Insight2 Impact
Advancing Financial Inclusion

# Innovative and sustainable data collection

## Reflections on financial inclusion pilot SMS surveys in Tanzania, Uganda and Zimbabwe

**Technical Note | June 2019**

Established and driven by

cenfri

FMT
FINMARK TRUST

Author

Shirley Jeoffreys-Leach

About insight2impact

insight2impact is a resource centre that aims to catalyse the provision and use of data by private and public-sector actors to improve financial inclusion through evidence-based, data-driven policies and client-centric product design.

insight2impact is funded by the Bill & Melinda Gates Foundation in partnership with The MasterCard Foundation.

For more information:

Visit our website at **www.i2ifacility.org**.
Email Mari-Lise du Preez (Partnerships Manager) at **mari-lise@i2ifacility.org**.
Call us on **+27 21 913 9510**.

# Contents

**The findings of the pre-pilots suggest that mobile surveys are a viable means to collect useful information.**

# Executive summary

**In late 2017, insight2impact facility conducted a pre-pilot SMS survey as a precursor to a large-scale multi-mode multi-country mobile pilot project. This pre-pilot was conducted to explore various aspects of setting up and running mobile surveys to measure different dimensions of financial inclusion. Pre-pilot surveys were conducted in Uganda, Tanzania and Zimbabwe.**

The aim of the pre-pilots was to gauge the range of constructs that could be measured using SMS surveys, how best these could be asked in 20 or fewer questions and how long it would take to design such a questionnaire, have it scripted, tested and fielded. inisght2impact also wanted to get a sense of how long it would take to collect that data and what the cost of SMS surveys would be relative to traditional survey methods. It was important to scan the environment for providers of SMS data collection services and the quality of sampling frames on offer. Of secondary importance was the issue of trying to exactly replicate indicators collected via traditional data collection methods or to design an appropriate weighting and adjustment plan to address known biases in the sample. This is the focus of the next round of full-scale piloting in eight countries, with an expected sample size of 34,000 respondents.

The results suggest that:

- It is possible to extensively measure various aspects of financial inclusion in separate modules in a short SMS questionnaire. In this pilot, three modules were explored, namely a Digital Financial Services (DFS) module based on the Financial Inclusion Insights survey (FII) DFS Indicators produced by Intermedia for the Bill & Melinda Gates Foundation, a module that measures the Financial Inclusion Access Strand developed by FinMark Trust (which is measured in all FinScope Surveys) and a Resilience module based on the needs-based framework developed by insight2impact, a programme jointly hosted by FinMark Trust and Cenfri.
- The cost of conducting an SMS survey is equal to 30% of a full-length 60-minute demand-side survey.
- The data was collected four times faster than a traditional demand-side survey.
- All mobile sample frames are limited and these limitations are not always well documented.
- A more sophisticated weighted and adjusted model is required to address the under-coverage of certain key populations.
- There was variation in the extent to which the various indicators matched those of the face-to-face surveys. The results are promising, considering we did not set out to develop perfect instruments, interlock the quotas, use additional mobile-data collection modes or use post-stratification modelling to adjust for the under-coverage of certain populations.

The findings of the pre-pilots suggest that mobile surveys are a viable means to collect useful information; and to this end we have planned an extensive piloting programme for 2018/2019 that will address the various methodological issues.

**The ubiquity of the mobile phone allows researchers new and distinct advantages when it comes to demand-side-survey data collection.**

# 1. Introduction

Data can shape and direct policy and product development in support of economic development. It also has transformative power, as it allows us to define appropriate indicators, track these indicators and use them to hold governments accountable. Arguably the greatest barrier to more and sustainable financial inclusion research is the cost of data collection. The high cost of survey data collection means that demand-side measurement occurs infrequently, as funders are becoming more reluctant to fund large, expensive surveys. Not only are these surveys expensive but they are also long and cumbersome, resulting in lower response rates and poorer data quality.

If the cost of demand-side data collection could be significantly reduced, researchers could design more regular, targeted and relevant surveys. This would provide data more often and ensure that surveys be more engaging for participants and that participants provide better-quality responses.

It is with this sustainability of data collection in mind that the insight2impact has set out to test what technology can offer in terms of faster, cheaper and more relevant data collection methods. In a previous note, we explored various mobile-data collection technologies. Based on this analysis, we selected SMS as our preferred method for these pre-pilots, as it is the most ubiquitous and the simplest to use.

## 1.1. The advantages of mobile research for financial inclusion

Mobile-phone penetration throughout the developing world continues to grow, and the GSMA (2018) predicts that in sub-Saharan Africa the regional subscriber base will grow at a compound annual growth rate of 4.8% for the next five years – more than double the global growth rate. Developing markets are in some instances at the forefront of the development of mobile technology, and the technology is changing how and where people interact and transact financially. The ubiquity of the mobile phone allows researchers new and distinct advantages when it comes to demand-side-survey data collection. These advantages are laid out in Table 1.1.

### Table 1.1. Mobile research advantages

| | |
|---|---|
| **Rapid response** | Traditional face-to-face surveys require detailed planning and logistics to get enumerators out into the field. Mobile surveys – be they interviewer-administered over the telephone or self-completed via the mobile device – can reach people as and where they are. Mobile phones are personal devices that travel with people wherever they go, so participants can complete mobile surveys anywhere and without the need for an enumerator to seek them out or schedule time when they will be available. |
| **Cost efficiency** | In the case of self-completion surveys, invitations to participate can be sent out in bulk. This means that sample sizes that would require months of face-to-face fieldwork to complete can be collected via mobile in a matter of days. Mobile surveys must also be short to ensure that people complete them.<br><br>The combination of mobile surveys being short, quick to turn around and in some cases self-completed means that they are cheaper to administer and manage, as the human resource and logistics requirements for data collection are smaller and the time required to design and analyse data is reduced. This is a core benefit of the mobile demand-side survey, as regular data collection, when it can be collected at a lower cost, is more sustainable. |
| **Short survey design** | The requirement that mobile surveys be short and engaging is a strength of the method. It forces researchers to focus on the key issues and ensures better-quality responses from the people participating in the surveys because they remain engaged throughout. |
| **Reach** | Mobile surveys can reach across wide geographic locations in a fraction of the time that traditional interviewer-administrated surveys can. Additionally, areas that are often difficult or impossible to access with enumerators (e.g. war-torn zones or gated communities) are reachable with mobile surveys. |
| **New data-collection** | Mobile surveys open up the possibility of reaching people "in the moment", as people tend to have their mobile phones on them at all times – thus giving researchers closer access to people's behaviour as and when it happens. |

With these benefits in mind, insight2impact launched a multiyear piloting programme in 2017 to develop a methodology that will allow for the collection of high-quality data through mobile-data collection technologies. The programme has three phases: 1) pre-piloting to explore some basic logistical aspects of conducting mobile surveys, 2) pilot testing in four African and four Asian countries, with a sample size of 34,000, focusing specifically on DFS indicators and 3) a possible second round in all eight countries to refine the methodology, instruments and models.

# 2. The 2017 pre-pilot objectives

The pre-pilots were implemented to kick-start insight2impact's thinking on how best to design a mobile survey programme that would produce high-quality reliable indicators. The primary aim was to get a sense of the lay of the land as it were, as opposed to trying to produce reliable indicators on the first attempt, despite the team having done a number of these surveys in areas other than FI. It was also decided that we would focus on a single mode, namely SMS, as we are more familiar with mobile CATI, which will be employed in the full pilots.

**Objectives**

- To determine the cost of SMS surveys relative to traditional survey methods
- To get a sense of how long it would take to collect the data
- To gauge the range of content modules that could be measured using SMS surveys
- To determine how best to design a questionnaire and represent the content in 20 or fewer questions
- To determine how long it would take to design such a questionnaire and to have it scripted, tested and fielded
- To scan the environment for providers of SMS data collection services and to assess the quality of sampling frames on offer relative to now population statistics
- To replicate and compare indicators collected via traditional data collection methods

[1] Plain Language is capitalised in this note to distinguish the legal meaning from the lay meaning (plain language is simple language).

# 3. Project design and implementation findings

## 3.1. Cost reduction

An SMS survey is 12 times cheaper than a traditional survey. Admittedly, they do not collect the same amount of data, but they do allow us to collect specific types of data from specific audiences faster and more cost-effectively. As mobile surveys are much cheaper, one could run four projects that cover different content in a year, and the price would still be a third of a single full survey.

**Table 3.1. The cost of different data collection modes[1]**

| Face-to-face, interviewer administered | Telephonic, interviewer administered | Mobile, self-completion |
|---|---|---|
| USD60 | USD15 | USD5 |
| Cost assumptions:<br>• 60-minute questionnaire<br>• Probability sampling<br>• High involvement of research supplier in questionnaire design | Cost assumptions:<br>• 10-to-15-minute questionnaire<br>• Random-digit dialling<br>• Low involvement of research supplier in questionnaire design | Cost assumptions:<br>• 10-to-15-minute questionnaire<br>• Quota sampling<br>• Low involvement of research supplier in questionnaire design |

Existing financial inclusion research programmes stand to benefit from the inclusion of alternative methods of data collection that can be administered at a low cost. Survey topics that work well on an alternative method can be removed from existing surveys that require probability sampling. Rationalising these surveys will make them quicker and more affordable to complete. This reduction in survey length can also lead to an improvement in data quality, as the burden on respondents to engage with them is lowered. The use of a supplementary alternative method, where appropriate, will ensure that richness of data is not lost through rationalisation. Further, the use of alternative, more affordable modes of data collection can encourage the development of new financial inclusion research programmes designed specifically for the method.

## 3.2. Timing

Fielding a mobile survey is around three to four times faster than a traditional survey, as illustrated in Table 3.2.

**Table 3.2. The timing of different data collection modes**

| | Face-to-face, interviewer administered | Telephonic, interviewer administered | Mobile, self-completion |
|---|---|---|---|
| Survey and sample design[2] | 6 weeks | 2 weeks | 2 weeks |
| Field preparation | 1 month | 2 weeks | 2 weeks |
| Data collection (n=5000) | 3 to 6 months | 6 weeks to 2 months | 2 weeks to 1 month |
| Data cleaning and weights | 1 month | 1 month | 1 month |
| Total time | 8 to 11 months | 3 to 4 months | 2 to 3 months |

## 3.3. Countries and modules

We started our pre-pilot programme by testing four different SMS survey modules in Tanzania, Uganda and Zimbabwe.

The modules developed were:

- The digital financial services (DFS) indicators from the Financial Inclusion Insights (FII) survey
- An overall access strand replication from the FinScope survey
- A short-form measurement of people's liquidity needs
- A short-form measurement of people's resilience needs

This initial phase of pilots included five mobile SMS survey pilots conducted in December 2017 and January and February 2018. An overview of the main aims and the broad design of each pilot is detailed in Table 3.3.

[2] This survey design time assumes no qualitative pre-testing. This step should ideally be included, particularly the first time a specific instrument is being implemented and will add 6 weeks to 2 months to design time across all modes.

**Table 3.3. 2017 SMS pilots**

| Country | Tanzania | Uganda | Tanzania | Zimbabwe | Zimbabwe |
|---|---|---|---|---|---|
| Main aim | Replicate DFS measures from FII | Replicate DFS measures from FII | Short-form measurement of liquidity needs | Short-form measurement of resilience needs | Replicate overall access strand from FinScope |
| Sample | 500 | 1000 | 500 | 500 | 500 |
| Location | Dar es Salaam. Morogoro, Pwani | National | Dar es Salaam. Morogoro, Pwani | Harare and Mashonaland | Harare and Mashonaland |
| Length | 25 questions | 26 questions | 23 questions | 25 questions | 21 questions |

## 3.4. Questionnaire design and use cases

Mobile self-completion questionnaires need to be significantly shorter than face-to-face and interviewer-administered telephonic questionnaires typically are. This is because people can very easily opt out of a mobile self-completion interview, as they do not have the social pressure of a person asking them to participate. Typical survey ethics also require that the length of the survey be clearly communicated at the beginning, and for mobile self-completion surveys it is very easy to ignore an invitation to a survey if it seems like it will take too long.

The length of mobile self-completion is not just limited to the actual number of questions, the questions themselves also need to be short. In the case of an SMS survey, the question – including all answer options – needs to fit into the character limit of a standard SMS (160 characters). Even with a mobile web survey, where there is no character limit, the size of the mobile-phone screen must be considered. Requiring people to scroll too much leads to the drop-off of responses. A design that fits the screen as people would typically use it, i.e. in portrait orientation, is therefore recommended.

These length limitations mean that SMS survey design can be challenging and requires a considered approach. Researchers are forced to be very specific about their information needs, as every character counts.

To ensure the best possible survey design, extensive cognitive testing of the questionnaire items is recommended. Further detailed script checking is essential. A simple error in the survey will cause people to drop out. Given that data collection can happen so quickly, errors might only be picked up when the full sample is completed, as there is no interviewer to report the error.

## Box 1. Potential uses for mobile SMS

### Short, specific modules

A specific topic of interest from an existing survey can be tailored specifically for the mobile survey mode. These modules can focus on specific topics such as digital FI or financial access. Modules developed for mobile will typically need to contain fewer and simpler measures than what is included in traditional face-to-face surveys on these topics but can be used effectively to collect high-level insights and indicators.

### Re-contacting

Given the right circumstances, mobile surveys can be a cost-effective method to supplement and build on insights gathered from traditional demand-side surveys. In a situation where a survey has been administered and the data is revealing an interesting insight or trend but more information on that specific topic would provide further answers, survey respondents can be re-contacted to answer a few additional questions on the topic.

To successfully achieve this, the main survey will need to include the mobile numbers of the respondents and respondents' consent to contact them for further information.

This method is as yet untested, and response rates are likely to be low – so reaching a large audience of people is unlikely. However, this could still be a cost-effective way to gain deeper insights quickly.

### Fast dip-stick surveys

Mobile surveys can be well used to gain quick insights into specific topics. Results will need to be considered with the limitations of the method in mind. However, these can be used for formulating ideas and hypothesis testing without requiring a large-scale investment in data collection.

### Tracking of monitoring and evaluation measures

Monitoring and evaluation can be challenging to manage over time, particularly where performance indicators are required on a regular basis. The affordability of mobile surveys means that monitoring and evaluation modules developed for mobile can be collected more frequently than with other data collection modes.

### In-the-moment measurement

People tend to carry their mobile phones with them as they go about their daily lives. This means that with mobile surveys people can be reached as and where they are. This makes mobile well suited to diary-based research where people can receive multiple survey contacts over a period where they provide information on their financial behaviour.

**SMS surveys underrepresent women, older people, those with less education, and less technologically savvy people. Sending reminders improved representativeness but offering shorter surveys or higher incentives did not.**

Lau, Lombaard, Baker, Eyerman and Thalji (2018)

# 4. Sample design and coverage findings

The strongest challenge to the usefulness of mobile research is lack of sample coverage. Sample coverage is influenced by several factors. The first and clearest barrier to participation in a mobile sample is that participants need to have access to a mobile phone. Mobile phones are widely used, and a growing number of people are entering the mobile world every year, but there are still many people who do not have access to a mobile phone.

Furthermore, the coverage of a mobile survey depends on what mobile numbers are available for surveying. Different modes of mobile surveying have different sampling approaches, and some allow access to more, or less, of the existing mobile market. This can be further limited by regulations that require anyone who is contacted to have opted in to contact – limiting reach to only those who have signed up and preventing sampling that would require cold calling. The sample list used for these pilots will be detailed further later in this section.

Non-response (when people are invited to take part in a survey but either miss the invitation, choose not to participate or forget to participate) is arguably the largest barrier to coverage. Non-response is a greater concern for self-completion surveys because it is easier for people to miss or ignore an invitation that comes to them digitally than if it comes from a person. The survey method with the best response rate is face-to-face, because people are more inclined to agree to take part in a survey if an actual person is standing in front of them and asking; however, telephonic interviews, when the calls are answered, are also more likely to be completed than when an invitation to participate is sent out by SMS.

The technology used for a self-completion mobile survey can also drive up non-response because people may not feel comfortable using it. SMS is widely used, but not everyone who has a mobile phone is necessarily literate enough to take a survey in this way. Further, even though self-completion surveys can be conducted at no cost to the respondent, it is likely that people do not fully trust that this is true, and concerns over the cost of responding may limit some people's participation.

Non-response is also not always due to the actual respondent's decision to ignore the invitation to participate in the survey. It is possible that people who live or work in areas that have less mobile network coverage do not receive the invitation immediately, and they may only be able to take the survey once the surveying period is closed. This can also be true for people with limited access to electricity, who do not keep their phones on all the time. They might only see the invitation once the survey is closed, or even not see the invitation due to it being cluttered amongst a group of other SMSs that have been sent to them while their phones were off (Lau et al., 2018).

Lau et al. (2018) found that driving up response rates improved overall sample representivity however larger incentives, shorter surveys and the time of day

that invitations were sent did not have any influence on non-response. The greatest improvement to non-response was seen when reminders to participate were sent to those who did not respond to the first invitation to participate in the survey.

The pre-pilots conducted in 2017 did not include reminders to drive response rates up and therefore it is likely that the sample achieved in these pre-pilots could have been improved. Reminders will be implemented in the full-scale pilot.

### Box 2. The types of mobile samples available

Random digit dialling (RDD) is a method for selecting respondents by generating telephone numbers at random. RDD has the advantage that it includes numbers that would be missed if the numbers were selected from a list, database or phone book.

While in theory RDD is the ideal sampling method for mobile surveys, when it comes to self-completion mobile surveys there are some practical limitations to the use of the RDD method.

Self-completion surveys have lower response rates than interviewer administered surveys, this low response rate is more pronounced when the contact is cold-calling, as would be the nature of an RDD sample. In addition, while the technology exists for sending surveys to randomly generated lists, agreements need to be made with Mobile Network Operators (MNOs) to ensure that respondents do not incur costs for responding and can receive incentives for their time. This adds an administrative step to managing data collection that is not widely offered by research suppliers.

Database sampling is a method for selecting respondents from an existing or developed database of numbers. Database sample providers offer large lists of numbers for sample selection and can have an existing relationship with MNOs that allows for respondents to respond at no cost and for the paying out of incentives. In a market where the database sample includes all the available numbers at all of the available MNOs, these databases can be used in the same way that an RDD approach would be. MNOs, however, can be reluctant to share their full subscriber bases and some MNOs can be excluded entirely from the database. Due to this lack of access to the full mobile universe database, sampling needs to take a non-probability approach using quota sampling to achieve representivity.

Database sample providers are likely to develop a smaller database audience within the larger database as people complete surveys. This is because they can 1) profile these people based on previous survey responses, thus allowing them to target them for surveys they qualify for in the future and 2) expect higher response rates from people who have been willing to take surveys in the past, thus driving down the cost to complete samples.

A panel sample is a group of people who have opted in to be surveyed. Panels are surveyed more than once and can be surveyed on the same topic over a period or on multiple different topics. Panels can be developed in many ways, but the ideal panel is recruited through a probability sampling approach to form a representative cross-section of the universe. However, this is not always the case, and panels can also become just a collection of people who have signed up. Well-developed and managed panels are rare in the developing world, as the time and logistics required for building and maintaining them are not commercially viable given that the majority of survey research is still conducted ad hoc and face to face.

Dynamic/river sampling is less common in developing markets, because it applies specifically to online sampling. A dynamic sample is one that is collected through the use of adverts calling for survey participants. Someone browsing the internet might see an advert, click on it and be redirected to take part in a survey. Dynamic sampling can be used for mobile-data collection, and the algorithms used to place ads that are seen by a wide cross-section of the population can produce a good representation of an online audience. They are, however, only able to reach people who have internet access.

## 4.1. The sample source

The 2017 pre-pilots were conducted by means of a database sample provider. The sample provider's database is profiled in terms of age, gender and area, and this data, along with their mobile numbers, has been collected directly from the MNOs and was then verified when people completed surveys.

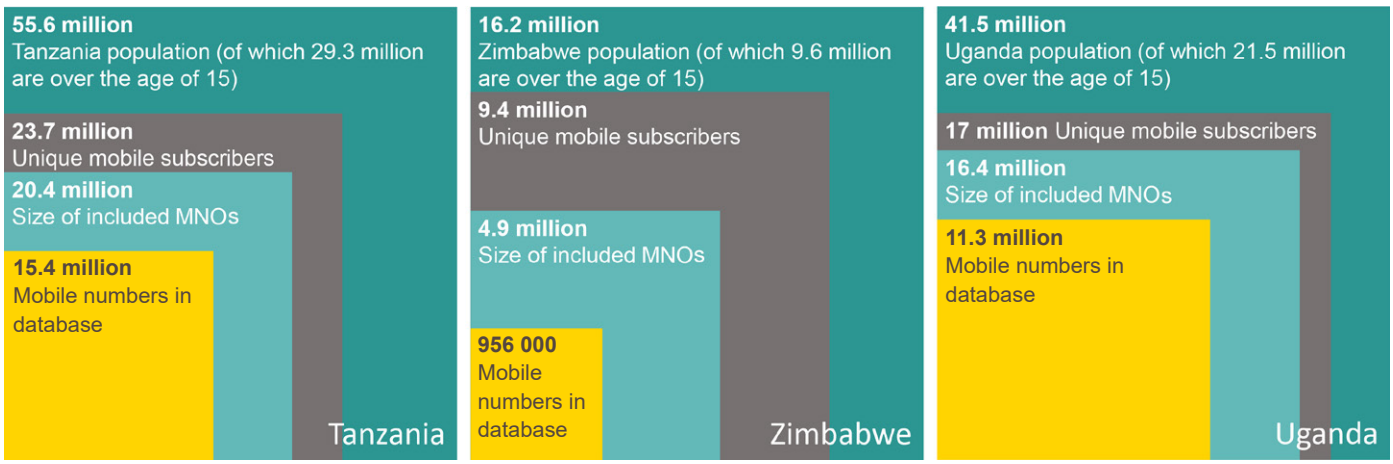
The sample provider has built its database by establishing relationships with the MNOs in each market. MNOs share their subscriber numbers and limited demographic information about them, if it is available, with the sample provider. The sample provider is then linked into each MNO's billing system so surveys can be conducted at no cost to the respondent and incentives can be paid immediately on completion of a survey. The relationship between the sample provider and the MNOs offers a strong benefit to respondents; however, they also lead to a potential compromise in the sample quality.

The first compromise is that the sample provider does not necessarily have partnerships with all MNOs in a particular market. Larger MNOs are typically included, but the exclusion of smaller MNOs can mean that significant portions of the population cannot be reached. Also, different MNOs might service different segments of the population, so with the exclusion of a particular MNO, it is possible that a particular type of person is excluded from the sample frame.

Secondly, the MNOs decide which numbers they will or will not share with the sample provider. Some MNOs willingly provide their entire databases, but others only provide a portion of this and without access to any information about who is and who is not included, it is difficult to control for bias or skews that might arise from this selection.

This means that in Tanzania the database represents around 65% of mobile subscribers, in Zimbabwe around 10% of mobile subscribers and in Uganda around 66% of mobile subscribers.

**Figure 4.1. The drop-off of representation**



## 4.2. The sampling frame

Only a portion of the national universe in each of these markets is represented when using a database sample source. Table 4.2 highlights how the overall database includes certain types of people in abundance but under-represents other types of people.

### Table 4.2. Mobile database profiles

| % | Tanzania | | Zimbabwe | | Uganda | |
|---|---|---|---|---|---|---|
| | Mobile database | Census estimates | Mobile database | Census estimates | Mobile database | Census estimates |
| **Gender** | | | | | | |
| Female | 28 ↓ | 50 | 35 ↓ | 51 | 28 ↓ | 50 |
| Male | 72 ↑ | 50 | 65 ↑ | 49 | 72 ↑ | 50 |
| **Age** | | | | | | |
| 16–20 | 13 ↓ | 19 | 5 ↓ | 18 | 11 ↓ | 22 |
| 21–25 | 39 ↑ | 16 | 15 | 16 | 42 ↑ | 19 |
| 26–30 | 24 ↑ | 14 | 18 | 15 | 24 ↑ | 15 |
| 31–35 | 10 | 12 | 19 ↑ | 14 | 12 | 11 |
| 36–40 | 6 ↓ | 10 | 16 ↑ | 9 | 5 ↓ | 8 |
| 41–45 | 3 ↓ | 8 | 13 ↑ | 7 | 3 ↓ | 6 |
| 46–50 | 2 ↓ | 6 | 8 ↑ | 4 | 2 ↓ | 5 |
| 51–55 | 1 ↓ | 4 | 3 | 4 | 1 ↓ | 4 |
| 56–60 | 1 | 3 | 2 | 4 | 1 | 3 |
| 61+ | 0 ↓ | 8 | 1 ↓ | 9 | 2 ↓ | 5 |
| **Area** | | | | | | |
| Dar es Salaam | 27 ↑ | 10 | - | - | - | - |
| Morogoro | 9 ↑ | 5 | - | - | - | - |
| Pwani | 4 ↑ | 2 | - | - | - | - |
| Harare | - | - | 30 ↑ | 16 | - | - |
| Mashonaland West | - | - | 13 | 11 | - | - |
| Mashonaland East | - | - | 13 ↑ | 10 | - | - |
| Mashonaland Central | - | - | 10 | 8 | - | - |
| Central | - | - | - | - | 54 ↑ | 27 |
| Eastern | - | - | - | - | 18 | 26 |
| Northern | - | - | - | - | 10 ↓ | 21 |
| Western | - | - | - | - | 18 | 26 |

↑↓ Indicates difference greater or less than 3 percentage points.

**The existing gold standard for financial inclusion demand-side surveys is a probability sampling approach, where all people in the desired universe have an equal chance of being selected to take part in the survey.**

The database in all three markets skews towards men living in major metro areas. People under the age of 20 are not well represented, but in Tanzania and Uganda the bulk of people that have been profiled in the database are between the ages of 20 and 35.

While these skews exist in the databases, the databases themselves are sizeable, meaning that in most groups where people are seriously under-represented there is still a significant number of people who can potentially be reached. This allows for a targeted quota sample to be applied that can be made to reflect a demographic profile that is relevant to the research objectives.

## 4.3. Quota sampling

The existing gold standard for financial inclusion demand-side surveys is a probability sampling approach, where all people in the desired universe have an equal chance of being selected to take part in the survey. For more information on this type of sampling, please refer to our survey implementation guide.

In the case of mobile surveys for FI, the desired universe and the sample source do not align – the desired universe is still the total national population of the country of interest; however, the sample source is limited to those who have mobile phones, those who are on the networks that the sample provider has partnerships with and those whose mobile numbers these networks have shared with the sample provider.

Representation of the desired universe will therefore actually be better if quotas are set, because random selection from the sampling database will produce a sample that represents the database and not the desired universe.

A quota sample is a sample formed on the basis of specific attributes derived from a pre-set standard. In this instance we wanted to ensure that demographically the mobile sample mirrored the desired universe as closely as possible.

The database used for sampling has a pre-existing profile that can assist with sending survey invitations to the right people based on the pre-determined quotas. The known variables in the database are age, gender and province/region. We used all of these as quota variables. Furthermore, in Tanzania and Zimbabwe, we applied a quota on income, but this quota could not be managed by using the existing database profile. To control for income, we had to use the actual survey, i.e. people were asked their income. Once a particular quota was met, any further people who answered that their income was within that quota were screened out of the survey.

Typically quota samples can be set up with interlocking quotas, where proportions of certain variables can be defined within others (e.g. a quota could be set for 60 urban males, 40 urban females, 30 rural males and 70 rural females) or as rim quotas where quotas are set only at an overall variable (e.g. 100 urban people, 100 rural people, 90 males and 110 females). Interlocking quotas are likely to produce better results, as they ensure correct proportions of people across multiple variables; however, if too many are set, they can become increasingly difficult to manage and achieve.

While limited interlocking quotas are the ideal, the sampling software available to the sample provider does not readily allow for interlocking quotas, and for these pre-pilots all quotas were treated as rim quotas. On reviewing the achieved samples, the rim quotas were not ideal, and interlocking of some variables is advised.

### Box 3. Quota achievement from a panel or database

In face-to-face, demand-side surveys, quota achievement is typically quite precise. This is because individual enumerators will be given a set list of quotas to achieve, and they are incentivised to not deviate from that. For example, an enumerator might be required to complete 10 interviews with rural women; and once they have found, recruited and interviewed these 10 rural women, their job is considered complete.

In a panel or database sampling setting where sample recruitment is being managed digitally, it is not as straightforward or accurate.

Given that not every person who is invited to participate in a survey necessarily responds, more invitations are sent out than the number of surveys that are completed. This means that if we set a quota for 100 males and 100 females and we expect a response rate of around 30%, approximately 300 men and 300 women will receive invitations to participate.

If on an occasion more or fewer than the estimated number of people respond, we may find that the quota achievement could be higher than what was set. For example, 105 men might respond. Obviously, if the achievement is lower than what is required, more invitations can be sent. If a quota goes over, however, then the additionally achieved surveys will remain in the overall sample. This can provide more data; however, it also makes perfect quota achievement nearly impossible to achieve.

# 5. Sample realisation and comparison to Financial Inclusion Insights and FinScope

## 5.1. Tanzania

The demographic sampling quotas for both pre-pilots in Tanzania were designed using the 2016 Tanzania FII survey (Wave 4). Quota achievement on the liquidity pre-pilot sample was more successful than for the mobile FII DFS pre-pilot sample. This is likely because this was the second survey conducted by using the same sampling frame; and, therefore, where quotas had over- or under-reached, quota targeting could be adjusted for a better result.

Given that the two pre-pilot samples produced very similar results and where there were differences the liquidity sample was actually better, the discussion that follows of the sample skews in Tanzania will focus specifically on the FII DFS mobile pre-pilots, as these best illustrate the possible limitations to mobile sampling in Tanzania.

**Table 5.1.1. Tanzania DFS demographic profile comparison**

| | Mobile DFS pilot | 2016 Tanzania FII[3] |
|---|---|---|
| | n=518 | n=603 |
| **Age** | | |
| 15–24 | 28 ↓ | 35 |
| 25–34 | 46 ↑ | 28 |
| 35–44 | 19 | 18 |
| 45–54 | 5 | 9 |
| 55+ | 2 ↓ | 11 |
| **Gender** | | |
| Female | 45 ↓ | 52 |
| Male | 55 ↑ | 48 |
| **Urbanicity** | | |
| Urban | 82 | 78 |
| Rural | 18 | 22 |
| **Region** | | |
| Dar es Salaam | 57 ↓ | 67 |
| Morogoro | 20 | 21 |
| Pwani | 23 ↑ | 12 |

↑↓ Indicates that mobile result is outside of the margin of error (precision 6.5%). Direction of arrows indicates higher ↑ or lower ↓

A comparison of the demographic breakdown from the FII DFS mobile pilot results against the 2016 Tanzania FII survey results did reveal a similar demographic profile with a few notable skews.

While overall the mobile sample skews to a younger audience, there is a skew away from people aged 15 to 24. Young people and older people are less likely to own mobile phones; but, in addition to this, the largest commercial opportunity for the sample provider is traditional market research, where the majority of surveys are targeted at those aged 18 and over. This likely confounds an under-representation of people aged 15 to 18, and therefore 15 to 24, in the sampling database.

It is unsurprising that people over the age of 55 are also under-represented, as they are less likely to have the technological inclination that might encourage people to participate in mobile surveys (Jeoffreys, 2015).

The mobile sample also skewed to a male audience. Mobile-phone ownership in Tanzania skews very strongly to men, and the gendered profile of the sample database is actually 70:30 male to female. Quotas largely accounted for this, but there is still a slight bias in the sample to men.

The mobile sample skewed away from people in Dar es Salaam and towards people in Pwani. This skew is a direct result of the quotas that were set on the target sample to ensure a large-enough sample in rural areas for analysis.

The 2016 Tanzania FII did not include a variable that could be used to compare claimed personal income; however, this variable was available in the 2017 Tanzania FinScope Survey, which follows a very similar sampling approach to the FII survey.

### Table 5.1.2: Tanzania DFS income comparison

| | Mobile DFS pilot | 2017 Tanzania FinScope[4] |
|---|---|---|
| | n=518 | n=1423 |
| **Personal Monthly Income** | | |
| TSH0–20,000 | 36 ↑ | 17 |
| TSH20,001–166,500 | 30 | 43 |
| TSH166,501–420,000 | 17 ↓ | 25 |
| TSH420,001–820,000 | 11 | 10 |
| TSH820,000+ | 5 | 4 |

↑↓ Indicates that mobile result is outside of the margin of error (precision 6.5%). Direction of arrows indicates higher ↑ or lower ↓

While our expectation was that mobile-survey participants would have a higher income profile, at first glance their claimed income from the survey indicates that mobile reaches a good spread of income groups and even skews to a lower-income group. We treat this finding with caution because it is possible that the lower-income skew is actually driven by the wording of the question.

[3] Filtered on mobile-phone users and appropriate regions

[4] Filtered on mobile-phone users and appropriate regions

**The face-to-face survey asks personal income in a far more comprehensive manner.**

Income was requested in the mobile surveys as shown below:

> Please estimate your total monthly personal income (TSH)?
> 1)0- 20000
> 2)20001- 166500
> 3)166501- 420000
> 4)420001- 820000
> 5)820001+
> 6)Don't know
> 7)Refuse

The face-to-face survey asks personal income in a far more comprehensive manner. It first asks people whether they earn any income from a list of nine possibilities (i.e. salaries/wages, selling products, selling a service, piece work, rental income, interest from investments, pensions, social grants, or receiving money from other people) and then this question is followed up with a question on how much income is earned per income source.

We therefore hypothesise that, without that detailed priming on multiple sources of income, people likely default their answer to income that is collected more formally. This hypothesis is supported when considering income from the 2017 FinScope Tanzania that is earned through salaries or wages or selling products or services only (Table 5.1.3).

### Table 5.1.3. Tanzania DFS formal income comparison

| | Mobile DFS pilot | 2017 Tanzania FinScope "formal income only" |
|---|---|---|
| | n=518 | n=1423 |
| **Personal Monthly Income** | | |
| TSH0–20,000 | 36 | 47 |
| TSH20,001–166,500 | 30 | 26 |
| TSH166,501–420,000 | 17 | 15 |
| TSH420,001–820,000 | 11 | 7 |
| TSH820,000+ | 5 | 3 |

A deeper look at income data by gender shows an interesting pattern that highlights a gap in the mobile sample. On average, women in the mobile sample earn the same amount as men, while women in the face-to-face sample earn on average 68% of what men do.

The similarities in income distribution across genders in the mobile sample indicates that a particular type of woman is being under-represented in the mobile sample – those with the lowest income.

[5] Filtered on mobile-phone users and appropriate regions

This is further evidenced in the urban-to-rural divide by gender, where we see that rural women account for 11% of the face-to-face sample and only 5% of the mobile sample.

While the skews in the mobile sample are not impossible to account for with a weighting framework, this lack of representation of poorer, rural women is an important factor to consider when using mobile surveys for financial inclusion research in Tanzania.

## 5.2. Uganda

For the pre-pilots in Uganda, we opted for a single pilot with a larger sample size. Given this larger sample size, we tested a national sample and did not specifically target certain areas. The national sample provided a good spread across regions in Uganda; however, the sample skews very strongly to urban areas.

**Table 5.2. Uganda DFS demographic profile comparison**

| | Mobile DFS pilot | 2016 Uganda FII[5] |
|---|---|---|
| | **n=1,135** | **n=3,000** |
| **Age** | | |
| 15–24 | 38 | 40 |
| 25–34 | 31 ↑ | 25 |
| 35–44 | 14 | 16 |
| 45–54 | 14 ↑ | 9 |
| 55+ | 4 ↓ | 11 |
| **Gender** | | |
| Female | 48 | 47 |
| Male | 52 | 53 |
| **Urbanicity** | | |
| Urban | 63 ↑ | 24 |
| Rural | 36 ↓ | 76 |
| Don't know | 1 | - |
| **Region** | | |
| Central I | 6 ↓ | 12 |
| Central II | 5 ↓ | 10 |
| Kampala | 16 ↑ | 6 |
| West / South West | 22 | 24 |
| Eastern | 17 | 15 |
| East Central | 7 | 10 |
| North / Karamoja / West Nile | 27 ↑ | 22 |

↑↓ Indicates that mobile result is outside of the margin of error (precision 4.5%). Direction of arrows indicates higher ↑ or lower ↓

In Zimbabwe and Tanzania, age quotas were set as three groups (15–24, 25–34 and 35+). In Uganda we split the top quota to be 35–45 and 45+. This appears to have a positive effect on the age spread with a good proportion of 45–54-year-olds being represented in Uganda and some people over the age of 55 still being included.

## 5.3. Zimbabwe

In Zimbabwe, we have two face-to-face data sources that can be used for sample comparison: the FinScope 2014 survey data and our 2017 face-to-face needs measurement pilot. Neither one of these datasets is a perfect point for comparison, because the FinScope data is a few years out of date and the needs measurement pilot was not recruited as a random sample but was rather recruited as a combination of a convenience sampling approach and from lists provided by a credit bureau.

### 5.3.1. FinMark Trust FinScope Access Strand pilot

Given that the access strand pilot was aimed at replicating some of the traditional FinScope measures, we have compared the sample achievement of this pilot against the 2014 FinScope sample. The resilience measurement was part of the needs measurement that was also piloted face to face in 2017. We have compared the sample achievement of the resilience pilot against the needs measurement pilot data.

[6] Filtered on mobile-phone users and appropriate regions

### Table 5.3.1: Zimbabwe Access demographic profile comparison

| | Mobile access pilot | 2014 FinScope Survey[6] |
|---|---|---|
| | n=556 | n=1,423 |
| **Age** | | |
| 18–24 | 15 | 17 |
| 25–34 | 30 | 27 |
| 35–44 | 36 ↑ | 23 |
| 45–54 | 16 | 14 |
| 55+ | 3 ↓ | 19 |
| **Gender** | | |
| Female | 50 ↓ | 57 |
| Male | 50 ↑ | 43 |
| **Urbanicity** | | |
| Urban | 59 ↑ | 36 |
| Rural | 36 ↓ | 64 |
| Don't know | 5 | - |
| **Region** | | |
| Mashonaland East | 24 | 26 |
| Mashonaland Central | 17 | 20 |
| Mashonaland West | 25 | 24 |
| Harare | 35 | 30 |
| **Personal Monthly Income (USD)** | | |
| No income | 37 ↑ | 7 |
| USD1–100 | 16 ↓ | 53 |
| USD101–200 | 14 | 12 |
| USD201–400 | 21 ↑ | 11 |
| USD401–1,000 | 11 | 5 |
| USD1,001–2,500 | 0.4 | 1 |
| USD2,501+ | 0.4 | 0 |
| Refused | - | 1 |
| Don't know | - | 9 |

↑↓ Indicates that mobile result is outside of the margin of error (precision 6.0%). Direction of arrows indicates higher ↑ or lower ↓

**The database used to source the sample in Zimbabwe is significantly smaller than the databases in other markets, therefore the skew that exists in the panel profile to more urban people is more difficult to avoid when setting up quotas.**

Similar to skews seen in Tanzania and Uganda, the mobile access sample in Zimbabwe skews to a younger profile, and there is an appearance of a lower-income skew that is likely driven by how the question is primed and asked.

The database used to source the sample in Zimbabwe is significantly smaller than the databases in other markets, therefore the skew that exists in the panel profile to more urban people is more difficult to avoid when setting up quotas.

There were no urban-to-rural quotas set for these pre-pilots as targeting urban or rural is not a pre-profiled variable in the database. Applying this quota will add time and cost to the research, but its addition will lead to an improvement in the overall sample profile of an SMS survey.

### 5.3.2. Resilience pilot

As both the face-to-face needs measurement pilots and the mobile resilience pilot had limited sample sizes, they were both conducted only in certain areas of Zimbabwe. These areas were not identically matched. For this reason, to compare the two we have filtered both datasets on their matching areas, Harare and Mashonaland West.

The resilience pilot further confirms a skew away from older people and an urban bias in the Zimbabwe database.

The Zimbabwe mobile sample also appears to skew to a lower-income profile; however, it is likely that this is driven by wording of the question, with no priming on multiple sources of income.

However, the formal income profile in the face-to-face Financial Needs and Usage Pilot is still higher than that of the mobile resilience pilot. This is to be expected, given that the sampling approach used credit bureau lists as the primary sample source – meaning that everyone in the sample has access to formal credit.

Like in Tanzania, there does appear to be an under-representation of rural women in the Zimbabwean mobile samples in Zimbabwe.

[7] Filtered on Harare and Mashonoland West

### Table 5.3.2.1. Zimbabwe Resilience demographic profile comparison

| | Mobile resilience pilot | 2017 Zimbabwe Financial Needs and Usage Pilot[7] |
|---|---|---|
| | n=315 | n=445 |
| **Age** | | |
| 18–24 | 19 | 14 |
| 25–34 | 29 | 29 |
| 35–44 | 45 ↑ | 33 |
| 45–54 | 7 ↓ | 16 |
| 55+ | - ↓ | 9 |
| **Gender** | | |
| Female | 56 | 59 |
| Male | 44 | 41 |
| **Urbanicity** | | |
| Urban | 81 ↑ | 64 |
| Rural | 19 ↓ | 36 |
| Don't know | - | - |
| **Region** | | |
| Mashonaland West | 39 | 46 |
| Harare | 61 | 54 |
| **Personal Monthly Income (USD)** | | |
| No income | 38 ↑ | 4 |
| 1–100 | 18 | 15 |
| 101–200 | 13 | 13 |
| 201–400 | 22 | 27 |
| 401–1,000 | 9 ↓ | 29 |
| 1,001–2,500 | 1 ↓ | 8 |
| 2,501+ | - | 6 |

↑↓ Indicates that mobile result is outside of the margin of error (precision 8.0%). Direction of arrows indicates higher ↑ or lower ↓

### Table 5.3.2.2: Zimbabwe Resilience formal income profile

| | Mobile resilience pilot | 2017 Zimbabwe Financial Needs and Usage Pilot 4 | 2017 Zimbabwe Financial Needs and Usage Pilot "formal income only" [8] |
|---|---|---|---|
| | n=315 | n=445 | n=445 |
| **Personal Monthly Income (USD)** | | | |
| No income | 38 | 4 | 17 |
| 1–100 | 18 | 15 | 9 |
| 101–200 | 13 | 13 | 10 |
| 201–400 | 22 | 27 | 23 |
| 401–1,000 | 9 | 29 | 27 |
| 1,001–2,500 | 1 | 8 | 8 |
| 2,501+ | - | 6 | 6 |

[8] Filtered on Harare and Mashonoland West. Only salaries, wages and money from farming and pensions included. Piece work could not be separated from salaries and wages, so it is included.

### Box 4. Response rates and respondent drop-off

Survey sample representation needs to be understood both in terms of the profile of those who respond and ideally the profile of those who do not respond.

Knowing who the people are that do not respond is not necessarily possible; however, monitoring of response rates gives an indication of the size of the potential gap between those who do respond and those who don't. A breakdown of response rates achieved in these pilots is detailed in Table 5.3.2.3.

#### Table 5.3.2.3. Response rates

| | Tanzania DFS | Tanzania Liquidity | Uganda DFS | Zimbabwe Access | Zimbabwe Resilience |
|---|---|---|---|---|---|
| Opted in | 13% | 14% | 13% | 13% | 30% |
| Completed | 7% | 7% | 7% | 6% | 16% |

Response rates were low for most of the pre-pilots, and only half of those who opted in actually completed the survey. A strategy to increase response rates and the rates of completion is critical to the success of future mobile SMS surveys.

Drop-off throughout the survey is inevitable; however, there are certain questions that prompt a higher drop-off than others, and an understanding of this can lead to more people completing the survey. Questions that are considered too personal and questions that are confusing are the most likely causes of respondent drop-off.

For most of the pilots, the significant majority of drop-off occurs in the first two or three questions. This is likely caused by people opting in but then finding the task unappealing when actually engaging in it. This early drop-off is impossible to avoid.

There were, however, several questions later in the surveys that showed spikes in respondent drop-off. Questions that showed high drop-off included questions around personal and household income – these questions probably made some people feel uncomfortable. Questions on mobile money, micro-finance and insurance also led to high drop-off. Cognitive testing on these types of questions has revealed that people do not understand them well, and this lack of understanding is likely what leads to respondent drop-off. For greater detail on the cognitive tests conducted, refer to our Cognitive Testing Note.

**While quota selection can account for some of the skews that occur, there are certain limitations in the database sampling frame that still influence the profile of a completed mobile sample.**

# 6. Considerations and recommendations for sampling

## 6.1. Reach and exclusion

Using a mobile database, while it is highly convenient and allows for superior mobile survey design, does limit the representation that can be achieved. Beyond simply excluding those people that do not have mobile phones, a database is limited by where it comes from. In the case of these pre-pilots, the limitations to representation happen across multiple stages:

- Only those with mobile phones are included.
- Only subscribers whose MNO has a partnership with the sample provider are included.
- Some MNOs share only a portion of their databases.
- Not everyone who is invited to participate completes the survey, and rejection of digital invites is higher than the rejection of in-person recruitment in more traditional research.

While quota selection can account for some of the skews that occur, there are certain limitations in the database sampling frame that still influence the profile of a completed mobile sample. Most notably mobile samples do not reach many older people, there is an urban skew and, while it is inconclusive, there may also be a skew to people with a higher personal income.

## 6.2 Achieving better representation

Achieving the best possible representation requires stricter quota control than was applied in these pre-pilots. An interim quota check is recommended. This should consider not only how the over-arching rim quotas look (i.e. the age profile or the gender profile) but should also consider how these interlock to avoid a sample that seems representative but actually over-samples some types of people to account for an under-sample of another, for example women in rural areas may be under-sampled but men in these areas are over-sampled and women in urban areas are also over-sampled.

We also recommend a compromise on field timings so that those people who are not as quick to respond are still included. Invitations to participate should be sent out in smaller batches while allowing people more time to participate. This will extend field times but will allow people who are less active on their mobile phones the time to be a part of the completed sample. Before moving on to a second batch of invites, a reminder message should be sent to those who do not respond to the first invitation to ensure that not only quick responders are included.

## 6.3 Adjusting for sampling limitations

The universe that we are able to sample (i.e. those with mobile phones) is far more likely to be included financially than those whom we cannot reach with the mobile methodology. Additionally, urbanicity plays a significant role in financial inclusion, and the mobile samples generally under-represent those people who live in rural areas.

Stricter quota control should improve representation; however, there are certain types of people that are more difficult to reach and therefore representative quotas are unrealistic, e.g. rural people over the age of 55. For this reason, we recommend setting realistic quotas for these people and then using weighting to amplify their voices in the final dataset. The use of a weighting structure can also account for the level of instability that occurs due to response rate variation across multiple waves of data collection.

While weighting is typically reserved for probability sampling, in the next phase of mobile piloting a multi-level regression and post-stratification model will be applied to weight the quota samples.

# 7. Comparison of key indicators and exploring the gaps

## 7.1. Digital Financial Services modules – Tanzania and Uganda results observations

The discussion of the 2017 pilot results will primarily focus on the results achieved in the DFS pilots in Tanzania and Uganda. These two pilots clearly demonstrate both the promise and the limitations of mobile SMS data collection. The face-to-face benchmarks for the access, liquidity and resilience measurement pilots provide for only limited comparison.

In 2013, the Bill & Melinda Gates Foundation conceived the FII programme in partnership with Intermedia to build meaningful knowledge about the financial landscape in eight countries across Africa and Asia: Kenya, Nigeria, Tanzania, Uganda, Bangladesh, India, Indonesia and Pakistan (Intermedia, 2018). The FII survey serves as a critical monitoring and evaluation tool for the Bill & Melinda Gates Foundation.

To track progress on their financial inclusion goals (using the FII survey), the Bill & Melinda Gates Foundation (BMGF) selected the key indicators below:

**Figure 7.1.1. BMGF FSP indicators (surveys of adults [15+])**

| |
|---|
| Can access a mobile phone (own or borrow) |
| Registered account owners |
| Active registered account owners |
| Own a phone |
| Have ever accessed a digital stored-value account |
| Actively using registered digital stored value accounts (on a 90-day basis)<br>• <USD2.5/day adults actively using registered digital stored-value accounts<br>• Females actively using registered digital stored-value accounts |
| Actively use digital stored-value accounts and have used at least one advanced financial service (beyond basic wallet and P2P)<br>• <USD2.5/day adults who actively use digital stored-value accounts and have used at least one advanced financial service (beyond basic wallet and P2P)<br>• Females who actively use digital stored-value accounts and have used at least one advanced financial service (beyond basic wallet and P2P) |
| Have ever accessed mobile money |
| Active mobile money account (on a 90-day basis) |

[9] Filtered on those who have access to a mobile phone and are living in Dar Es Salaam, Morogoro or Pwani

Our DFS pilot modules in Tanzania and Uganda aimed to replicate this list of indicators using an SMS methodology. The results do not perfectly replicate the face-to-face measurement of these indicators, but the comparison is promising for the future of mobile surveys as a part of a sustainable data collection programme for the FSP indicators.

### Tanzania

In Tanzania, the overall trends appear consistent across mobile SMS and face-to-face data; however, there are some notable differences. These differences are driven by both the sample and the survey design.

Mobile SMS produced low measurement for:

- ever accessing a digital stored-value account
- actively using a digital stored-value account with at least one advanced financial service
- having accessed mobile money

These lower scores are likely driven by the design of the survey.

**Table 7.1.1. Tanzania BMGF FSP indicators (surveys of adults [15+])**

| | Mobile n=518 | Face-to-face[9] n=603 |
|---|---|---|
| Can access a mobile phone (own or borrow) | 100 | 100 |
| Registered account owners | 75 | 73 |
| Active registered account owners | 66 | 61 |
| Own a phone | Unknown | 80 |
| Have ever accessed a digital stored-value account | 70 ↓ | 81 |
| Actively using registered digital stored value accounts (on a 90-day basis) | 60 | 61 |
| • <USD2.5/day adults actively using registered digital stored-value accounts | Unknown | 53 |
| • Females actively using registered digital stored-value accounts | 64 ↑ | 52 |
| Actively use digital stored-value accounts and have used at least one advanced financial service (beyond basic wallet and P2P) | 38 ↓ | 51 |
| • <USD2.5/day adults who actively use digital stored-value accounts and have used at least one advanced financial service (beyond basic wallet and P2P) | Unknown | 42 |
| • Females who actively use digital stored-value accounts and have used at least one advanced financial service (beyond basic wallet and P2P) | 39 ↓ | 42 |
| Have ever accessed mobile money | 68 ↓ | 80 |
| Active mobile-money account (on a 90-day basis) | 54 | 60 |

↑↓ Indicates that mobile result is outside of the margin of error (precision 8.0%). Direction of arrows indicates higher ↑ or lower ↓

Using the face-to-face survey to construct these measures allows for them to be constructed from multiple variables, while in the mobile survey they are asked far more directly. For example, access to mobile money is asked at an intuitional level when face to face but as a single binary question via SMS, as illustrated in Figure 7.1.2.

**Figure 7.1.2. Comparison of access to mobile-money questions face to face and mobile SMS**

| Face to face | Mobile SMS |
|---|---|
| MM4. ASK IF AT LEAST ONE PROVIDER MARKED IN MM2 OR MM3. OTHERS SKIP TO MMP1. Have you ever used this mobile money service for any financial activity?<br>MTN Mobile Money<br>Airtel Money<br>M-Senta<br>Ezee Money<br>Vodafone M-PESA<br>Africell money<br>Safaricom M-PESA<br>Other (Specify) | Have you ever used your own OR someone else's account from a mobile money provider e.g. M-cash, MTN, Airtel etc.?<br>1)Yes<br>2)No<br>3)Don't Know |

There is no prompting on specific institutions for SMS respondents. People who are less engaged with mobile money because they used it long ago, for a very brief period or infrequently are more likely to forget that they have ever done so. Those who are more active or more recent users are more likely to claim they have ever done so. This is confirmed, as we see the claim for active usage shows no significant difference in the mobile SMS sample against the face-to-face sample.

The lower measurement of ever using mobile money compounds then to a lower usage of a digital stored-value account, as mobile-money usage is a core variable of this.

Similarly, in the face-to-face survey, advanced financial services are asked at a service level but in a singular variable on mobile SMS, as follows:

> Does the main bank you use or used offer at least one of: savings, money transfer, insurance or investment services?
>
> 1)Yes
>
> 2)No
>
> 3)Don't know

This leads to lower claim in the mobile sample.

[10] Filtered on those who have access to a mobile phone

[11] The PPI index could not be included in the 2017 pilots and therefore this indicator cannot be replicated.

The differences in responding to the mobile survey in Tanzania are largely driven by differences in how the questions were asked. The sample achieved produced a very similar profile, and therefore the only notable difference driven by sampling is that the women reached in the mobile sample are more likely to have a digital stored-value account with advanced financial services. Poorer rural women are under-represented in the sample, and with a boost on that target audience we expect to see more consistent results.

As a general rule, mobile samples do skew to urban people; and our sample in Tanzania, being targeted to only three states (with one of them being Dar es Salaam, which is completely urban), meant that getting a close match in terms of geo-location was more achievable. In Uganda, where we went for a national sample, the urban skew is far more highly pronounced – and this sample skew is highly evident in the achieved results.

### Uganda

The Ugandan mobile sample, when compared to those who have access to mobile phones in the face-to-face sample, produces significantly higher scores on every measure. The people reached by means of the mobile SMS method are quite clearly a different profile to those reached with the face-to-face method. The core driver of this is that in the face-to-face sample, 76% of people are living in rural areas; however, in the mobile sample, only 36% claim to be living in rural areas.

**Table 7.1.2. BMGF FSP indicators (surveys of adults [15+])**

| | Mobile n=1135 | Face-to-face[10] n=2254 |
|---|---|---|
| Can access a mobile phone (own or borrow) | 100 | 100 |
| Registered account owners | 76 ↑ | 51 |
| Active registered account owners | 75 ↑ | 44 |
| Own a phone | Unknown | 67 |
| Have ever accessed a digital stored-value account | 82 ↑ | 67 |
| Actively using registered digital stored-value accounts (on a 90-day basis)<br>• <USD2.5/day adults actively using registered digital stored-value accounts<br>• Females actively using registered digital stored-value accounts | 70 ↑<br>Unknown[11]<br>68 ↑ | 42<br>25<br>34 |
| Actively use digital stored-value accounts and have used at least one advanced financial service (beyond basic wallet and P2P)<br>• <USD2.5/day adults who actively use digital stored-value accounts and have used at least one advanced financial service (beyond basic wallet and P2P)<br>• Females who actively use digital stored-value accounts and have used at least one advanced financial service (beyond basic wallet and P2P) | 59 ↑<br>Unknown[4]<br>58 ↑ | 22<br>10<br>15 |
| Have ever accessed mobile money | 71 ↑ | 66 |
| Active mobile money account (on a 90-day basis) | 61 ↑ | 42 |

↑↓ Indicates that mobile result is outside of the margin of error (precision 8.0%). Direction of arrows indicates higher ↑ or lower ↓

Comparing the urban-only face-to-face sample to the SMS results produces a far more consistent story.

### Uganda – urban

The Ugandan results highlight a critical and challenging problem for mobile surveys. Using the mobile SMS method as the only source of data collection in some markets is unlikely to provide sufficient reach among those who are the most likely to be financially excluded. Reaching rural populations is critical to develop a realistic view of FI; however, they are under-represented in mobile-survey samples.

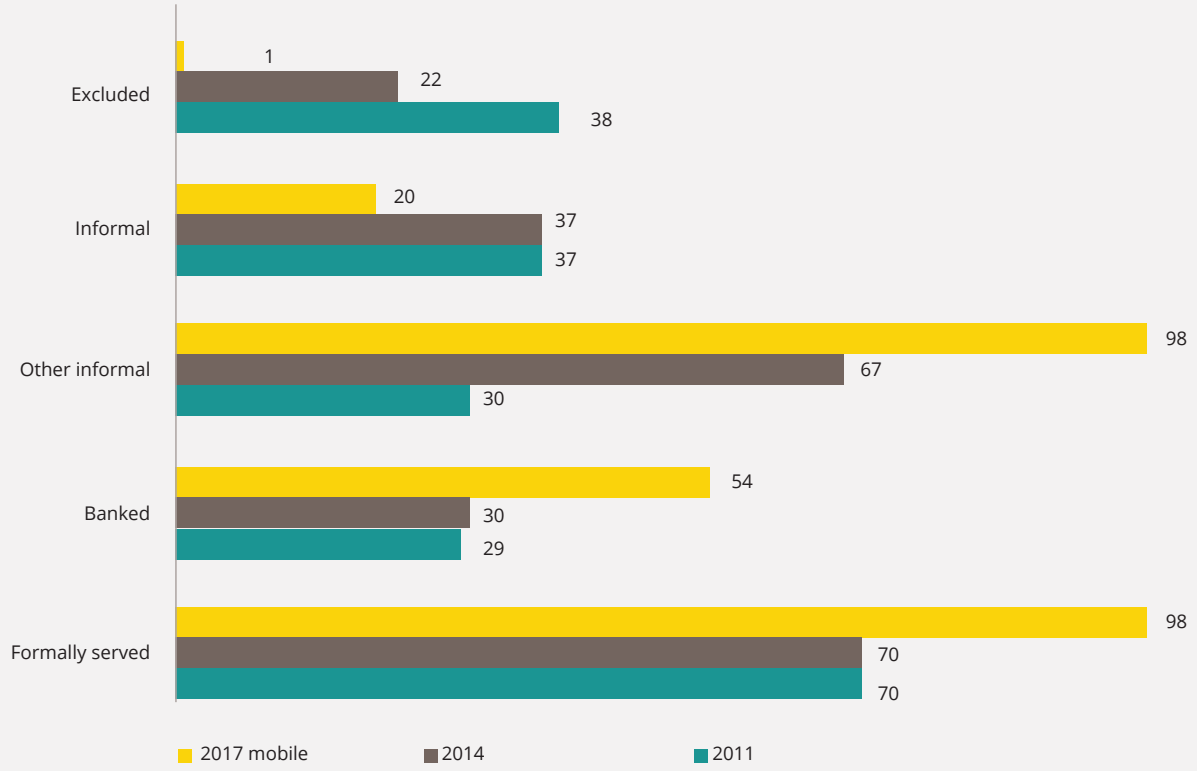
**Table 7.1.3: Urban Uganda BMGF FSP indicators (surveys of adults [15+])**

| | Mobile n=1135 | Face-to-face[12] n=650 |
|---|---|---|
| Can access a mobile phone (own or borrow) | 100 | 100 |
| Registered account owners | 76 | 68 |
| Active registered account owners | 75 ↑ | 61 |
| Own a phone | Unknown4 | 80 |
| Have ever accessed a digital stored-value account | 82 | 83 |
| Actively using registered digital stored value accounts (on a 90-day basis)<br>• <USD2.5/day adults actively using registered digital stored-value accounts<br>• Females actively using registered digital stored-value accounts | 70<br>Unknown4<br>68 ↑ | 62<br>38<br>54 |
| Actively use digital stored-value accounts and have used at least one advanced financial service (beyond basic wallet and P2P)<br>• Poor who actively use digital stored-value accounts and have used at least one advanced financial service (beyond basic wallet and P2P)<br>• Females who actively use digital stored-value accounts and have used at least one advanced financial service (beyond basic wallet and P2P) | 59 ↑<br>Unknown<br>58 ↑ | 37<br>19<br>27 |
| Have ever accessed mobile money | 71 | 83 |
| Active mobile money account (on a 90-day basis) | 61 | 60 |

↑↓ Indicates that mobile result is outside of the margin of error (precision 8.0%). Direction of arrows indicates higher ↑ or lower ↓

[12] Filtered on those who have access to a mobile phone and are living in urban areas

## Box 5. Observations on Access and Measurement frameworks pilot results




In Zimbabwe, we sampled a largely urban sample from a comparatively small database of mobile numbers. This sample skew meant that the financially excluded population was not represented in the mobile data.

Mobile SMS surveys in Zimbabwe are currently limited to a focus on formal inclusion.

Our mobile measurement frameworks tests did not have representative face-to-face data benchmarks for comparison. It is expected that the sample limitations that exist in other tests will have a similar impact on how these pre-pilots represent the overall population. This being said, the pre-pilots are considered successful in terms of building a questionnaire designed to capture needs indicators by using a very short and simple survey design. Key indicators can be gained by using a mobile survey.

Best practice with regard to measurement frameworks remains face-to-face research where all elements of the needs measurement tool can be included in detail. Mobile is likely to best serve to track key performance indicators on these over time, following a detailed face-to-face benchmark survey. For detail on implementation of the needs measurement frameworks, refer to our toolkit.

# 8. Implications of the pilot results

Critical implications from the results of the 2017 surveys are being used to develop the next steps of insight2impact's sustainable data collection pilot programme. These are as follows:

- The reach of mobile surveys is limited. Strategies to boost response rates, especially among rural people, need to be explored and implemented.
- The financially excluded are present in mobile samples, but there are few of them. Post-stratification weighting can be used to amplify their voices.
- Even with a clear plan to boost response rates, there is likely to still be under-representation of certain people in a mobile sample. Weighting is unlikely to be able to fully account for this lack of representation. Sample blending with other, more representative modes of data collection is necessary to take advantage of cost saving where possible but to still ensure that everyone has a voice in the data.
- Mobile survey design can lead to under-claim, but specific behavioural questions track well against face-to-face surveys. SMS survey design needs to be approached in a way that has a stronger emphasis on specific behavioural measures, and we need to conduct detailed testing of the survey to ensure that questions are properly understood.

Beyond the direct implications of the 2017 pilot results, there are additional design considerations that we are paying special attention to.

### Poverty Probability Index

Income measurement in the 2017 mobile pilots produced difficult-to-interpret results due to our inability to include in the income batch of questions the same level of detail that would be included in a typical face-to-face financial inclusion survey.

The face-to-face FII survey includes the Poverty Probability Index (PPI) generated by Innovation for Poverty Action as a core measure in the survey. The PPI is a statistically sound, yet simple-to-use measure of poverty. It uses the answers to approximately 10 questions, depending on the market, about a household's characteristics and asset ownership. These are scored to compute the likelihood that the household is living below the poverty line (Innovation for Poverty Action, 2018).

The variables that are typically used for poverty measurement are unfortunately too numerous to be practical on an SMS survey, and we therefore require a short-form poverty measurement. To this end, for our next wave of pilot tests we are working with Innovation for Poverty Action to develop short-form, five-question PPI indicators. These will be designed with the mobile method in mind but can thereafter even be applied in face-to-face surveys.

### Urban or rural measurement

A core design challenge is the problem of measuring where people are located, i.e. are they rural people or are they urban people. In a traditional survey this is typically measured by an enumerator who already knows whether the location of the survey is considered urban or rural, or the data is collected using GPS technology on the data-collection device, e.g. a tablet. Further definitions of urban and rural are based on what the statistics office defines them as, and this may not match with what people define as urban or rural in common language.

With SMS surveys, this variable depends entirely on respondent claim and needs to be based on a question that people will be able to answer correctly. In the 2017 pre-pilots, we tested two different ways of asking urbanicity.

In Tanzania urban/rural was determined through three questions:

> In which region do you live? 1)Dar es Salaam 2)Morogoro 3)Pwani 4)Another region
>
> Do you live in the city of Morogoro? 1)Yes 2)No, I live elsewhere in the region
>
> Do you live in Kibaha? 1)Yes 2)No, I live elsewhere in the region

If the respondent chose Dar es Salaam in the first question, they were coded as urban and skipped the next two questions. If they mentioned Morogoro, they then saw the question specific to the city of Morogoro; and if they mentioned Pwani, they saw the question specific to Kibaha. All those who mentioned yes to the city questions were also coded as urban, and those who said no to these questions were coded as rural.

In Tanzania then, only those in the main city of each region are being classified as urban, and there are likely people in towns who are being classified as rural. This suggests that perhaps the self-reporting method is a more accurate way to measure urbanicity; however, it does place greater power in the hands of the respondent in terms of their classification.

In Zimbabwe and Uganda, urban/rural was determined by asking:

> Which best describes where you live?
>
> 1)City
>
> 2)Town
>
> 3)Village
>
> 4)Farm land or countryside
>
> 5)Other
>
> 6)Don't know.

Those who selected city or town were classified as urban.

Both of these ways of asking seemed to provide fairly accurate results, and the skews seen in the sample towards urban people in Uganda and Zimbabwe are reflected in answers to the rest of the survey being more similar to the urban face-to-face samples – highlighting that it is not likely that the way of asking the question led to people classifying themselves incorrectly.

This being said, neither of these ways of asking can be validated, and each fails to be an ideal solution going forward.

While the way in which urbanicity was asked in Tanzania is more likely to be accurate as it is less open to the interpretation of the person taking the survey, it is not practical in most markets. The question was possible in the Tanzania pilot because we were limited to three regions, and these regions had a few clear urban areas. In a market like India where there are 29 states or Uganda where there are 56 districts and within those states and districts there can be multiple locations that classify as urban, the question becomes impossible to administer on the mobile mode.

The way in which the question was asked in Uganda and Zimbabwe is more open to interpretation and therefore people may not accurately classify themselves.

Our ideal strategy to overcoming this problem is, along with our sample provider, to engage with the MNOs to release the location of the mobile tower used either most often or at the time of the survey by each person that participates. If we can influence the MNOs to allow this, the location information will only be available once the survey is completed, so it cannot be used for quota filling. This will be our most accurate way of using this data for post-stratification weighting of data.

A common-language question that is clearly understood and produces good results is still required in the survey in order to get as close an estimate of urban and rural as we can from respondent claim. To this end, insight2impact has conducted some qualitative cognitive testing on the survey to get to a question that is most universally understood as an accurate reflection of urban and rural. The best measure tested is:

> Reply with 1 or 2:
> 1)My home is in a village
> 2)My home is in a city

Ideally, if we successfully influence the MNOs to share location data with us, we can then validate the respondent claimed questions against what the mobile tower location information tells us. This will determine whether it is indeed a good-enough measure that can be applied even in markets where we have been unsuccessful in our engagement of the MNOs.

# 9. Conclusion: SMS mobile has limitations but can be a viable solution

The 2017 pilots revealed several limitations to the mobile methodology; however, the potential for mobile SMS surveys as a sustainable data collection method is promising. With a view towards the development of a credible and long-term mobile SMS data collection solution, insight2impact is continuing to pilot SMS surveys.

The next phase of piloting includes new waves of data collection in Bangladesh, China, India, Indonesia, Kenya, Nigeria and Pakistan and conducting a second wave in Tanzania and Uganda. The pilots cover the FII DFS module specifically.

Taking what we have learned from the 2017 pre-pilots into the next phase of piloting includes several steps aimed at ensuring success.

As a starting point, cognitive testing has been conducted on the survey questions to ensure that they are well understood. In addition, insight2impact is working with Innovation for Poverty Action to produce a short-form PPI for the mobile survey.

Lau et al. (2018) found that the inclusion of reminders to participate has a positive effect on representation. To this end, insight2impact is implementing a reminders strategy that ensures that everyone who is invited to take part in the survey but does not respond immediately is prompted to take part in the survey up to three additional times.

Further, the pilots will include both the SMS self-complete mode of data collection and an interviewer-administered telephonic approach. The majority of these markets also have an existing FII face-to-face survey that was conducted in 2018 or 2017.

Using these three data-collection methods, we will build a detailed plan and model for mixing the data collected with these different modes. The aim is to rely on the cheapest mode (mobile SMS) for the bulk of the sample while ensuring sufficient reach of the excluded population using telephonic and face-to-face interviews where necessary. This optimal mix of data collection methods, along with the appropriate modelling to weight the data, can then be applied to future waves of data collection, making them more affordable yet sufficiently robust in terms of reach.

With regard to the modelling to weight the data, the anticipated approach to addressing the non-representativeness of SMS data involves multilevel regression and poststratification (Ghitza & Gelman, 2013; Gelman, A., Goel, S., Rothschild, D and Wang, W., 2017). Multilevel regression and poststratification are widely used across the social sciences to adjust non-representative data. The use of these relies on external data from representative surveys to gather population characteristics and then the SMS sample can be corrected to match the appropriate population proportion.

The use of multilevel regression and poststratification will be applied to SMS data alone, as well as to a mix of SMS data and telephonic data. If necessary, the test may also include a small sample of face-to-face data. Using this approach, we aim to determine whether the modelling approach can allow for SMS data collection on its own. In the likely event that this is not sufficient, we will determine what the optimal mix of data collection modes is to produce reliable results at the lowest possible cost.

# 10. Bibliography


Bhat, A. (n.d.) Types of sampling for social research. QuestionPro. Available: https://www.questionpro.com/blog/types-of-sampling-for-social-research/

Gelman, A., Goel, S., Rothschild, D and Wang, W. (2017) High-Frequency Polling with Non-Representative Data. *In Political Communication in Real Time: Theoretical and Applied Research Approaches*, Dan Schill, Rita Kirk and Amy E. Jasperson (eds). New York, NY: Routledge.

Ghitza, Y. and Gelman, A. (2013) Deep Interactions with MRP: Election Turnout and Voting Patterns among Small Electoral Subgroups. *American Journal of Political Science*, 57(3): 762–776.

GSMA. (2018) The mobile economy: Sub-Saharan Africa 2018. Available: https://www.gsma.com/r/mobileeconomy/sub-saharan-africa/

Intermedia. (2018) *About Financial Inclusion Insights*. Available: http://finclusion.org/about/

Innovation for Poverty Action. (2018) About the PPI: *A Poverty Measurement Tool*. Available: https://www.povertyindex.org/about-ppi

Jeoffreys, S. (2015) The influence of "joiners" and "techies": *Improving the representation of mobile research through opt-in sample control and calibration*. Presented at SAMRA 2015 conference, Drakensberg.

Lau, C., Lombaard, A., Baker, M., Eyerman, J. and Thalji, L. (2018). How representative are SMS surveys in Africa? Experimental evidence from four countries. *International Journal of Public Opinion Research*, 8. DOI: 10.1093/ijpor/edy008

Robertson, G. and Jeoffreys, S. (2017) Sustainable data collection: mobile modes. Available: https://cenfri.org/publications/sustainable-data-collection-mobile-modes/

# How to find us:

Get involved. Contact us.

+27 21 913 9510

**i2ifacility.org**

## Join the conversation:

**@i2ifacility**
**/insight2impact**
**/insight2impact**
**/i2ifacility**

**About Cenfri**
Cenfri is a global think-tank and non-profit enterprise that bridges the gap between insights and impact in the financial sector. Cenfri's people are driven by a vision of a world where all people live their financial lives optimally to enhance welfare and grow the economy. Its core focus is on generating insights that can inform policymakers, market players and donors who seek to unlock development outcomes through inclusive financial services and the financial sector more broadly. For more information, visit www.cenfri.org

**About FinMark Trust**
FinMark Trust is an independant non-profit trust. Its purpose is "Making markets work for the poor, by promoting financial inclusion and regional financial integration". Its programmes aim to unlock financial inclusion and sector development through a symbiotic relationship between rigorous data collection and research activities. Its work can be found in South Africa, throughout the SADC region and the global arena. For more information, visit www.finmark.org.za.

Established and driven by

cenfri

FMT FINMARK TRUST

Sponsored by

BILL & MELINDA GATES foundation

mastercard foundation