kwa mfano katika wiki. Shughuli za mwili za kiwango cha wastani zinakufanya upumue kwa kiwango cha wastani kuliko kupumua kwa kawaida. Mfano ukiwa unatembea kwa haraka, kucheza muziki, kulima, kazi za nyumbani. Usijumuishe kutembea kwa kawaida. Fikiria tena shughuli/kazi ambazo ulizifanya kwa angalau dakika kumi mfululizo.**	
In a typical week, do you do any moderate physical activity? If Yes, Press 1. If No, press 3. **Kwa kawaida katika wiki unafanya shughuli/kazi za mwili za kiwango cha wastani? Kama NDIYO bonyeza 1, kama HAPANA bonyeza 3.**	1 = YES 3 = NO 0 = REFUSED OTHER = (Error_0_1)
In a typical week, on how many days do you do **moderate** physical activities? Please enter the number of days now. **Kwa kawaida katika wiki, ni siku ngapi unafanya shughuli za wastani za mwili? Tafadhali ingiza idadi ya siku.**	MC RANGE (1-7) 0 = REFUSED OTHER = (Error_0_1)
How many hours do you usually spend doing moderate physical activities in a day? If you do 1 hour or less than 1 hour, press 1; if you do about 2 hours, press 2; if for 3 hours, press 3 and so on. **Ni masaa mangapi hua unatumia kufanya shughuli za wastani za mwili kwa siku? Kama unafanya kwa saa 1 au chini ya saa 1, bonyeza 1; kama unafanya kwa masaa 2, bonyeza 2; kama kwa masaa 3 au zaidi, bonyeza 3 na kuendelea**	MC,RANGE (1-9) 0 = REFUSED OTHER = (Error_0_1)

GeoPoll

Reliability in CATI Surveys

Mafoud et al. (2015)

- Recontact of respondents from a face-to-face survey in Lebanon (82% recontact rate)

- Median time between FTF and CATI: 2 months

- 8-minute CATI survey (abridged questionnaire)

- 630 respondents participated in both surveys

Comparisons between FTF and CATI:

- Near-perfect concordance for age, health insurance, cigarette smoking, diabetes ☺

- Substantial for education, water pipe smoking, past-year alcohol, hypertension, others ☺

- Lower for physical activity (could be because of time difference between surveys or measurement error) : /

GeoPoll

CATI Reliability: Summary

- Good reliability for most measures

- But be careful about cognitively burdensome questions!

- Testing is important
  - Expert review
  - Cognitive interviewing
  - Pilot tests

- More research is of course needed ☺

GeoPoll

4. Challenges and Solutions: SMS

GeoPoll

SMS Questionnaire Design

| Aspect	Description
One survey question per SMS	• Question must include question and response options
Character limit	• Usually 160 characters • Special characters can take up double the space: this depends on mobile operator and software tool: be sure to test!
Response options	• Can be embedded in question or can be listed separately • Be sure to clearly identify the responses • Numbers are associated with response options – list them BEFORE the responses in SMS
Acceptable responses	• Oftentimes, numeric responses are only acceptable response • Non-valid responses trigger error messages • Many systems will terminate after 3-4 error message • More advanced systems can cover custom responses (e.g., "some" instead of "2")
Preamble or section header	• It is possible to have SMS step as a simple transition statement, e.g.., "The next set of questions are about voting. Press 1 to continue." Don't forget the "Press 1 to continue!" Or you could set this up as multiple SMS step.
Multi-select question	• Multi-select questions are possible, but not recommended because respondents sometimes don't know how to respond + Probing element missing as it is self completion
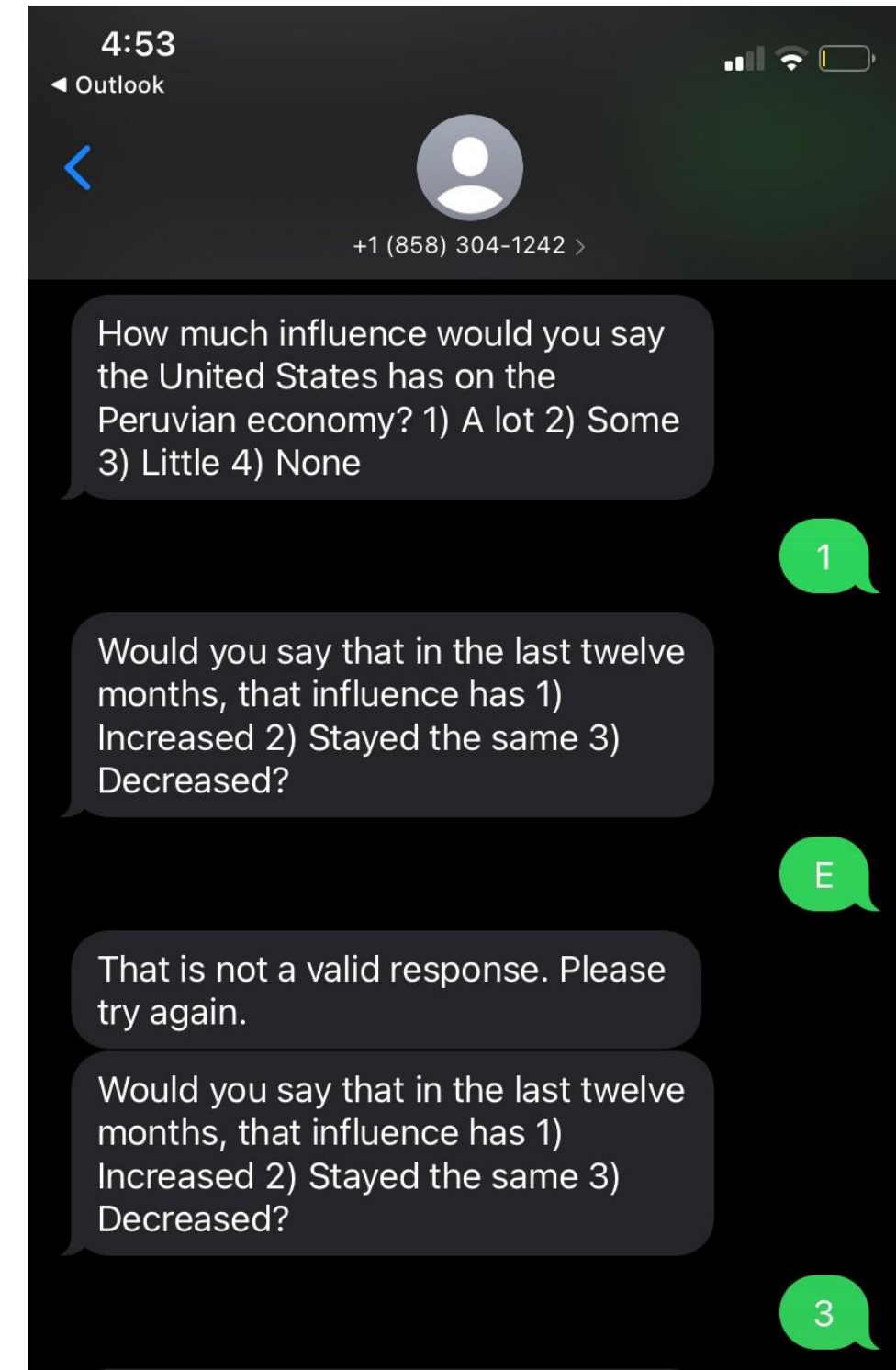
SMS: Remember the User Experience!

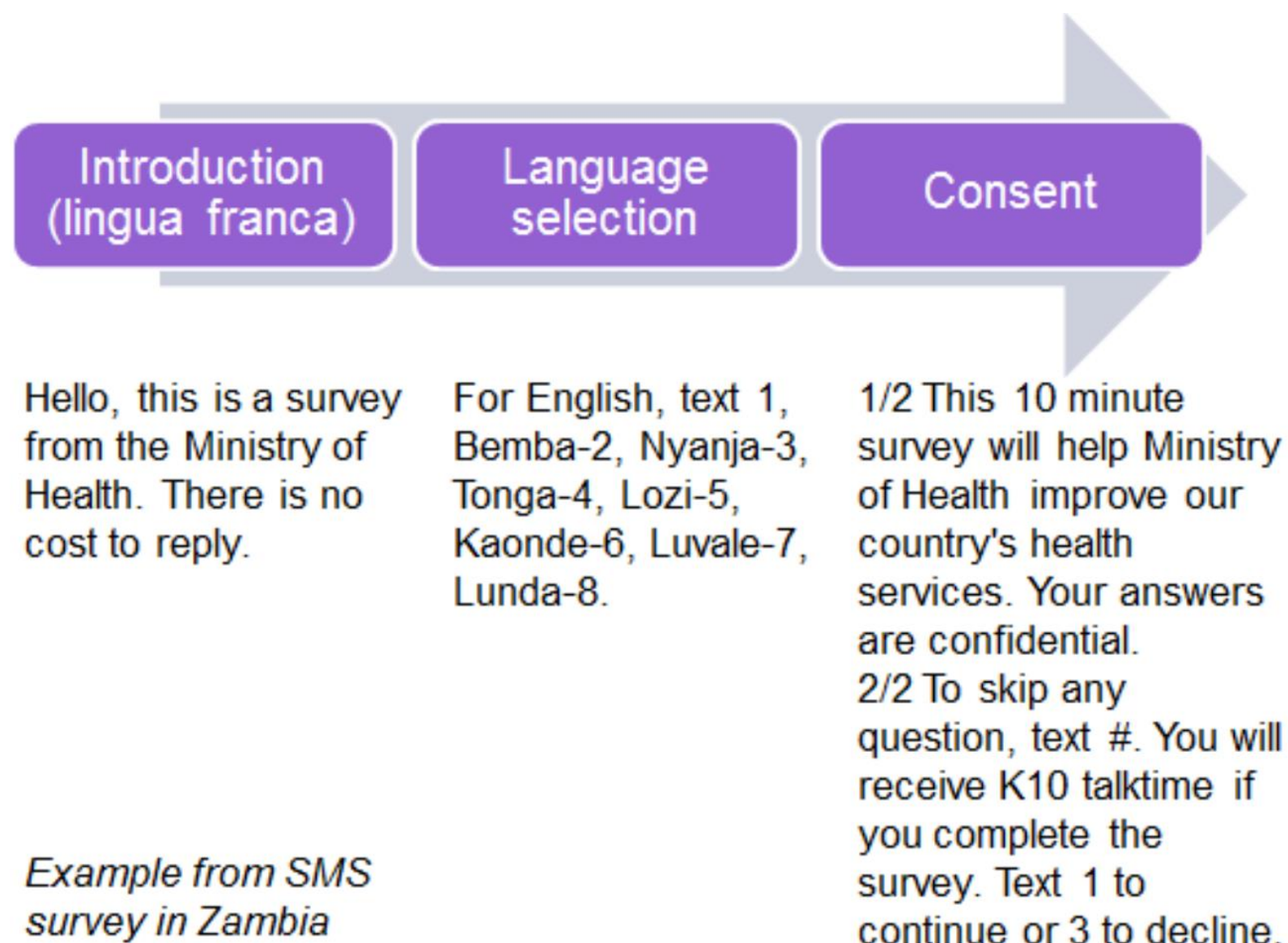*Multiple messages may come out of order*
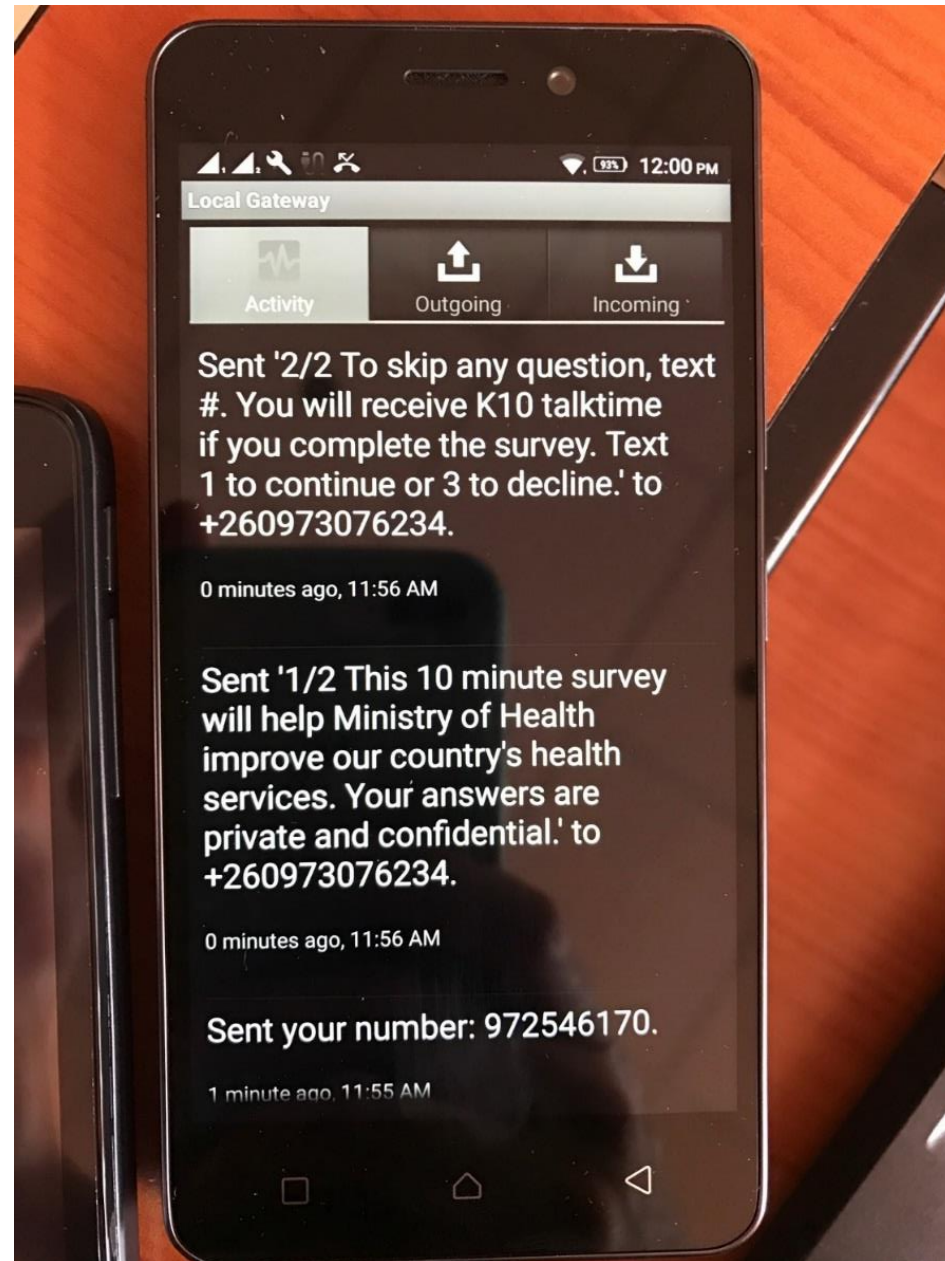
Respondents may have difficulty scrolling

GeoPoll
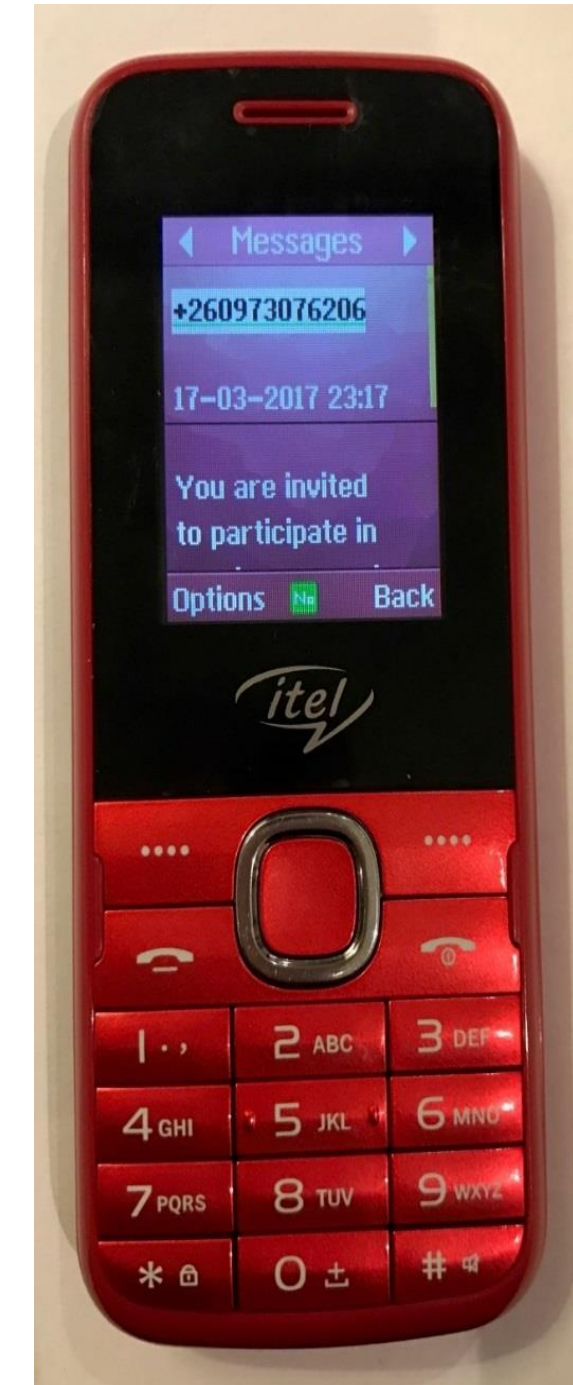
SMS: Language Selector

- Ask for language selection before official consent (in all languages)

- Language selector is in a "lingua franca"

- Consent is in selected language

- One challenge: How much should you include in the lingua franca?

Introduction (lingua franca)

Hello, this is a survey from the Ministry of Health. There is no cost to reply.

Example from SMS survey in Zambia

Language selection

For English, text 1, Bemba-2, Nyanja-3, Tonga-4, Lozi-5, Kaonde-6, Luvale-7, Lunda-8.

Consent

1/2 This 10 minute survey will help Ministry of Health improve our country's health services. Your answers are confidential. 2/2 To skip any question, text #. You will receive K10 talktime if you complete the survey. Text 1 to continue or 3 to decline.

GeoPoll

SMS Questionnaire Design: Order Effects

- Lau, Sanders, Lombaard (2019) randomized response options in an SMS survey

- The order in which we present response options in SMS surveys matters

- For example: in Uganda, the first response option was selected 26% of the time. The fifth response option was selected 16% of the time. This is independent of the content of the response option.

Experiments in 4 African SMS surveys show consistent evidence of primary effects

Likelihood of Endorsing Response, by Position (Uganda)

26%	21%	18%	16%	16%
1st	2nd	3rd	4th	5th
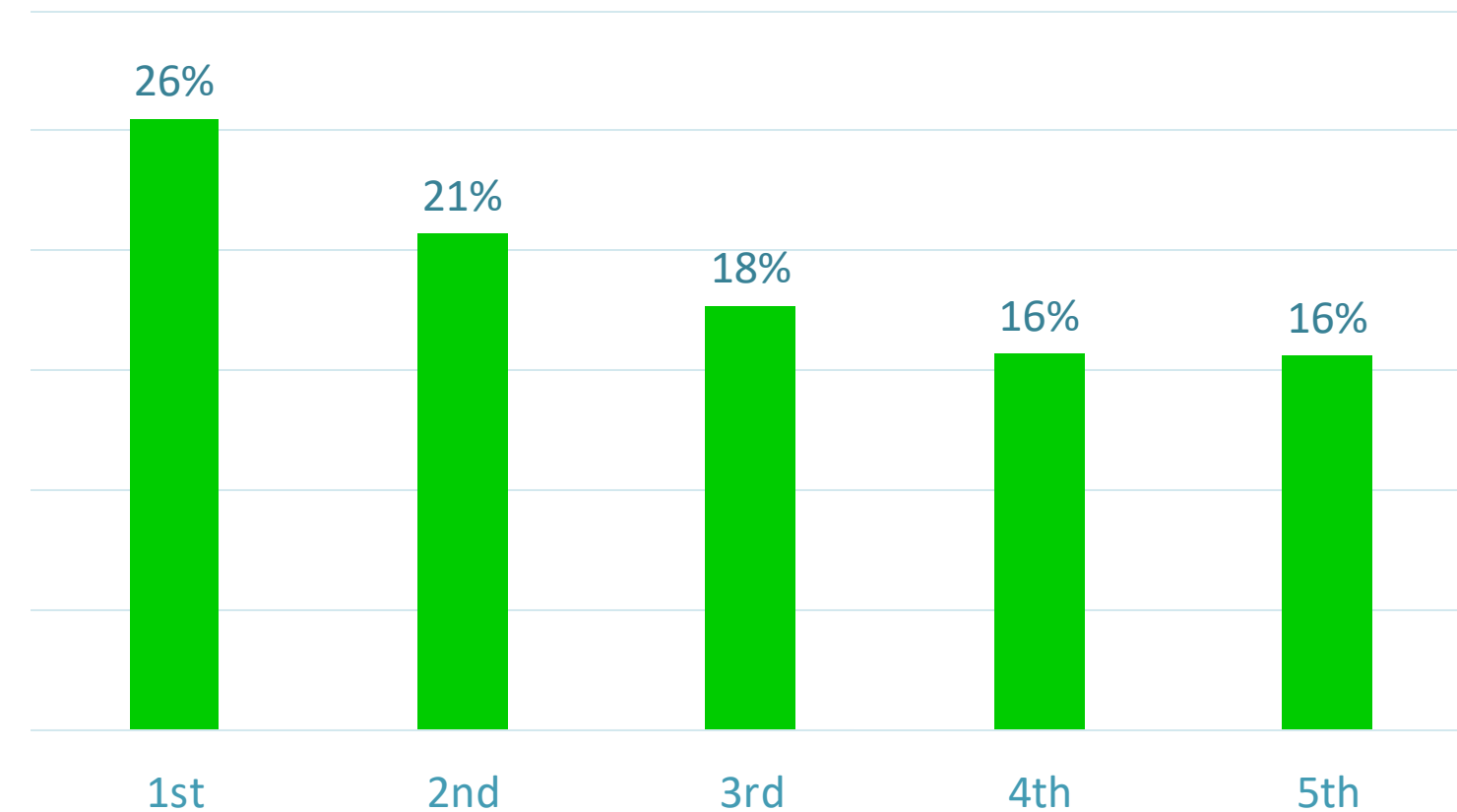
Source: Lau, Sanders, and Lombaard, 2019

GeoPoll

SMS Questionnaire Design: Other Tips

Test **locally, and in-person**
(Firchow and Mac Ginty, 2017)

Avoid embedding response options in question
(L'Engle et al., 2018)

Consider a modular design for longer surveys
(West, Ghimire, Axinn 2015)

For numeric questions, considering using categories instead of exact numeric responses (L'Engle et al., 2018)

GeoPoll

SMS Implementation

Channel to route messages via telecommunications network

Approval of/by mobile network operator (MNO)

Incentive delivery – Airtime/ Mobile Money and Turnaround time

Another method: USSD, but this approach has high drop-off rates and limitations on question structure

iGeoPoll

5. Q&A

GeoPoll

CONTACT
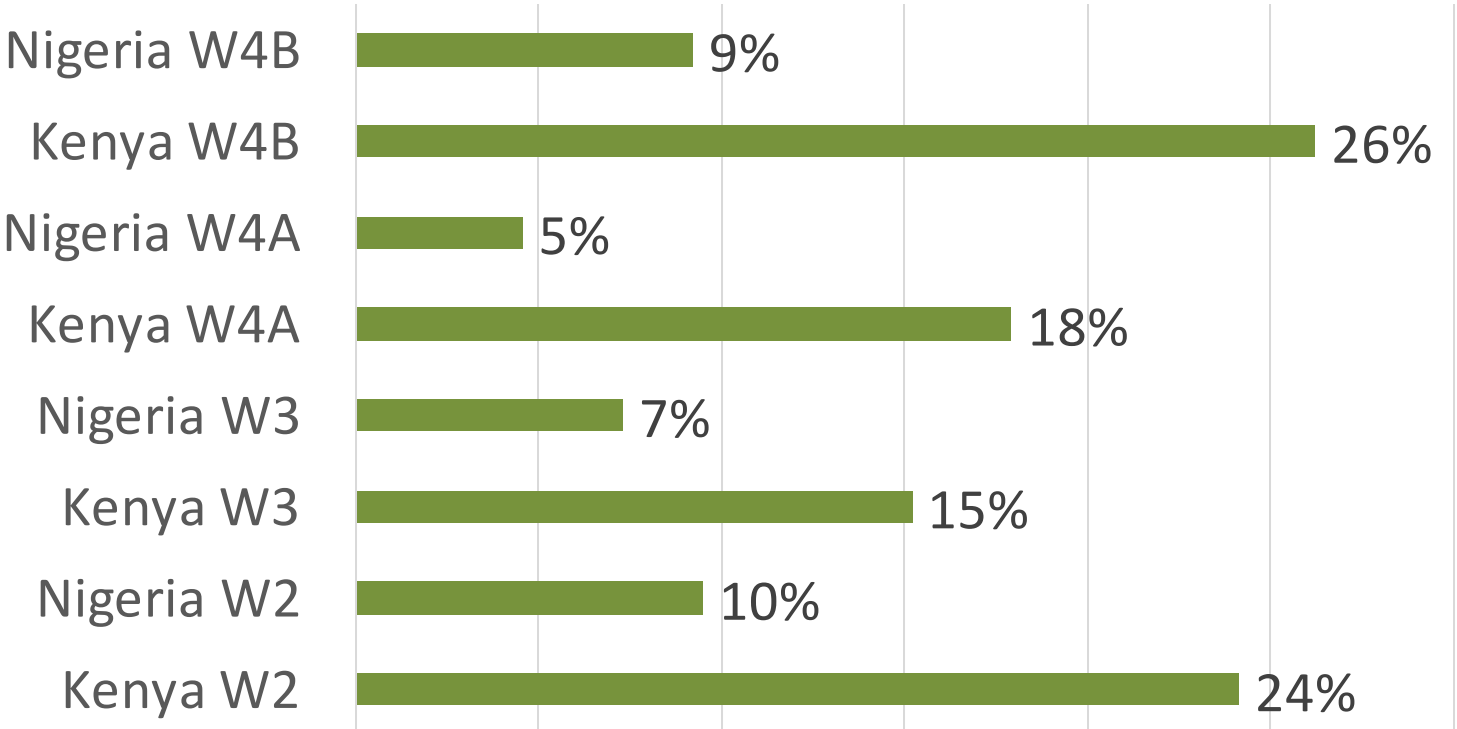
✉ JPMurunga@geopoll.com
Charles.Lau@geopoll.com

🌐 www.GeoPoll.com

Breakoff

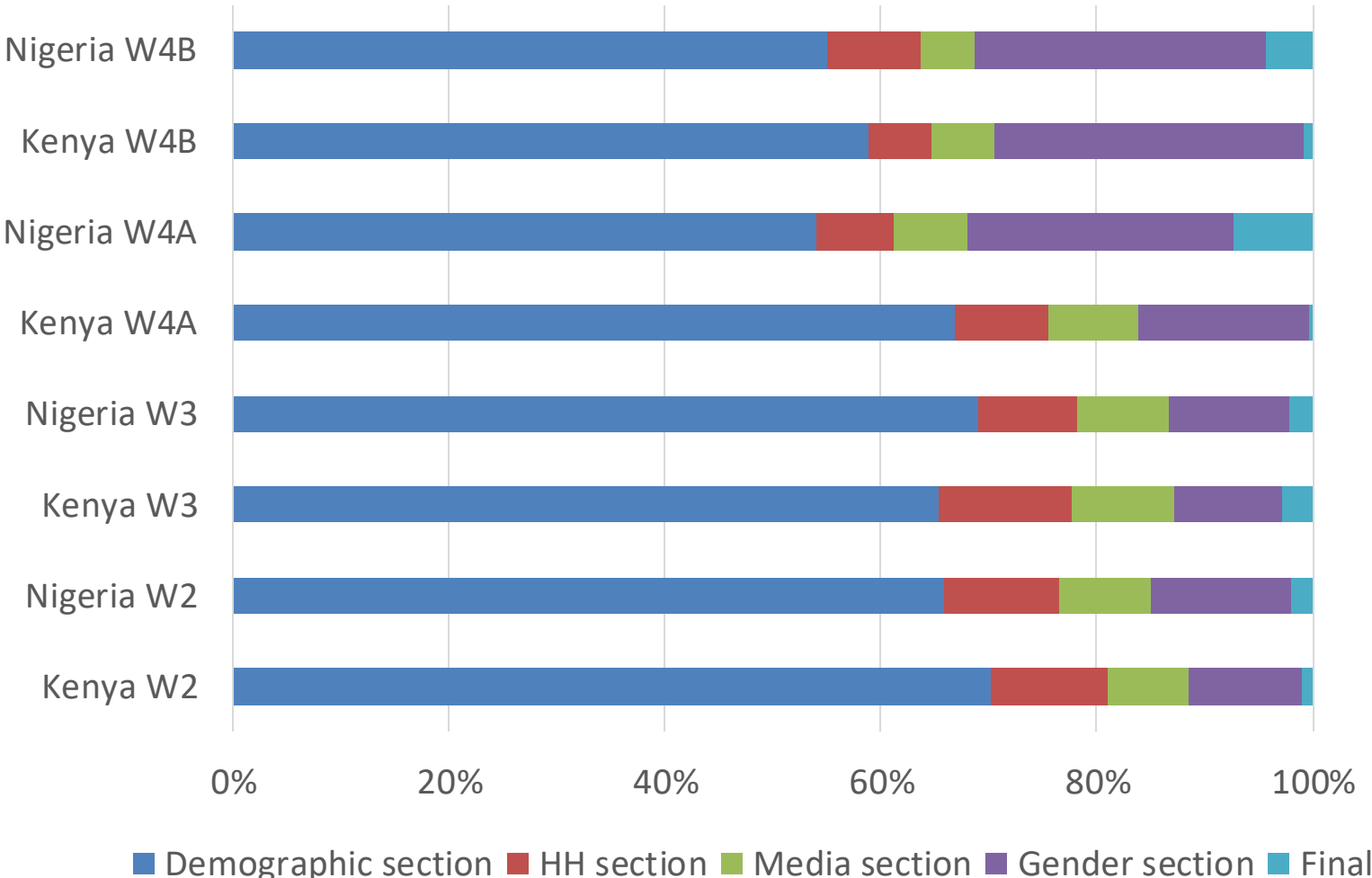## Breakoff: Respondents who start but do not complete

- GeoPoll Research on 8 CATI RDD surveys in Kenya and Nigeria: 2023-2024. Surveys are long (median >30 min) and ask about sensitive topics related to gender issues
- Significant variation in breakoff

Breakoff Rates

Survey	Rate
Nigeria W4B	9%
Kenya W4B	26%
Nigeria W4A	5%
Kenya W4A	18%
Nigeria W3	7%
Kenya W3	15%
Nigeria W2	10%
Kenya W2	24%

- Most breakoff occurs early in the survey

Location of Breakoff

Legend: Demographic section, HH section, Media section, Gender section, Final
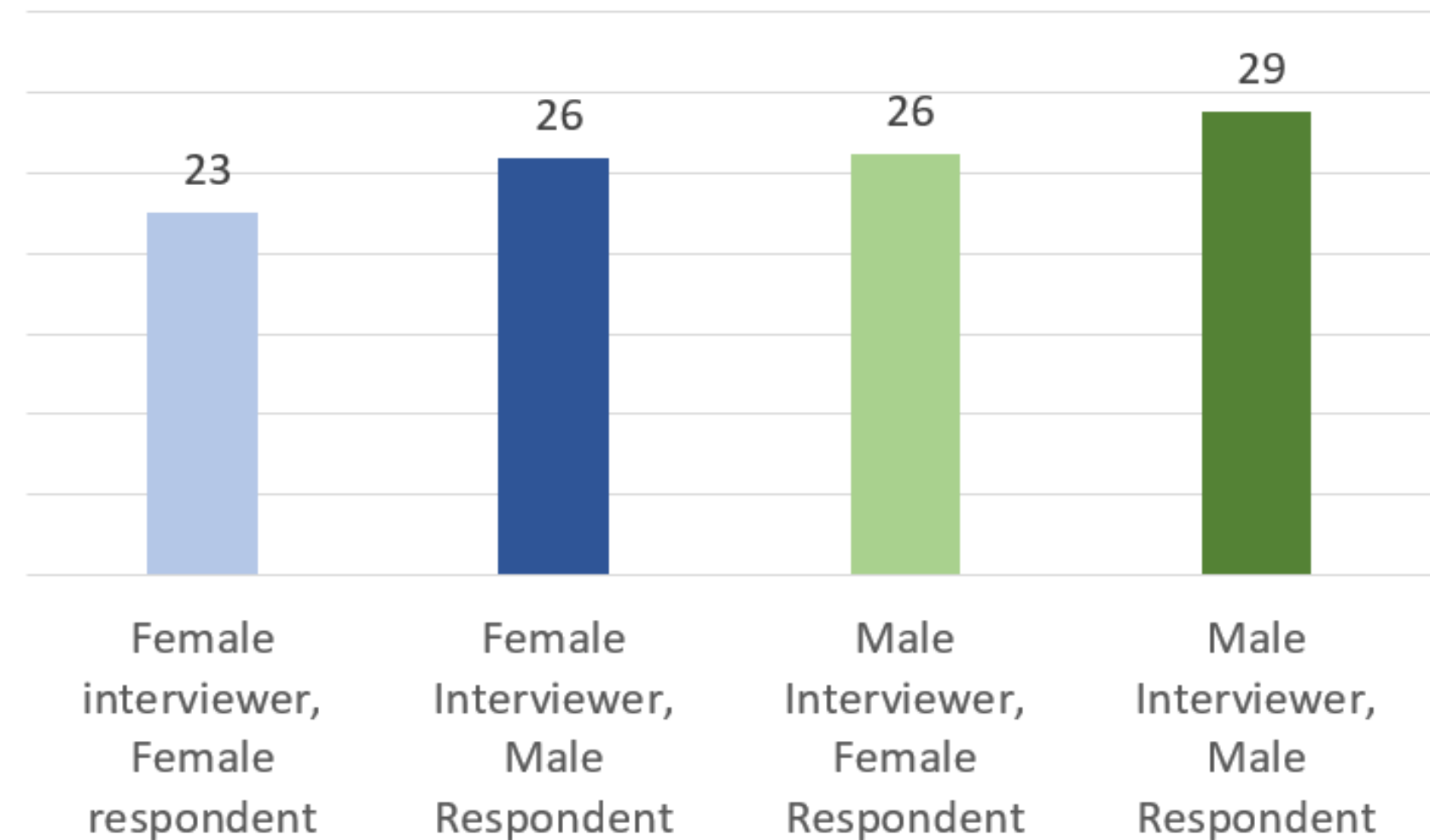
Breakoff

GeoPoll research (continued)

Factors Associated with Breakoff

- **Respondent:** Low education
  - But not gender, rural, age)
- **Interviewer experience**
  - But not gender or quality)
- **Individual interviewer**
- **Call centre partner**

Breakoff Rates, by Interviewer and Respondent Gender (W4B only; percents shown)

Female interviewer, Female respondent	Female Interviewer, Male Respondent	Male Interviewer, Female Respondent	Male Interviewer, Male Respondent
23	26	26	29

Breakoff: Tips and Recommendations

Measure and monitor

Coach interviewers

Use breakoff as signal for interviewer and partner quality

Place more engaging questions earlier in survey when possible

Avoid long introductions to modules

Use question-level breakoff rates to flag specific questions

More research on optimal questionnaire length, especially considering quality