TRANSACTION COST INDEX

A Toolkit for Measuring the Real Costs of Digital Financial Services

Innovations for Poverty Action

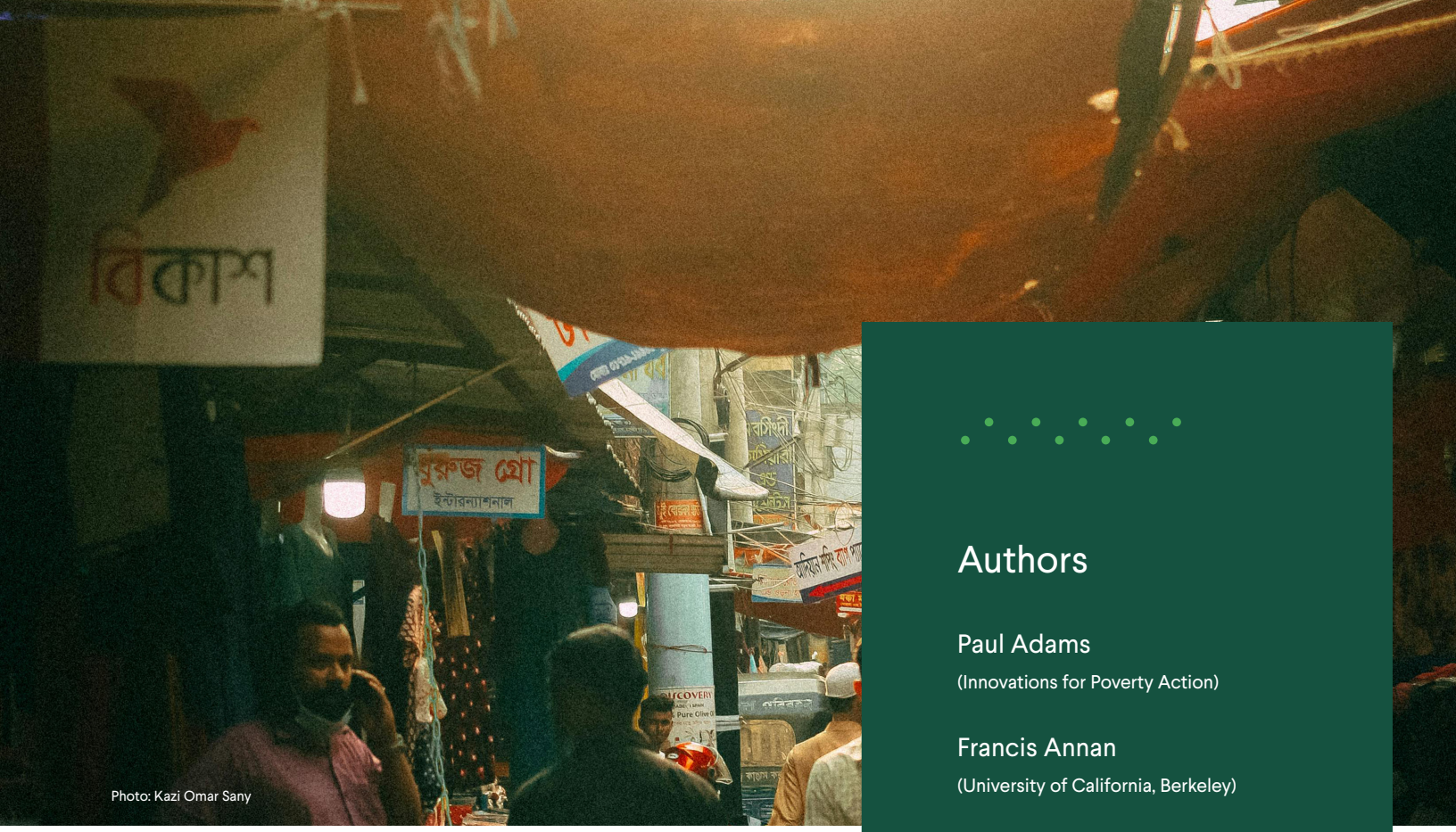


Photo: Kazi Omar Sany

Authors

Paul Adams
(Innovations for Poverty Action)

Francis Annan
(University of California, Berkeley)

William Blackmon
(Innovations for Poverty Action)

Xavier Giné
(World Bank)

Lina Hochhalter
(Innovations for Poverty Action)

Brian Mwesigwa
(Innovations for Poverty Action)

Arianna Zapanta
(Innovations for Poverty Action)

Acknowledgments

This work was carried out by Innovations for Poverty Action (IPA) with support from the Gates Foundation. We gratefully acknowledge excellent research management and support by Furaha Abwe, Akmam Binte Arif, Hashir Arif, Julia Chen, David Rodriguez Estrada, Kyu Khin Gar, Veronica Kakengi, Anthony Kamwesigye, Lauren Perlik, Rachel Pizatella-Haswell, Tyler Spencer, Jerin Tasnim, and Victoria Yan.

Disclaimer

Innovations for Poverty Action bears sole responsibility for the content of this report.

The findings and views expressed do not necessarily reflect the views of the Gates Foundation.

Table of Contents

I. Introduction

Evidence suggests that digital financial services (DFS) can reduce poverty and improve user welfare through a variety of causal channels. There are many factors that influence the usage and uptake of DFS, but one considerable factor is the cost of using these services. It is therefore of interest to regulators and policy makers to understand and monitor the actual costs that consumers face. We developed the Transaction Cost Index (TCI) to measure the costs of using DFS. This toolkit is designed to help regulators, firms, consumer advocacy organizations who wish to recreate the TCI in their markets.

What is the Transaction Cost Index?

The TCI systematically measures the true costs that consumers face related to DFS. "True" costs go beyond official charges. Non-monetary costs, such as opportunity cost of time wasted on failed transactions, and quality of the service provided are both crucial aspects that impact consumers' user experience. This experience impacts the value consumers see in DFS as well as the trust they have in DFS providers and the system as a whole, which ultimately affect the uptake and continued usage of these services. An important objective of the TCI is to conduct this measurement in a cost-effective way, balancing the quality of the insights with the costs of constructing the index. Results from our work testing and refining tools to measure these costs can be found in our Transaction Cost Index reports, along with country briefs and datasets available for download. These resources can be found on IPA's Transaction Cost Index webpage.

In our fieldwork, which put different methods to the test, we focused specifically on mobile money. The goal was to understand the experience of the majority of DFS users, with a particular focus on consumers of lower socioeconomic status, who in our focus countries primarily use mobile money as their primary financial service provider. However, the TCI can be adapted to measure customer experience with other types of frontline financial service providers such as bank tellers at bank branches, bank agents, or merchants accepting digital payments.

What are the consumer costs associated with DFS usage?

Monetary.

Official fees and taxes charged by DFS providers, and unofficial off-the-books fees that may be charged by agents.

Non-pecuniary.

The opportunity cost of time wasted on failed transactions and exposure to consumer protection risks (including fraud and limited price transparency).

Service quality.

Quality of telecommunication channels and network connectivity, liquidity problems experienced by agents, system outages, or agent error or misunderstanding.

Who is this toolkit for?

The TCI is a useful tool for policy makers, regulators, or researchers what want to understand the financial and non-financial costs consumers face when using DFS. It can be used to monitor firm compliance with regulatory frameworks related to pricing and pricing transparency, detect consumer protection risks such as overcharging, or evaluate the success rates of transactions, and understand how front-line DFS agents treat consumers. The TCI's purpose is to shed light on the consumer experience with DFS.

Examples of questions that the TCI can address are:

Agent compliance with mobile money and consumer protection regulations

- Do agents apply illicit fees?
- Do they conduct transactions according to appropriate local regulations?
- Are price lists displayed at the agent location?

Reliability of mobile money transactions

- How often do agent transactions fail, and for what reasons?
- Which transaction types are least likely to succeed?
- How often can a transaction not be completed due to agent absence?

Variations based on customer or transaction characteristics

- How does the typical consumer experience differ by...
  - **Gender, experience level, or age of the customer**
  - **Size or type of the transaction**

Price disclosure and ease of accessing prices

- Are fee lists available from provider websites?
- Does the fee list reflect information on whether taxes are inclusive or exclusive?

How to use this toolkit

This toolkit lays out four different options that can be used to measure DFS costs. The methods and recommendations in this document are based on a two-year project in three countries, but the methodologies can be modified and applied at the desired scale (whether national, regional, community, or firm specific level) and timeframe (as a one-off exercise or repeated data collection, for example to assess the impacts of policy changes). Ultimately, the choice of method will depend on the goal of the exercise, the variables of interest, and the resources available.

Official fees charged by DFS providers can be measured through a review of price lists that providers make available on their websites, which also allows for measurement of some pricing transparency indicators. However sometimes these listed prices are not reflective of the official price in the market. To capture the current official price with certainty, and costs beyond official fees – including informal fees and non-pecuniary costs incurred when using agents – fieldwork is necessary. Options involving fieldwork include interviews with mobile money customers who are intercepted outside agent locations or mystery shopping visits with agents. Mystery shopping can be done either by professional enumerators or by locally engaged and trained consumers.

Each methodology comes with its own advantages and disadvantages in terms of what it can measure, with what accuracy, and at what cost. The toolkit provides a menu of options to choose from based on these considerations, together with recommendations and tools for their practical application

The remainder of this toolkit is structured to give a short overview of each method: (1) tracking of official fees from provider websites; (2) interviews with consumers; (3) mystery shopping by professional enumerators; and (4) mystery shopping by local consumers.

For each method we also discuss their individual challenges and advantages which can be used to determine the method best suited to your particular needs. The last section then provides a detailed guide on how to implement the methods in practice, including questionnaires and protocols for each method.

Method types

TRACKING OF OFFICIAL FEES FROM PROVIDER WEBSITES

INTERVIEWS WITH CONSUMERS

MYSTERY SHOPPING BY PROFESSIONAL ENUMERATORS

MYSTERY SHOPPING BY LOCAL CONSUMERS

II. Data collection approaches: a menu of options

In this section we describe four different ways to measure consumer costs when using mobile money agents. We begin with a pure desk exercise, tracking pricing information from mobile money providers' websites. Next we turn to fieldwork options, including interviewing consumers and mystery shopping, either by trained professionals or local consumers.

Method 1:

Tracking of official prices

This method is designed to gather official, listed prices from major mobile money providers' websites. Prices include both fees charged directly by providers for transactions and government taxes (where applicable and publicly listed). The goal of this approach is to document the full official monetary cost consumers incur when making transactions, excluding any extra fees levied by agents. To lower collection costs, the process of collecting price lists from providers' websites and monitoring for changes in prices can be automated. Raw data for each provider can be collected using three methods: a) direct web scraping for fees listed in HTML tables, b) web scraping and PDF-to-HTML conversion for fees listed in PDF tables, and c) manual compilation of data in excel for fees not listed in tables (e.g. in FAQs, customer care inquiries and images).

In addition to collecting data at a single point in time, it is also possible to use website monitoring tools, such as Visual Ping and Distil to alert users to specific website changes. Finally, historical prices can also be found using the Wayback Machine, a digital archive of the World Wide Web. This creates a few options – a true "real time" monitoring system can rely on website monitoring tools to detect changes which could alert staff to re-run the web scraping process and update the pricing dataset for that provider (or even potentially automatically trigger this scraping and dataset updating). Alternatively, the price dataset can be updated periodically (quarterly or annually, for example). If intermediate changes in prices are of interest, results from website monitoring or archiving tools can be used to fill in this information, though this can involve a manual process if automated scraping tools are not adjusted to work with outputs from the website monitoring or archiving tool.

Challenges and opportunities

i) Costs

Scraping of official prices is a quick and inexpensive method of price monitoring, as no fieldwork is necessary. When automating the process is deemed appropriate, there is a level of technical knowledge required to build and maintain the backend code, which may or may not pose significant labor costs. A basic proficiency of R Studio, an integrated development environment for R, along with adequate knowledge of the mobile money market context, is sufficient to carry out the automated exercise. The sophistication of the programming may be determined by the existing formats of pricing information. If all pricing sites are in html format and there is bandwidth to upgrade, it may be worthwhile creating a program that runs the script at the backend in real-time so it produces the reports at your desired frequency. Otherwise, there will still be some level of manual check and input. In our case, it required roughly four months of full-time staff time to build the backend code, set up the pages in the web monitoring tools and collect data for two quarters.

There are typically membership fees associated with web scrapers and visualization tools, such as Visual Ping and Distill. Annual membership fees for business use range from $25-$250 per month, depending on the number of pages and users, among other things. For our project, we used a free subscription in Distill to track 25 provider pages and a paid basic annual membership ($120) in Visual Ping for the remaining providers, including those whose pricing information are in image formats. We used this combination because it was more cost-effective at the time of set-up - Distill's free subscription has a page limit and can only track images in the pro version which was more expensive than getting a supplementary paid basic membership in Visual Ping that can track images. Between the two tools, we found Distill tracking to be more accurate than Visual Ping - the former shows more clearly the changes versus the "fuzziness" that sometimes show up in the latter. Similar to the programming effort level, the choice of the web scraper would depend on the formats of the fee pages to be monitored and budget.

1. At the minimum, a completely manual process of data collection can be employed where someone copies the official prices from provider websites. This may be a cost-effective approach for an exercise covering a single country with 2-3 providers, for example, though it is more labor-intensive and poses increased risk of human error.

ii) Reflection of real-world consumer experiences

Among all methods aimed at capturing the full transaction costs, the online scraping of official prices is the least reflective of real-world experiences consumers have when making mobile money transactions. It relies on the display of online prices which may not be up-to-date and does not capture informal, extra fees charged by mobile money agents or non-monetary costs particularly related to time wasted on failed attempts to make transactions.

iii) Data quality

If prices are displayed in an appropriate format then the automated nature of this method increases the data quality, as it is less prone to human error. Calculations of pricing can also be done with ease, if adjustments need to be made to cover additional components like taxes. However, the web scraping code is also prone to some errors, and so we recommend that the automatic process is periodically monitored to ensure changes are captured correctly. Regulators can play a role in ensuring data quality by requiring providers to make their prices available online in a standardized, machine-readable format.

Outcomes that can be measured

Relative to in-person surveys, web scraping captures a more limited set of data. This sort of web scraping only captures official listed pricing information of different transaction types and services. It rests on the assumption that online pricing pages are updated in real-time and reflect actual costs on the ground, which may not be true. Even if prices are not displayed, this approach gives important information about the underlying quality of disclosure in financial markets.

From a market monitoring standpoint, similar to the concept of "sludge audits" in behavioral public policy, the web scraping approach shines a light on specific consumer protection issues, which include:

Price transparency: Are fee lists available from provider websites?

Ease of accessing information: Does the fee list reflect information on whether taxes are inclusive or exclusive?

These unjustified frictions, or sludges, might end up depriving consumers access to financial services, especially the most vulnerable members of society.

Method 2:

Consumer interviews

Consumer experiences with DFS can be captured by interviewing ordinary mobile money customers. To conduct the surveys, enumerators station themselves outside agent locations and complete interviews with exiting customers. By stationing enumerators at agent locations during normal business hours and interviewing any person leaving the location, with the only requirement that they have just completed a mobile money transaction with the agent, a relatively representative sample of consumers can be generated. Enumerators ask customers questions about their recent transaction, including the success of the transaction, fees incurred, and quality of service received.

Challenges and opportunities

i) Costs

Intercept surveys can be costly, as recruiting customers takes time. Especially in rural areas, agents receive few customers per day. Many are not dedicated agents and also operate a shop or other business. Therefore, waiting for customers that made a mobile money transaction (as opposed to a shop purchase) can be time consuming. In more urban areas where customer traffic is high, consumers often do not have time to answer questions, or are reluctant to share information about their financial dealings with a stranger. Low survey productivity results in high per-survey costs. Based on our experience, limiting the wait time to between 40 minutes and an hour per respondent proved to be just enough to accurately assess the traffic flow at the agent location. This duration strikes a balance: it's not too short, which could lead to incorrect conclusions about low traffic, and it's not too long, which could significantly reduce productivity and unnecessarily increase costs. Of all the TCI methods piloted, intercept surveys were the most expensive.

Outcomes that can be measured

In-person intercept surveys can provide direct insights into the customers' interactions with agents, the types of transaction they attempt, and how much they are charged. In addition, they capture customer profiles – such as age, gender, educational level and DFS experience.

Outcomes that can be captured in intercept surveys include:

Transaction outcomes: Whether the transaction was successful, what fees were paid including extra informal fees, how transparently fees were disclosed, and the quality of service the customer received from the agent.

Characteristics of typical transactions in the market: Common transaction types and values.

Characteristics of DFS consumers (used for segmentation): Gender, age, education, user sophistication, and other variables related to DFS usage, such as the types of challenges the consumer has experienced while using DFS.

ii) Adaptability

Consumer intercept surveys can only reflect the actual transactions that consumers make. Variables such as transaction types, values, DFS provider, or customer characteristics cannot be artificially controlled or manipulated. One can set targets to generate a sample of consumers balanced by, for example, gender or educational achievement, though this can reduce productivity in contexts where few mobile money customers are available.

iii) Reflection of real-world consumer experiences

Of all the methods in this toolkit, consumer surveys most accurately reflect the DFS transactions of regular consumers. This includes not only the characteristics of typical transactions, but also incorporates behavior and decisions that are based on local knowledge of the agents they transact with. For example, consumers may know which agent is most likely to succeed with a given transaction scenario, or agents' customary opening hours. Mystery shopping methods, which pre-assign mystery shopping visits rather than relying on real-world transactions, don't capture this local knowledge, so may overestimate actual average failure rates experienced in the market. We discuss mystery shopping methods in a subsequent section.

iv) Data quality

A central component of the TCI is the monetary cost of DFS transactions, which includes overcharging by agents. Fees and charges come in various forms – official provider fees are typically automatically deducted from mobile money accounts, and an agent may ask for an extra unofficial fee to be paid in cash, or may deduct the extra fee directly from a customer's account. A key challenge is that ordinary consumers are more likely to suffer from recall bias, making it difficult to capture costs incurred. Additionally, they often lack the knowledge necessary to accurately disaggregate the fee components, creating measurement errors and biases in their perceptions of overcharging. Nonetheless, insights into perceived overcharging (i.e. what consumers estimate that they were overcharged), even if incorrect, is valuable in and of itself, as concerns about pricing could affect customer trust and potentially influence their DFS uptake and usage.

In theory, the recall and knowledge issues mentioned above can be reduced if enumerators review consumers' transaction histories on their devices – in practice, this could be considered invasive, and would likely make potential respondents less likely to participate in the survey.

As with any survey, and with methods three and four discussed below, standard data quality measures like CAPI implementation, consistency checks, and numeric range validation should be applied. For detailed guidance on these, see the implementation guide in the next section, or IPA's "Minimum Must Dos".

v) Observer effects

Observer effects (also known as Hawthorne effects) can impact the behavior of agents in a way that skews findings away from the experience of typical consumers. Agents that become aware of being monitored may reduce their misconduct – including overcharging, but also other aspects of their service, such as keeping customer information private. This is particularly the case if agents perceive the work to be associated with their mobile money service provider or a government agency, both of which could take enforcement action based on any observed misconduct. Enumerators waiting outside agent locations for prolonged periods of time, and that are seen interviewing exiting customers, are easily noticed. Agents frequently share information about unusual or suspicious activity with other agents in the area, which might prime agents to behave in a certain way. As such, to avoid detection, enumerators should dress to blend in with local populations; travel separately; and disembark their vehicle out of sight of agents.

Method 3:

Mystery shopping by professionals

Mystery shopping is often considered the gold standard for collecting accurate data on frontline service provider conduct, including overcharging and other misconduct by DFS agents. During a mystery shopping visit, an enumerator conducts a set of predetermined transactions with an agent, simulating a regular customer interaction, and records their experience in a survey.

Challenges and opportunities

i) Costs

Relative to other fieldwork approaches such as consumer intercept surveys, professional mystery shopping can be a relatively inexpensive way to collect data, mainly because survey productivity is high, which results in lower per-survey cost. Effective sample size (visits where the agent was open and a survey could be filled out) can be increased through revisit protocols of closed agents.

ii) Adaptability

Mystery shopping methods allow researchers to directly control the types of scenarios investigated. This includes the type and size of the transaction to be conducted, the mobile money provider used, and the observed characteristics of the mystery shopper, such as gender. By controlling these variables, you can ensure that differences in outcomes are not driven by these factors by holding them constant, or vary them in a controlled manner to understand how they impact agent behavior. For example, randomly varying shopper gender makes

Outcomes that can be measured

Outcomes can be captured through mystery shopping include:

Transaction outcomes: Whether the transaction was successful, what fees were paid including extra informal fees, how transparently fees were disclosed, and the quality of service the customer received from the agent.

Impacts of transaction type, transaction value, DFS provider, and shopper characteristics on visit outcomes

it possible to explore potential gender discrimination. Similarly, including two different transaction values allows you to explore differences in overcharging (and other outcomes) by transaction value. This could be important particularly if certain segments (e.g., women or rural populations) typically conduct different valued transactions on average. Our own work suggests that outcomes, such as success rates and overcharging differ significantly by transaction type, so being able to design scenarios that explore these differences is valuable.

iii) Reflection of real-world consumer experiences

While mystery shopping offers researchers the ability to adapt scenarios to their particular areas of investigation, this comes at a cost: by directly controlling the types of scenarios tested, mystery shopping does not reflect consumers' knowledge as well as consumer intercept surveys. Data collected by professional enumerators conducting mystery shopping visits is likely to overestimate actual transaction failure rates, since enumerators do not have knowledge of agents' ability (and willingness) to complete transactions – something that local consumers will know and act on. Additionally, mystery shopping often struggles to capture real-world consumer decisions and to accurately reflect representative consumer demographics, typical transaction types, and transaction amounts. However, researchers can endeavor to bridge this gap by designing scenarios that closely mirror real-life situations, as documented in representative surveys.

iv) Data quality

Professional enumerators are better at recording responses accurately and tend to report mobile money fees that are consistent with official fees. To further enhance data quality, enumerators can take screenshots of their pre-and post balances as well as automated generated post transaction receipts. These records serve two verification purposes: confirming official fee calculations and proving enumerators visited their assigned agents. The automated receipts contain agent identifiers (such as phone numbers or provider-assigned IDs), and comparing these identifiers across different enumerator visits can reveal potential errors. However, this verification method isn't perfect, as agents may use multiple phone lines throughout the day, and receipts are only generated when transactions are successful. For suggestions on how this can be done, see Online Appendix 2: Transaction Verification Resources.

v) Observer effects

While all methods discussed in this toolkit are susceptible to observer effects, professional mystery shopping is particularly prone to this issue due to the conspicuousness of the enumerators. Mobile money agents tend to be on the lookout for suspicious behavior as scams and robberies are relatively common occurrences at mobile money agent locations. They may also be on alert for mystery shopping visits which can be carried out by FSPs or regulators for enforcement purposes. Enumerators are especially likely to be noticed in rural areas with few mobile money agents and narrower customer bases, where outsiders are easily identified. Even in busy markets where enumerators are less likely to stand out as strangers, unusual visit patterns – such as repeated transaction requests involving the same transaction amounts – can still alert agents. In general it is difficult to estimate the impact of observer effects on research outcomes.

Method 4:

Mystery shopping by local consumers

The final method involves mystery shopping by regular consumers who live in the study areas, rather than by enumerators. The local mystery shoppers can be recruited through intercept surveys or household surveys. Once recruited, they receive training, and conduct visits similar to the professional mystery shopping visits described above. Professional enumerators accompany the local shoppers to each visit, wait outside the agent location until the visit is complete, and fill out the survey together with them.

Challenges and opportunities

i) Cost

Local mystery shopping is more costly than mystery shopping done by professionals. While standard compensation rates for consumers tend to be lower than enumerator salaries, professional enumerators still need to be employed to assist the shoppers. On top of this comes the added monetary and time cost of recruiting the participants. Nonetheless, survey productivity is significantly higher than for consumer intercept surveys (each enumerator can handle several local consumers, and several consumers are active in the market at the same time), which drives down per-survey costs. Ultimately, the cost per survey is likely to lie in-between that of professional mystery shopping and consumer intercept surveys.

Outcomes that can be measured

Outcomes can be captured through mystery shopping by local consumers include:

Transaction outcomes: whether the transaction was successful, what fees were paid including extra informal fees, how transparently fees were disclosed, and the quality of service the customer received from the agent.

Impacts of transaction type, transaction value, DFS provider, and shopper characteristics on visit outcomes

Familiarity with the agent: Do outcomes, such as overcharging and agent conduct, differ between regular and new (or less frequent) customers?

Local knowledge: What are the agent presence rates when mystery shoppers have knowledge of agents?

Representative shopper demographics: What are the effects of different customer characteristics on outcomes (including education, age, level of user sophistication)?

ii) Adaptability

As with all mystery shopping, it is to a certain extent possible to adapt the scenarios that are being tested (such as transaction types and values) according to need. However, regular customers often have other occupations, limited time to spend on the activity, and may find the research itself more challenging. So researchers may wish to limit the number of visits or the number of transaction types.

iii) Reflection of real-world consumer experiences

Mystery shopping is a stylized version of a regular customer visit and does not reflect real-word conditions. However, some of the effects of shoppers' local knowledge can be harnessed by adjusting incentive structures and scenario designs. For example, you can vary monetary compensation per visit according to whether agents are present or not, encouraging shoppers to visit agents when they believe they are most likely to be open. Local mystery shopping can thereby be designed to better reflect the typical consumer experience, compared to professional mystery shopping.

iv) Data quality

Opposed to consumer interview respondents, local mystery shoppers receive training prior to their visits. They are therefore better at identifying transaction fees and accurately answering questions relating to other TCI core outcomes, such as the presence of price lists. Nonetheless, the research work may still be challenging to them, particularly when filling out a survey programmed on unfamiliar data collection platforms. Navigating new software can be daunting, as it requires understanding specific user interfaces and input methods that may not be intuitive, affecting the efficiency and accuracy of the data collection. Professional enumerators therefore play an integral part in local mystery shopping, by helping the shoppers to find their assigned agents, assisting them with questions, and filling out the survey on the consumer's behalf. With these measures in place, local mystery shopping can achieve levels of data quality similar to that of professional mystery shopping.

v) Observer effects

Local customers may not stand out as much during mystery shopping visits in the area. However, the presence of professional enumerators in the field to offer them support means that concerns about observer effects persist. Requiring enumerators to dress appropriately and remain out of sight of the agent (both when waiting for the shopper, and when filling out the survey) can help with mitigating observer effects. An additional concern is that local shoppers who are well acquainted with (and loyal to) agents might reveal to them that they are taking part in a study. Emphasis should be placed on mystery shopping as a research method during training.

Photo: Dennis Sylvester Hurd

III. A guide for implementation

In this section we review the same four methods described in the prior section, for each providing practical information on how to go about carrying out each method. We encourage readers to use Section II to inform their decision about which method is most appropriate for their needs and this section to be used as a guide for implementing the chosen method. We conclude with a discussion of data quality assurance, data management, and analysis, with guidance that applies regardless of the choice of method(s) used.

Method 1:

Tracking of official prices

Systematically collecting official mobile money fees from major mobile money providers' websites can prove useful in tracking pricing movements especially after regulatory changes. Throughout the two-year duration of this project, we documented our experience carrying out this price tracking exercise with the goal of contributing to effective mobile money price monitoring in LMICs where mobile money use is prevalent. Promoting a harmonized system for collecting this pricing information across these markets, as characterized by consistent terminology and clear price transparency guidelines, allows consumers and regulators alike to conduct easy fee comparisons to aid decision making. This is most significant in areas where price transparency issues still exist.

The methodology followed can be divided into five main components: a) planning, b) monitoring, c) scraping, d) cleaning, and e) visualization. We discuss this in detail in the next section. Prior to this, delineating at the start the prerequisite information (transaction types, mobile money providers covered) and optional data (reference transaction value, market share), if analysis is desired, will make the process more efficient. As you read the discussion, you may treat our process as a baseline scenario in which you can build upon to customize according to your objectives and context.

Step one: planning

Determining the objective for the price monitoring and understanding the local mobile money context are the first steps in planning a successful exercise. While we limited our data collection to basic information given the scope of our study, the components can be extended to fit the country context. We discuss below our selections and note how it can be modified.

LINKS TO RESOURCES:

Tracking of official prices

The following resources are available for those interested in tracking official prices from provider websites in the Online Appendix 1

- Converting online pricing data to excel format
- Appending provider/country datasets to one master dataset
- Creating codebook

Prerequisite information

i) Transaction types: Mobile money providers often include a wide range of services that can be paid for through mobile money, such as bill payment, merchant payment, airtime purchase, e-commerce, etc. In our study, we limited our collection of fee information to four core mobile money transaction types: cash-in at an agent, cash-out at an agent, on-network person-to-person transfer, and off-network person-to-person transfer.

Transaction types

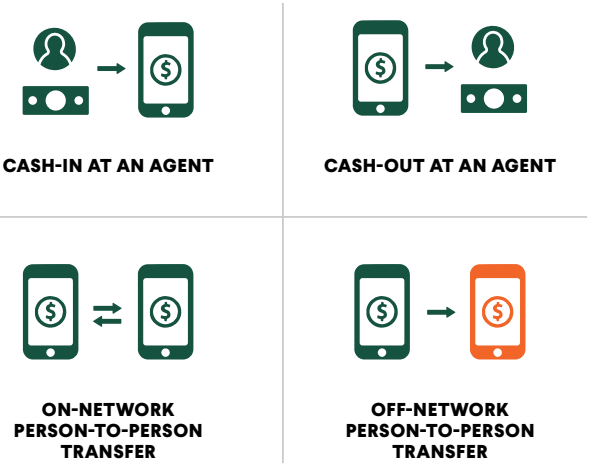


TABLE 1: Core transaction types studied

TRANSACTION TYPES	DEFINITION
Cash-in at an agent	Depositing cash into a mobile money wallet with an agent
Cash-out at an agent	Withdrawing cash from a mobile money wallet with an agent
On-network person-to-person transfer	Transferring money from one mobile money wallet to another wallet with the same provider (self-serve, no agent needed)
Off-network person-to-person transfer	Transferring money from one mobile money wallet to another wallet with a different provider (self-serve, no agent needed)

Given the heterogeneity within these core transaction types and across providers, we found it helpful to be clear with the definition of the transactions we intended to capture. For example, we have seen that cash-in and cash-out can be conducted in multiple ways, such as through an ATM, an app, convenience store, pawn shops or other payment facilities - each of which may be charging a different fee. While it may be excessive in a project scope covering 16 LMICs like ours, these additional types can be covered in a country-level monitoring especially if it is a common service among mobile money providers.

ii) Mobile money providers: We gathered official transaction prices from major mobile money providers, defined as those that make up a combined market share of at least 80 percent in each country, which include between one or three providers.

There may be interest to include not just mobile money market players but also other digital financial services providers (FSPs), such as traditional banks and other Fintechs that may have overlapping services of interest. It may be helpful to create a listing of all provider information, such as market share, number of users and company type, to understand the universe of players.

Optional information (for analyses)

i) Reference transaction value: To compare fairly across countries, we computed the fees as a proportion of a reference value. The choice of transaction value is key because many providers' pricing structures are quite regressive, meaning that the cost in percentage terms varies significantly depending on the size of the transaction value. The simplest approach would be to set a single USD value, as is done in similar price measurement work, such as the World Bank's Remittance Prices Worldwide. However, our intent is to replicate the experience of the typical mobile money user in each country as much as possible. Because of differences in the economic development of each country, typical mobile money transaction sizes differ significantly across countries so having a single reference value in USD amounts is not ideal.

Because data on the distribution of mobile money transaction sizes is not publicly accessible in most countries, we had to find a different approach to setting reference values. Using self-reported transaction data from IPA's consumer protection surveys, we found that median transaction sizes tended to equate to approximately 15 times the mean daily income per capita for the bottom 40 percent of the population. Although this is a very rough approximation, we believe it reflects the typical transaction size more accurately than a set USD dollar amount used globally, and we tried to recreate this for each country. We implemented this approximation in the following way.

We used World Bank data on the daily mean income per capita for the bottom 40 percent of the population to approximate the median transaction size. We converted this value to local currency in 2017 using World Bank's PPP conversion factor for 2017, then inflated it to current local currency using local CPI levels. Lastly, we multiplied this income per capita in current local currency by 15 to obtain our reference value.

Using a reference value simplifies comparisons across countries, and across providers within a given country, by condensing complex pricing structures into a single fee expressed as a percentage of the reference value. By using reference values that approximate typical transaction amounts in each country, the true costs consumers face are more accurately reflected. This approach is more representative than alternatives like setting a single global reference value (e.g., USD $20) or calculating the average cost across the entire price schedule.

2. Remittances Prices Worldwide uses two reference values: USD $200 and USD $500, converted to local currency. See https://remittanceprices.worldbank.org/methodology.

ii) Market share: We looked at estimates of market share from both official government sources and third-party market reports. If these were not available, we used qualitative information from market reports, selecting the leading two to three providers identified as the major providers in the country. In our study, we explored the relationship between mobile money prices and the level of competition across countries included in our review. We plotted the aggregate cost by country in percentage terms against each country's Herfindahl-Hirschman Index (HHI), a standard measure of market concentration where larger values indicate a more concentrated market.

iii) Mobile money regulations: In our study, we also tracked relevant regulations related to mobile money pricing to contextualize trends in providers' prices in specific markets. We assessed a set of basic indicators that can be compared across the 16 countries studied. Namely, the themes covered are interoperability, pricing caps, pricing transparency, redress, and taxation.

Identifying relevant policy changes and when they happened can be useful in explaining price movements. Combining it with other administrative datasets such as transaction volume or subscriber base, if available, could also deepen understanding of potential policy impacts.

Step two: monitoring

Once the information to be collected is identified, online webpage monitoring tools, otherwise known as change detectors, can be set-up to monitor the provider websites of interest. Each tool has varying features and limitations depending on the tool and which subscription you choose. In our project, we signed up to a free subscription in Distill to track 25 provider pages and a paid basic annual membership ($120) in Visual Ping Ping for the remaining providers, including those whose pricing information are in image formats. We used this combination because it was more cost-effective at the time of set-up - Distill's free subscription has a page limit and can only track images in the pro version which was more expensive than getting a supplementary paid basic membership in Visual Ping that can track images. Signing up is quick – it only requires an email address and credit card information, if you are getting a paid subscription.

Once subscribed, you can add websites that you want to be monitored for changes. For Visual Ping, you can make adjustments on the job name, frequency, threshold of change, and section format to be monitored. In our project, we set the frequency of checks to every week, indicated that any change be detected (versus only medium or large changes) and chose the element type – visual, text or element-depending on the provider page. For Distill, we could only adjust the intervals of the checks, which we set every 14 days. You can access the dashboard and view the summary of the checks anytime.

Both tools can trigger an instant notification via email or SMS when there is a detected change during a scheduled check – the availability of the alert platform (SMS notifications are possibly in paid subscriptions only) and maximum number of notifications (only 30 email alerts per month are offered in Distill's free subscription) would depend on the subscription type. In our case, we found the email notifications, including the limitation on email alerts from Distill, to be sufficient as we check emails every work day and prices do not move as frequently to necessitate unlimited alerts. Typically, each email notification would show previous and current snapshots of the page with the detected changes highlighted.

Between the two tools, we found Distill tracking to be more accurate than Visual Ping – the former shows more clearly the changes versus the "fuzziness" that sometimes show up in the latter. That said, these online monitoring tools do not support a few formats we have seen in this project where manual monitoring needs to be done. Examples of formats that need manual monitoring are pricing information contained inside FAQs, pricing provided through customer care response, and those obtained through fee calculators.

A somewhat manual and less stringent approach to monitoring prices is making use of a web page archiving service such as the Wayback Machine. Wayback Machine is designed to access historical information and can thus allow researchers to recover price changes not recorded in real time. It automatically snapshots web pages at various points in time, which are then stored, attached to timestamps and made accessible to users. However, this approach has three main drawbacks. First, because Wayback Machine relies on users to manually suggest webpages to archive, not all webpages with mobile money fees are archives. Second, for the same reason, web pages with mobile money fees are not typically archived on a set schedule and long gaps can exist between archive dates, meaning some price changes may be missed. Finally, the Wayback Machine struggles to archive dynamic web pages, such as those that require the user to select items from a dropdown list before viewing prices. Paid services, including Archive It, address some but not all of these drawbacks.

Step three: scraping

Table 2 shows the scraping methods that can be used to capture various machine-readable and non-machine-readable pricing formats. Given our scope of 16 LMICs, we encountered a wide range of formats (pdf, image, html, calculator, embedded in FAQs, through customer service) – each format necessitating a unique scraping approach.

TABLE 2: Scraping method by format

USE OF MACHINE-READABLE FORMAT	FORMAT	METHOD
Machine-readable format	HTML	Data can be directly scraped into a dataframe using a programming software like R.
	PDFs	Automate getting data from PDFs through an API (Application Programming Interface), which can be called from any programming language. To get the API, you create an account with pdftables, wherein you get 50 free pages of pdf conversion with each account.
		Another option is to use pdf-to-excel converters online. Excel format can then be imported into the programming software.
	Image (JPG and PNG)	Many image-to-excel converters, such as Smallpdf and iLovePDF, are available online.
Non-machine readable format	Information is under various FAQs	A manual check for fee changes was conducted by comparing current fees to last collected fees. Fees are recorded in excel format.
	Information is provided through customer care response	
	Information is obtained through fee calculators	

Machine-readable formats such as HTML and PDFs allow for the quickest and most accurate scraping. Nonetheless, the other methods are useful in specific scenarios.

Our exercise reveals that maintaining multiple methods to cover different pricing formats could be cumbersome. From a price monitoring standpoint, having clear, harmonized guidelines on price transparency, particularly from online websites, could simplify the process. The exercise also rests on the assumption that providers update their pricing pages regularly. Mobile money providers vary in the frequency of the update from real-time to every year, often with no indication of the last date of the update. We cannot be fully certain if the official prices posted reflect the actual costs, despite regulatory policies requiring providers to display prices at relevant locations – this is the major limitation of the price scraping method. Mystery shopping methods could complement price tracking by also monitoring posted prices at physical agent locations, and actual prices paid when making transactions.

Step four: cleaning

Data cleaning and processing would benefit from effective data management. The initial work involves setting up the folders and software code – in our research, we used R Studio. Regardless of the method of scraping, data from every provider is converted into the standard dataset format in R Studio. For some countries, all providers' prices need to be scraped manually, and so a single excel is manually created by country and by provider. All provider/country datasets are then appended into one master dataset using an R script.

When there are changes detected through notifications from Distill or Visual Ping (as part of the monitoring stage), we manually trigger the re-run of the R script that cleans each provider's pricing data to produce the revised provider dataset – it is not a real-time change. If the scraping was done manually, the Excel sheet needs to be manually revised. Since our exercise covered multiple countries, the price conversion to USD was integrated in the R script. All price conversion is based on Oanda, a reliable online foreign exchange source, and documentation of the exchange rate is also saved. Table 3 shows a resulting summary of the quarterly dataset.

TABLE 3: Sample of the summary quarterly dataset

PROVIDER: Bkash									
TRANSACTION TYPE	CHANNEL	CUSTOMER TYPE	CURRENCY	VALUE MIN	VALUE MAX	FEE	FEE PCT	TAX	TAX PCT
cash-in	agent	Registered	Taka	50	30000	0	–	0	–
cash-out	agent	Registered	Taka	50	25000	–	1.61	–	0.24
off-network p2p transfer	self-service	Registered	Taka	0	200000	–	0.43	–	0.07
on-network p2p transfer	self-service	Registered	Taka	.01	100	0	–	0	–
on-network p2p transfer	self-service	Registered	Taka	100.01	25000	4.35	–	0.65	–
on-network p2p transfer	self-service	Registered	Taka	25000.01	200000	8.7	–	1.3	–

TABLE 4: Codebook of variables

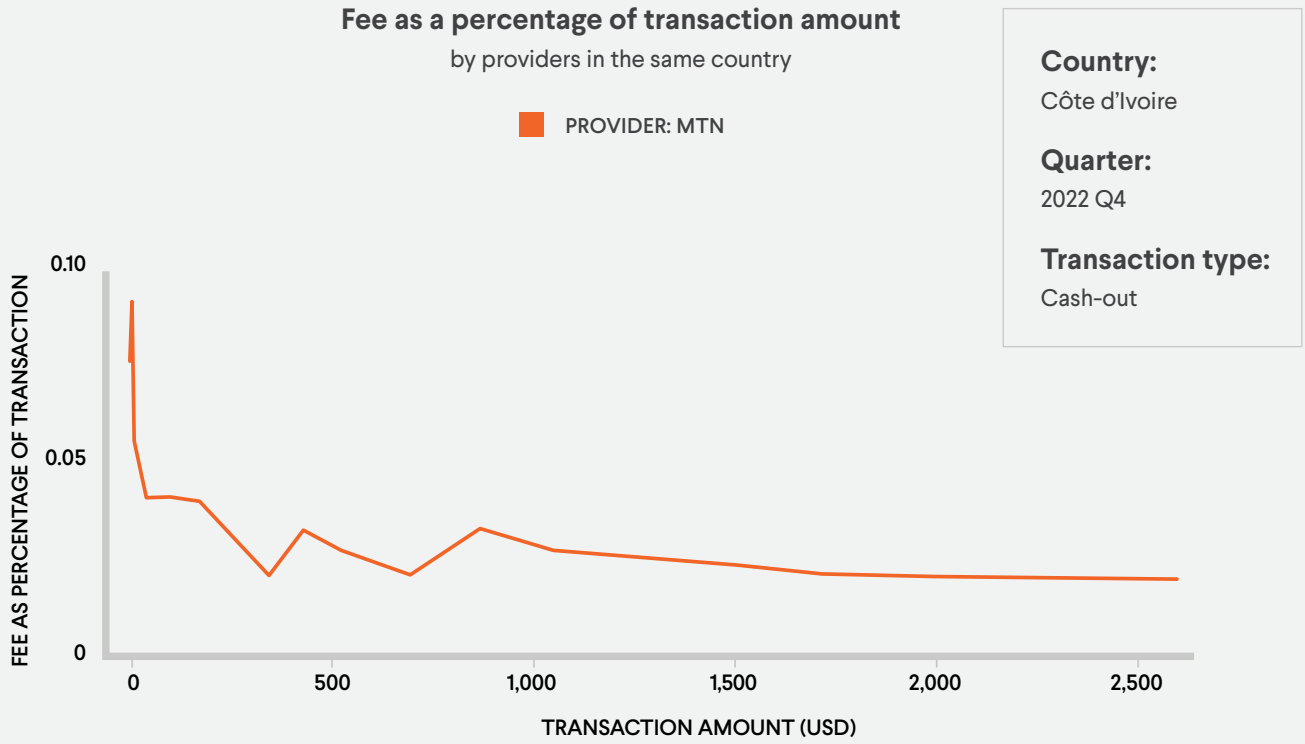
VARIABLE NAME	VARIABLE LABEL
country	Country
mobile_money	Name of mobile money service
provider	Service provider
transaction_type	Type of transaction
channel	Channel to initiate the transaction
customer_type	Whether the fee is applicable to a registered or unregistered user
value_min	Minimum transaction amount
value_max	Maximum transaction amount
fee	"Slab based" fee charged by provider
fee_pct	Percentage based fee charged by provider
tax	Government tax charge, if tax is "slab" rather than percentage
tax_pct	Government tax charge, if tax is percentage rather than "slab"
currency	Currency of pricing information on provider's website
exchange_rate	Exchange rate to 1 USD on January 1, 2022 (Source: Oanda.com)
value_min_USD	Minimum transaction amount in USD
value_max_USD	Maximum transaction amount in USD
fee_USD	Slab based fee charged by provider in USD
tax_USD	Government tax charge in USD, if tax is "slab" rather than percentage
fee_pct_USD	Percentage based fee charged by provider in USD
tax_pct_USD	Government tax charge in USD, if tax is percentage rather than "slab"
date_collection	Date of access for pricing information
web_address	Webpage URL of the pricing information

The approach we employed in this research required basic proficiency in R and R Studio. This technical skill set enables the maintenance and revision of the R script to address needs. That being said, we fully recognize that there are other ways and softwares that can be used, some of which may even process data faster. One option may be using programming languages to create a program that runs the script 24/7 and produces reports at a desired frequency. The set-up cost may likely be more expensive than using our basic code in R, but it may be less labor intensive in the long run.

Step five: visualization

We employed Microsoft applications, such as Excel and Power BI, to create graphs and visualizations on pricing trends and country comparisons. Use of Power Bi service requires a license and a capacity. If accessible, Power BI is a powerful tool to turn data into visuals with advanced data analysis tools and a user-friendly report creation tool. For our project, we used Power BI to create an interactive visualization tool using Power BI to examine the data in real-time. We imported our Excel workbook into our Power BI workspace, and built separate line graphs for country-level and provider-level data - the Power BI provides a way to link graphs such that when you select a country, corresponding providers in that country will also be reflected. It also has a helpful "tooltip" feature that displays information when you hover over visuals. Power BI is designed to be user-friendly, and learning data analysis in the application may be quicker and simpler than learning data analysis with R, though it can be challenging to those with no prior experience with data analysis tools. To those who do not have access to Power BI, Excel offers many advanced charts and graphs to help present your data more effectively.

FIGURE 1: Sample visualization output using Power BI



Note on taxation

In our exercise, we wanted to measure the full monetary costs consumers incur when making transactions, exclusive of any extra fees levied by agents. This meant that tax fees, if available, would be included. As it turns out, the level of tax disclosure varies significantly by country. Some countries do not impose any type of tax on mobile money transactions; of those that do tax mobile money, providers will do one of the following: (a) exclude taxes from their price lists, leaving it up to consumers to determine how much extra they will be charged in taxes, (b) include taxes in their listed prices, but without separating the taxes and the direct provider fees, or (c) include taxes in their listed prices, separating out the taxes and the direct provider fees. If taxes are not explicitly included, we conducted a separate review to determine the applicable tax rate.

There are two broad types of consumer taxes that can apply to mobile money transactions, depending on the country. First, broad-based taxes including sales tax and value-added tax (VAT) are applied to most goods and services in an economy, including mobile money transactions. These broad-based taxes are typically applied to the fees FSPs charge consumers for completing mobile money transactions. Second, targeted taxes apply to only a segment of transactions in the economy. Targeted taxes that are applied to mobile money transactions in the countries we reviewed include excise taxes, stamp duties, and mobile-money specific levies. The tax base for these targeted taxes can either be the transaction fee or the value of the transaction itself. The TCI Year 1 Comparative Report discusses this in detail. To accurately capture the prices, it is key to understand the specifics of taxation in the country of interest in order to adjust the calculations accordingly. If desired and relevant, the extra fees levied by agents can also be included. As seen in Table 3. Codebook of Variables in the section on Cleaning, the components of the price can be distinguished to see its breakdown.

Photo: Esther Ruth Mbabazi

Fieldwork methods

Before delving into each of the three fieldwork methods (consumer interviews, mystery shopping by professionals, and mystery shopping by local consumers), we cover two topics that apply across the board: market and agent selection.

Market selection

DFS agents tend to cluster in limited geographic zones, for example in markets selling food and other goods or in transportation hubs, which offers convenience for consumers. These markets can vary greatly in size – from large markets in urban centers that host many agents to small rural markets or village centers with only one or two agents. As a first step to defining the sample of agents to include in your fieldwork, we recommend selecting a set of these markets for inclusion in your study. This section will describe that process, starting with a discussion of sample size, then moving to sampling approaches, and finally discussing the mechanics of selecting and defining individual markets.

i) Sample size determination

Sample size determinative involves weighing tradeoffs between precision with which key outcomes are measured and budget. Larger sample sizes allow for more precise estimation of consumer outcomes but also require larger budgetary commitments. Power calculations can be used to determine the sample size required to measure outcomes with a certain level of precision. For example, if you would like to estimate rates of overcharging with a certain confidence interval, or you would like to be able to measure a certain difference in reliability between urban and rural agents with statistical significance, power calculations can tell you the number of visits required to achieve this level of statistical precision. These calculations can be quite complex for mystery shopping visits, however, as they need to take into account agent- and scenario-specific factors as well as characteristics of outcomes you are interested in measuring, and the type of subgroup analysis you plan (for example, comparing outcomes between rural and urban agents or male and female consumers), and the size of expected differences between these groups. As a point of reference, the TCI fieldwork typically involved sample sizes of about 400 agents per country and 1,000 - 2,000 surveys or mystery shopping visits per country.

LINKS TO RESOURCES:

Fieldwork methods

The following resources are available in the Online Appendix for those interested in fieldwork methods.

- Agent census survey instrument
- Customer intercept survey instrument
- Professional enumerator mystery shopping instrument
- Local consumer recruitment survey instrument (for recruiting local mystery shoppers)
- Local consumer mystery shopping instrument

Photo: Mark Pickens

ii) Sampling approaches

Market selection is dependent on both the questions you seek to answer and your resource constraints. We suggest three alternatives:

1. **Selection of urban markets only:** Conducting fieldwork only in a few large urban markets would allow you to achieve a relatively large sample size without incurring high travel costs. However, if you believe the consumer experience with DFS agents differs by geography – perhaps urban and rural agents are systematically different or differences exist by region – then restricting data collection to urban centers will bias your results somewhat. Still, if resources are limited or you have confidence that geographic differences in consumer outcomes are limited, this offers a viable option. Of course, if you are exclusively interested in the experience of rural consumers, you might choose to only conduct fieldwork in rural areas, though this would be more resource intensive.

2. **Purposive market selection:** If the research team believes that there may be substantial geographic variation in consumer outcomes, one relatively low-cost approach that allows for that geographic variation to be captured is to purposively select markets with some geographic variation. For example, if capturing differences in urban and rural consumer outcomes is a top priority, you might split your sample equally between urban and rural markets, concentrating both the urban and rural market into one or two regions each to minimize field travel costs. This approach would allow you to capture differences by urbanicity but not other regional differences.

3. **Nationally representative market selection:** To generate results that can be confidently reported as nationally representative, you may decide to select a nationally representative set of markets. This typically involves working with a government statistics office to identify a representative set of enumeration areas (perhaps stratified by region, population density, and/or other characteristics of interest), and then selecting one market within each enumeration area. This approach produces the most representative results but will be the most resource intensive as field teams will be required to travel to locations across the country.

In our TCI fieldwork, we selected option two, purposive market selection. We elected to split our sample equally between urban and rural markets to maximize our ability to measure differences between urban and rural agents. To minimize field team travel costs, we conducted all urban data collection in each country's two largest cities and all rural data collection in two more rural regions. To ensure we selected rural locations with varying population densities, we used census data to select regions at the 25th and 50th percentile in terms of population density (where lower percentiles represent regions with lower population densities).

Once a broad sampling strategy has been identified, you need to select individual markets to visit. Ideally, these should be a random, representative set of markets in the locations you elect to carry out fieldwork. If a dataset is available containing information about individual agents and their locations – an agent registry, for example – this can be used to generate clusters of agents and then a set of these clusters can be randomly selected. If this is not available, population clusters can be identified through geospatial analysis and a random subset of these clusters can be selected, with the assumption that most markets where agents are concentrated are positioned near population centers. Finally, purposive sampling offers a simpler approach, though it will generate a less representative sample. In this case, local field teams may be empowered to select markets that they believe represent typical markets in each area selected for fieldwork.

In our TCI fieldwork, because a complete listing of individual agents was not available in the countries we worked, we chose to use a combination of geospatial analysis and purposive sampling to select individual markets. In rural areas, we chose to use geospatial analysis (including both night light data and 'built up area' data) as this allowed us to select markets in areas with varying population densities, from very small villages to mid-sized towns, which was important for us as we were interested in understanding how varying levels of agent competition influenced consumer outcomes. A detailed technical description of the geospatial analysis techniques we deployed is available online here. In urban areas, where we were less concerned about differences in agent density, we employed a purposive sample. We instructed field teams to select five 'typical' market locations in each city we included in our fieldwork, one in the central business district and one in each of the four cardinal directions emanating out from the central business district.

iii) Defining individual markets

Once a sampling strategy has been determined to identify markets, these markets need to be clearly and consistently defined. This market definition is used to define the boundaries of a census of agents in each study market, described in the next section. There are many ways that markets may be defined. We describe the approach taken in our TCI fieldwork, though many variations are possible and we do not claim our approach is best for all situations. In rural areas, we defined markets as circles centered around central points identified by field teams (typically the center of a village or market area) with a radius of 300 meters. The research team selected this radius as it generates markets that can typically be walked across in less than 10 minutes. Because we were particularly interested in competition between agents, and we expect that consumers are unlikely to walk more than ten minutes to choose a preferred agent, we felt this radius was appropriate. In urban areas, using the same definition would have in some cases yielded markets with very large numbers of agents. To avoid this situation, we set a cap of 20 agents per market, selecting the 20 closest agents to the central point.

Agent selection

After defining the market boundaries, researchers must determine what constitutes an agent and then decide on how agents will be selected within these markets. An agent can be defined either by the individual managing the location or by the physical location itself. Given the high turnover rate among agents and the likelihood that a single location can be staffed by multiple employees working different shifts—making it challenging to time visits to specific individuals—we recommend defining an 'agent' as the physical location rather than the individual providing services. This section describes the process of selecting these physical agent locations for inclusion in your study.

i) Generating a list of agents in each market

When it comes to agent selection, we recommend one of two approaches: if comprehensive agent lists from providers are available, these can be used as the basis for selecting agents, otherwise you should conduct a census of the agents within the defined market boundaries.

For the TCI fieldwork, we decided to conduct an in-person census and include all agents within the boundaries of selected markets, and not select a random sample. During the census, it is essential to collect information that ensures the accurate reidentification of selected agents. This is especially important in markets with high agent density, where many locations may appear similar, though it is valuable in all situations. Here are the three key options available based on our experience:

1. **GPS Coordinates:** These can be recorded during the census, provided a device that is able to collect accurate geo coordinates and the enumerator is close to the agent's location. These coordinates can later be loaded into the survey, allowing enumerators to be directed to the location via Google Maps or a similar application. GPS is most useful in rural areas where agents are dispersed and harder to locate. However, in crowded markets, GPS coordinates will not be as effective in differentiating between agents positioned just a few steps apart.

2. **Written descriptions:** In markets where GPS alone is insufficient, detailed written descriptions become essential. These descriptions should include information about surrounding landmarks (e.g., buildings, shops, infrastructure) and details about the agent's location (e.g., type of building, building material, color, signposts).

3. **Photos:** Photos can further confirm that the correct agent has been located. However, when using photos, it's important to consider the potential risks, as agents may be sensitive about being photographed, especially given that they often hold significant cash on hand and so are wary of theft. Taking pictures could raise suspicion or alarm, which can be problematic both for enumerators conducting the census as well as the field teams that will later conduct surveys or mystery shopping visits in the area.

The census survey can also be used to capture location characteristics for analysis, such as whether the agent operates from a makeshift or permanent structure, and whether the location is shared with another business or is a standalone business, as these factors can significantly impact agent availability, cost structure, as well as customers' perception and trust.

ii) Selecting agents for inclusion in the research sample

Once you have a complete list of agents in each market, you will need to decide whether to include all agents or to use a sample. This decision depends on several factors:

Variation between markets vs. within markets: Consider whether the greatest variation in agent behavior occurs between different markets or among agents within the same market. If most variation is between markets and agents within a single market behave similarly, it is more important to cover a wide range of markets, sampling only a few agents in each. However, if there is significant variation within a single market, it becomes crucial to sample a larger number of agents within each market. All else being equal, visiting fewer markets and sampling more agents is more cost effective because of travel costs.

Market size: Larger markets may necessitate sampling due to the sheer number of agents involved. Attempting to include all agents in a large market can be impractical and resource-intensive.

Resource availability: The time, budget, and personnel available for the study will significantly impact your choice. Comprehensive coverage of all agents might be ideal but may not be feasible given resource constraints. In such cases, sampling becomes a practical solution.

Study objectives: Your study's objectives will guide your decision. If the goal is to capture a complete picture of agent behavior and market dynamics, including all agents might be necessary. However, if the aim is to generalize findings across the market, a well-designed sample should suffice.

If you decide on sampling, it is essential to make the sampling process random to ensure that your findings are unbiased and representative. To do this, you can use statistical software programs (R, STATA, python, etc.) that can generate and assign random numbers to the agents, and output a random sample based on your desired sample size.

Agent selection plays a critical role in determining where users will be intercepted for consumer intercept surveys or which agents will be visited in the case of mystery shopping. In the next section, we will discuss key tenets related to implementing these different field methods.

Method 2:

Consumer interviews

When implementing consumer intercept surveys, several important considerations must be taken into account. We discuss these below:

A. Target sample

Prior to conducting any consumer intercept surveys, it is crucial to establish clear eligibility criteria for your target sample. Participants should be of legal age to give consent to participate in the study, typically 18 years or older. Additionally, participants should have just attempted or completed a transaction at an agent location, considering the survey aims at capturing relevant and recent experiences with agents, including those related to failed transactions.

B. Recruitment of participants

Participants can be intercepted using a purely convenience-based approach, where intercepts happen one after the other, or a more systematic approach can be adopted, such as intercepting every nth person (for example, every 4th person that attempts a transaction) to reduce subjective decision making about who to intercept. However, a key limitation of the systematic approach is that low volumes of customers at agent locations can make this process slow and expensive.

You may be interested in particular subgroups within your sample. If so, it can be beneficial to set sample quotas based on specific characteristics to ensure sufficient variation, allowing for meaningful comparisons between these subgroups. Gender is often one such characteristic, and logistically, it can be easier to set quotas by gender because it can be observed by the enumerator before initiating a survey. This is more challenging for characteristics like education or socioeconomic status, which aren't as readily identifiable on sight. Additionally, the gender of the enumerators can influence participants' willingness to engage in the survey. For instance, female customers might feel more comfortable speaking with female enumerators, and the same could apply for male participants.

For our TCI fieldwork, participants were recruited based on two key criteria: they had to be 18 years or older, and they needed to be at the agent location to conduct a financial transaction, even if it was unsuccessful. Enumerators waited two to four hours for an eligible participant, and if none was found, they made one additional attempt on the same or a different day. After two such attempts, the intercept survey was discontinued at that location due to resource constraints. We also aimed to achieve gender balance, as we intended to recruit shoppers through these intercept surveys to understand variations in shopper experience by gender. This required enumerators to coordinate, adjusting whom to intercept based on the gender of previous participants to maintain as much gender balance as possible. In more conservative contexts, we preferred that enumerators intercept participants of the same gender as themselves, which helped facilitate more comfortable interactions.

C. Scenario design

When it comes to consumer intercept surveys, predetermined scenario designs are not possible. Researchers are not able to control key variables such as the providers customers choose to use, the transaction amounts involved, the types of transactions conducted, or the number of transactions a customer completes during their visit, or even which agents they choose to visit. This lack of control means that each customer interaction can vary significantly, but reflect real world consumer decisions more closely.

D. Questionnaire design

When designing and implementing a consumer intercept survey, there are several key considerations to keep in mind.

1. **Structure:** Demand-side surveys offer a unique opportunity to gather valuable demographic information about consumers, such as age, gender, educational level, and DFS usage patterns. Consumer intercept surveys are no different in this regard. However, given potential recall issues, it may be best to split the demographic questions into two parts: a brief set at the beginning to ease into the conversation, followed by the main outcome questions relating to the attempted transaction they just completed, and then additional demographic information at the end. This minimizes the time between the transaction and the questions about that transaction.
2. **Survey Length:** It's important to keep the survey brief, as customers intercepted outside agent locations are often on their way to other engagements. This is particularly true for urban respondents, who may become impatient and interrupt the survey if it is perceived as too long. A concise survey that is not more than 20 minutes also helps minimize recall issues, especially for questions related to mobile money fees, which require the respondent's focused attention.
3. **Sensitive Topics:** Consumers may be reluctant to share sensitive information about their transactions. For example, someone who has just withdrawn a large sum of cash might be unwilling to disclose this to a stranger. It is crucial for enumerators to approach these situations with care. While taking screenshots of respondents' digital transaction receipts can be a good way to verify fees paid and transaction amounts, this approach should be optional, and enumerators should avoid requesting screenshots if the respondent appears uncomfortable with the survey questions.
4. **Establishing credibility and rapport:** Enumerators' ability to establish credibility and rapport with respondents is key to achieving reliable outcome data from respondents. Enumerators should take the time to clearly explain the purpose of the study, ensuring that respondents understand the importance of their participation. Enumerators should also carry clear identification to establish credibility and build rapport, making respondents feel more comfortable. Additionally, conducting the interview in a relatively private space identified before intercepting the customer can help foster a sense of security, leading to more open and honest interactions.
5. **Making sure enumerators are familiar with the topics:** Enumerators should be well-prepared to assist respondents with complex questions, particularly those related to mobile money fees. Intercepted customers may struggle with understanding and recalling specific charges, such as provider fees, taxes, and informal fees, which need to be separated from the transaction value. Additionally, some concepts included in the survey, such as sexual harassment, may have nuanced translations in the local language. It is important to spend the time during the translation and back-translation to ensure these concepts are clear in the translation and are context-appropriate.

E. Compensation structure

When designing a compensation structure for participants in a consumer intercept survey, it is important to ensure that respondents are reasonably compensated for their time and effort. The compensation can be guided by the hourly wage rate for the country, calculating a proportionate amount based on what enumerators are being paid relative to the length of the survey, or benchmarking typical compensation structures for similar surveys within the country. For the TCI fieldwork, we benchmarked our compensation structures against those used in similar surveys conducted by our country offices.

It is essential that the compensation is approved by an accredited ethical review committee. This ensures that the compensation is fair and does not coerce respondents into participating solely because of the monetary incentive, compromising the voluntary nature of the survey.

Photo: Mark Pickens

Method 3:

Mystery shopping by professionals

In this section we will discuss consideration for implementing mystery shopping by professionals.

A. Shopper recruitment

When selecting professional shoppers, it's essential that they possess some form of post-secondary education. This educational background is necessary to help them grasp the complex concepts related to DFS much easier. Along with education, these shoppers should demonstrate strong digital literacy and have some prior experience with DFS, which will make the training process smoother and the learning curve less steep.

Language proficiency is also critical; the recruited shoppers must be able to speak the language typically used in the market with the agents such that they can blend in relatively easily and avoid detection as an outsider which can lead to agents modifying their behaviours.

For the TCI fieldwork, we recruited our shoppers from databases of enumerators that we have worked with on other surveys through our country offices. These enumerators typically hold a university degree, have digital literacy (considering that nearly all IPA surveys are conducted using SurveyCTO, a digital data collection platform), and we specifically selected them based on the region where the mystery shopping would take place and their ability to speak the local languages.

B. Shopper training

It's crucial to allocate sufficient time for training shoppers, ensuring a comprehensive and well-rounded preparation. The training should begin with a classroom setting where shoppers can understand the study objectives, thoroughly review the survey instrument, and engage in role plays, with one shopper acting as an agent and another as a customer.

Additionally, shoppers must gain field experience through piloting to fully embody their assigned profiles. They should also be provided with a script that is detailed enough to answer basic inquiries but vague enough to avoid revealing their true purpose in the market. This script should be rehearsed multiple times to ensure fluency and confidence.

Moreover, training should emphasize that shoppers should not visit markets in groups, dress to fit in, and be able to recall key survey questions without needing to fill out forms at the agent location. It should also capture aspects regarding how they would manage their cash-flow without raising alarms: this could be done through doing practice cash-in or cash out transactions. Based on our experience, this training process typically takes five to six days.

C. Scenario design

A key advantage of mystery shopping methods, as discussed earlier, is that they allow researchers to have greater control over various aspects of the mystery shopping visit, such as the types of transactions shoppers perform, the providers they interact with, the value of the transactions, and factors like customer-agent gender. In this section, we will discuss important considerations in scenario design.

DFS providers: Mystery shopping requires setting up and funding DFS accounts with the providers that will be used during the study. Using a limited number of providers simplifies this process. In practice, a small set of providers will capture most of the market share in most markets, so by limiting your selection to these key providers, you can still obtain results that are representative of the vast majority of the market.

For the TCI fieldwork, we conducted financial transactions with the largest mobile money providers in each of our focus countries, specifically targeting those that account for at least 80 percent of the overall market share. Information on market presence of providers can be found from official sources including reports from governments or financial institutions. In most cases, agents were not exclusive to a single provider, so we randomly assigned each agent to one provider, ensuring an equal distribution of providers at the market level. Professional shoppers would then make transactions using the provider to which the agent was assigned.

Transaction types: A key question to consider is whether to use the same transaction types across all countries or to develop country-specific transaction types. Choosing the latter will reduce the ability to make cross-country comparisons but will ensure that the transaction types closely mirror the typical user experience in each market. To gather information about typical transactions in each country, you could look at data from providers' websites, mobile money regulators, or consumer surveys. For the TCI fieldwork, we used data from consumer intercept surveys to inform both the types of transactions conducted and their distribution. Below are four transaction scenarios we deployed.

TABLE 5: Common mobile money transaction scenarios

TRANSACTION TYPE	DESCRIPTION
Cash-in	Depositing physical cash into a mobile money account.
Cash-out	Withdrawing physical cash from a mobile money account.
Over-the-counter cash-to-account transfer	Mystery shopper brings physical cash and asks the agent to send it to the recipient's mobile money account.
Account-to-account transfer	Mystery shopper asks the agent for assistance in transferring electronic funds from their mobile money account to the recipient's account, either with a different mobile money provider ("off-network") or with the same provider ("on-network").

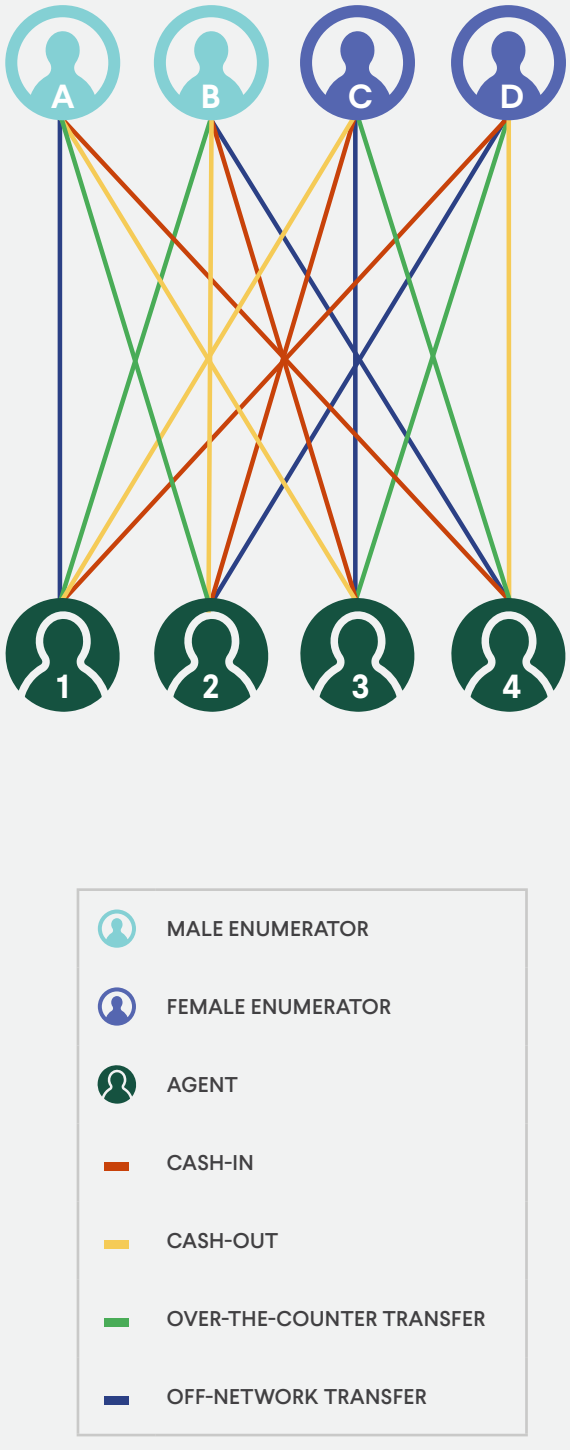
Transaction values: The choice of transaction value to use during the mystery shopping is crucial. You have the option of setting a single USD value, as done in the World Bank's remittance work, or finding a way to identify typical transaction values in each country. While a single USD value would be straightforward, it wouldn't reflect the user experience in each country, given the different levels of economic development and varying typical transaction sizes. Additionally, given the regressive nature of mobile money fees in many countries, it would be beneficial to use two values—one high and one low—to capture variations in pricing more effectively.

For the TCI fieldwork, we opted for a more tailored approach. We found from IPA's consumer protection surveys that median transaction sizes often equate to roughly 15 times the mean daily income per capita for the bottom 40 percent of the population. Although this is a rough approximation, we believe it more accurately reflects typical transaction sizes than a global USD value, and we aimed to replicate this for each country. To implement this approximation, we used World Bank data on the daily mean income per capita for the bottom 40 percent of the population. We then converted this value to local currency using the 2017 World Bank PPP conversion factor, adjusted it for inflation to current local currency using local CPI levels, and multiplied the income per capita by 15 to obtain our high value, and half of this high value constituted our low value.

Visit schedule: In designing the visit schedule, the number of visits per agent and enumerator involves balancing the benefits of learning as much as possible across multiple interactions with agents against the costs, complexity, and risk of observer effects (which can arise from unusual transaction patterns). During each visit, you may vary several elements, such as the shopper's gender, transaction type, transaction value, and the provider used. These variations create different scenario types that can be highly informative. We recommend that different shoppers take on different scenario types, allowing for only a single interaction with each agent.

For the TCI fieldwork, we had four enumerators active in each market—two females and two males. Each enumerator visited an agent once, attempting one of the transaction types. This resulted in four visits per agent (two by female shoppers and two by male shoppers), ensuring that every agent was exposed to all transaction scenarios. Half of the visits per agent were conducted with a low transaction value and the other half with a high value. See Figure 2.

FIGURE 2: Enumerator to agent allocation



It is also important to structure visits in a way that minimizes non-research transactions. For instance, a mystery shopper might need to withdraw money to complete an assigned cash-in scenario. If no agents outside the study are available, this adds an extra visit, increasing the risk of observer effects and adding non-research-related costs, which can accumulate quickly in markets with many agents and may require "top-ups" of shopper funds. One way to mitigate this is by structuring the research transactions so that a cash-out transaction is conducted first, followed by other officially "free" transaction types, and then a final cash-in transaction. However, this approach may not always work, as some transaction attempts may fail (either because the agent is unavailable or the transaction is unsuccessful).

D. Questionnaire design

Structure: The mystery shopping survey should include two key sections. The first section is the setup, which contains directions to the agents, information on the assigned scenario, transaction value, and provider, as well as mobile money balance checks, and this should be completed before the visit. The second section covers the core visit experience, focusing on transaction charges, agent behavior, and disclosures, and is to be filled out after the transaction attempt has been completed.

Given the complexity of visit assignments, which involve varying transaction type, transaction value, provider, and agent, the survey should be prefilled with as much information as possible to minimize shopper inputs and reduce errors. For example, details such as the transaction type, value, and agent ID should all be auto-populated in the survey.

Survey length: Mystery shopping surveys are typically filled out from memory after the visit to an agent. To reduce errors due to recall issues, it's important to keep the survey to a length that can be completed comfortably within 10-15 minutes.

E. Compensation structure

Ideally, your compensation structure should boost the shopper's productivity and incentivize them to use their market knowledge to identify the best times for making successful transactions at agent locations. However, implementing productivity-based compensation with mystery shoppers can be challenging, as low productivity (measured by the number of successful transactions) might be due to factors beyond their control. Additionally, professional shoppers may lack the local knowledge needed to determine when agents are most likely to be available.

For the TCI fieldwork, enumerators were hired for a full day and compensated regardless of whether a transaction was successful. Our compensation was based on wage rates from similar surveys conducted by our country offices. In your case, you could determine enumerator wage rates by referencing rates from similar surveys conducted by other institutions or by considering the national labor rate for comparable job categories.

Photo: Nichika Yoshida

Method 4:

Mystery shopping by local consumers

In this section, we discuss key elements related to deploying mystery shopping using local consumers, starting with how to recruit the consumers, followed by designing scenarios, survey instrument structure, and establishing a compensation structure.

A. Recruitment of participants

As a first step, you will need to determine how to recruit shoppers. You have the option of conducting a household survey in homes close to the market or using consumer intercept surveys.

Consumer intercept surveys can be slow, particularly in rural areas where traffic at agent locations is low, leading to higher recruitment costs. Additionally, these surveys may skew the sample toward "heavy users," who may not represent the typical customer. On the other hand, household surveys may skew the sample toward individuals more likely to be at home, which often means more women. To address this, you would need to make additional efforts to reach excluded users at times they are more likely to be at home—such as over the weekend or early in the morning before they leave.

Another challenge you may encounter is how to create scenarios that match customers with agents they regularly use or have a relationship with. One way to address this is by having shoppers visit all agents in the market and then, after the visit, indicate whether they regularly use the agent or not.

For the TCI fieldwork, we initially recruited shoppers through consumer intercept surveys but later shifted to household surveys to reduce recruitment costs and achieve better representation of shoppers. To conduct the household survey, enumerators used the central point of the market as a starting point and then visited the nearest residential area for a door-to-door survey. Respondents were eligible if they were:

1. Aged 18 years and above,
2. Had a mobile money account,
3. Didn't work as mobile money agents,
4. Had some basic numeracy skills determined through a simple test that asks respondents to subtract fees from a common transaction value.

For efficiency, only individuals present in the household at the time of the interview were considered. We recruited one person from each household with the goal of achieving gender balance in each market. To account for potential attrition, we also over-recruited (aiming at 50 percent above the desired number of mystery shoppers) both within households and across the village, ensuring we had backups that could be randomly selected if needed.

B. Training

When it comes to training local shoppers, there are two approaches you may consider:

Just-in-time training: This involves training shoppers immediately before their visits. The key advantage of this approach is its flexibility, as it allows you to train shoppers when they are available, accommodating their varying schedules.

Classroom training: This involves gathering all shoppers at once for training in a classroom setting. The primary benefit of this method is that it allows for more control over what everyone is learning, ensuring consistency across all participants. However, this approach can be challenging to implement in practice due to the differing availability of shoppers.

Regardless of the approach, it is essential that shoppers are given the opportunity to rehearse their profiles and conduct some "dummy" transactions before embarking on the mystery shopping field exercise. Enumerators should also shadow them at a distance during the actual exercise and offer support in filling out the survey.

In our TCI fieldwork, we assigned two enumerators to each market to support the local mystery shopping activities. They were responsible for providing just-in-time training to the shoppers before they conducted their visits, ensuring the shoppers had dummy sessions to practice and feel comfortable. The training sessions were conducted outside the market to avoid alerting the agents. During the training, we emphasized the importance of being discreet and focused particularly on understanding the different types of mobile money fees and charges. For an example of a training guide, see the appendix.

C. Scenario design

DFS providers: For local shoppers, the providers they use for their scenarios will be dictated by the accounts they already have rather than the providers you might assign to the agents. In cases where there is a mismatch, there isn't much that can be done, especially if the goal is for the shopper to visit every agent in the market. However, if you plan to visit only a subset of agents, you can use the agent census data to match shoppers with agents who serve the same providers.

Alternatively, you could ask shoppers to open accounts with multiple providers, but they may not be comfortable doing so, and depending on the country, this process can be complex. In our TCI fieldwork, for consistency, we allowed shoppers to decide which provider they would use ahead of their visit to an agent, reducing the complexity of managing multiple accounts. This approach worked well since we needed the shoppers to visit all agents in the market, and fortunately, many agents were not exclusive to a single provider.

Transaction types: To determine the types of transactions to include in your mystery shopping scenarios, you have several options. You can reference data from providers, regulators, or similar surveys that, while deployed for different purposes, capture information on typical transactions. Another approach is to use data from the household recruitment survey, staggering the survey to allow this information to be incorporated into the scenario design.

For the TCI fieldwork, we used data from our recruitment household surveys and set up a visit assignment workflow that allowed us to create individual shopper scenarios almost immediately after the recruitment process was completed. For example, if 50 percent of survey respondents reported that their most recent agent transaction was a cash-out, we would assign 50 percent of all scenario visits as cash-outs. Similarly, if no respondents reported performing an account-to-account transfer, that transaction type was excluded from the mystery shopping scenarios.

Transaction values: While there are advantages to having a high and low transaction value, local shopping using a single value would reduce the complexity and ensure shoppers don't mix things up. The median transaction amount as reported by the household survey respondents can be used instead of two values, to reduce complexity.

For the TCI fieldwork, we used the median transaction value observed in our consumer intercept surveys and verified it with field teams to ensure it was a typical transaction amount. If this value was a decimal or not representative of a typical transaction, we rounded it to the nearest whole number that was more typical.

Visit schedule: You might consider having each shopper conduct either a single visit per agent or multiple visits to the same agents. Multiple visits per agent can increase suspicion and may lead to shoppers mixing up their visits. On the other hand, single visits can help offset these challenges.

Additionally, you can decide whether to have shoppers visit a subset of agents within the market or all the agents in the market. Visiting a subset might be the best option if the market is large with many agents. However, visiting all agents in the market can provide insights into how outcomes vary depending on the shopper-agent relationship, as the shopper would encounter both agents they use regularly and unknown agents. This approach can be useful in understanding shopper behavior after exposure to better-quality agents and in assessing whether there is any switching of agents as a result.

In our TCI fieldwork, shoppers made a single visit to all the agents within the market because we were interested in understanding the shopper-agent relationships and how they affect the outcomes we were examining. Additionally, to minimize the risk of shoppers being detected, we opted for single visits to each agent.

D. Questionnaire design

Structure: Similar to the discussion of the survey structure for professional shoppers, the mystery shopping survey

for local shoppers should include two key sections. The first section is the setup, which contains directions to the agents, details of the assigned scenario, transaction value, and provider, as well as a mobile money balance check. This section should be completed before the visit. The second section focuses on the core visit experience, covering transaction charges, agent behavior, and disclosures, and should be completed after the transaction attempt.

To simplify the process, given the complexity of visit assignments involving different transaction types, values, providers, and agents, the survey should be prefilled with as much information as possible. For instance, details like transaction type, value, and agent ID should be auto-populated in the survey to minimize shopper input and reduce errors.

Survey length: Like the professional mystery shopping survey discussed earlier, the local shopper survey should be kept short, ideally no longer than 15 minutes. This ensures that there are no recall issues, as these surveys are typically filled out from memory after the visit.

E. Compensation structure

As discussed in the professional mystery shopper section, your compensation structure should ideally boost the shopper's productivity and incentivize them to use their market knowledge to identify the best times for making successful transactions at agent locations. A significant advantage of using local shoppers is that they are familiar with the market dynamics, such as when agents are typically available. You can leverage this knowledge by creating incentives beyond the base payment that encourage shoppers to use their local insights when determining the best times to visit agents. Additionally, it's important to ensure that shoppers complete all their visits and do not drop out after receiving their first payment.

Payments to local shoppers should consist of four components, two of which are disbursed before any mystery shopping begins, and two of which are disbursed after the completion of all attempted mystery shopping visits:

Disbursed before mystery shopping begins

Transaction value: This is the amount shoppers will use for all their transactions across the different agents they are visiting.

Funds for fees and charges: This includes funds to cover transaction fees and charges, both for the assigned transactions and any potential non-research transactions that need to be completed (for example, if the shopper needs to withdraw money in preparation for an assigned cash-in transaction). This amount is given in advance and topped up by enumerators if necessary.

Disbursed after mystery shopping is completed

Per-visit payments: Shoppers should be paid for each mystery shopping visit they attempt, both to compensate them for their time and to incentivise participation. We recommend making all per-visit payments at the end of the exercise, both to encourage completion of all activities and reduce the administrative burden of making multiple small payments. We also recommend making smaller payments for mystery shopping visits when the agent is not present than visits when the agent is present. This encourages shoppers to make visits when they believe agents are likely to be available (mirroring their real-world behavior) and discourages shoppers from strategically visiting at times when they know the agent is unlikely to be present. We recommend payments for visits when the agent was not present to be approximately 30 percent of the value of per-visit payments when the agent is present.

Completion payment: This is an incentive payment dispersed after all visits have been completed, designed to motivate shoppers to carry out all assigned visits.

Here is a formula you can use to determine total compensation to the shoppers:

Payment before mystery shopping begins = [Transaction Value] + [Funds for Fees and Charges]

Payment after mystery shopping is completed = [Per-visit Payment]*[No. visits where the agent was present] + 0.3*[Per-visit Payment]*[No. visits where the agent was absent] + [Completion Payment].

Where:

[Transaction Value] + [Funds for Fees and Charge should be set based on the scenario and typical market fees.

[Per-visit Payment] and [Completion Payment] should be calibrated to local survey respondent payment norms.

The sum of expected [Per-visit Payment] and [Completion Payment] should be more than the payment made before the mystery shopping begins ([Transaction Value] + [Funds for Fees and Charges]) to discourage taking the upfront payment without completing the work.

Data quality assurance, management and analysis

In this section, we will explore key aspects of data quality assurance, data management, and analysis. We specify important considerations to ensure the accuracy of data collected. Additionally, we will discuss crucial aspects related to data analysis.These considerations generally apply regardless of method selected.

i) Software considerations

Based on IPA's experience, using paper surveys to collect data is not only laborious but also challenging in maintaining real-time data quality. Therefore, we recommend collecting data electronically using specialized software. Several options are available, including SurveyMonkey, Google Forms, SurveyCTO, KoboToolbox, and even WhatsApp surveys via Twilio. The choice of software will depend on factors such as budget, team familiarity, data security, survey complexity, and the level of after-sales support. Initially, we attempted to deploy WhatsApp surveys using Twilio, hoping to leverage a platform familiar to local shoppers. However, due to the survey's complexity and challenges in programming constraints within Twilio, we opted for SurveyCTO. Although SurveyCTO is not open-source and costs about $200 per month to host a server, it offers robust data quality controls and dependable support.

For data monitoring, cleaning, and analysis, a range of software options are also available, including R, Python, SPSS, STATA, and MS Excel. We generally advise against using Excel for large datasets as it can be inefficient and prone to errors. In our case, we chose STATA given our team's proficiency with STATA, its robust data management and statistical tools, and because our data assurance programs that have been refined over the years are STATA based.

ii) Bench testing

Conducting bench tests of the survey instrument prior to its launch is essential for ensuring that the tool functions correctly, that the survey flow operates as intended, and that any programming, formatting, or language errors are identified and corrected. This process involves having a few team members simulate real interviews and submit test responses. Rigorous bench testing is critical because it helps prevent survey-related issues from emerging during fieldwork, where they would be far more time-consuming and challenging to resolve.

iii) Survey translations and back-translations

When a survey is conducted in a language or dialect different from the one in which it was originally written, translation becomes necessary—sometimes into multiple languages, depending on the survey area. While hiring a professional translator is one option, it can often be more practical to have members of the field team handle the translation. If field staff are tasked with translation, it is beneficial for several team members to collaborate on the process. This allows them to reach a consensus on the most appropriate and easily understood translations for specific words and concepts.

Ensuring consistency between the translations and the original language is of utmost importance to guarantee that respondents correctly understand the questions. To further ensure accuracy, a back-translation should be performed. This involves having someone who was not involved in the initial translation translate the survey back into the original language. By comparing the back-translated version with the original, any discrepancies in words or concepts can be identified and addressed.

3. Because tracking of official prices does not involve substantial fieldwork, some tools discussed in this section do not apply to this method.

iv) High frequency checks

A high-frequency check (HFC) is a systematic review of certain elements in the data collection process, conducted regularly (preferably daily) as new data is received in order to address any emerging issues. These checks are crucial for assessing the quality of the data collected, identifying potential threats to validity, evaluating enumerator performance, and pinpointing data flow issues that could stem from the survey programming. It's important to perform these checks daily and address any emerging issues promptly. In the TCI fieldwork, we established a Stata file that generated an Excel sheet to facilitate this process. This output included the number of surveys completed in each market to ensure targets were met, outlier values (especially for numeric variables like transaction fees), distances to agents within the market based on GPS coordinates to verify agent localization, and survey durations, with exceptionally long or short durations flagged for further investigation with the responsible enumerator.

v) Using screenshots of transaction receipts for verification

Screenshots of transaction receipts can be an effective tool for cross-checking transaction details. These receipts may be in the form of automated text messages received after an agent transaction, such as sending money, depositing, or withdrawing funds. The information contained in these receipts—such as transaction amounts and official provider fees—offers an opportunity to verify the accuracy of survey responses to related questions.

In our experience, accurately entering transaction fees has been a particular challenge for both enumerators and local shoppers. Therefore, using screenshots to cross-check inconsistencies can be especially valuable, at least for fees that are directly deducted from mobile money accounts (as opposed to cash-based fees). It is advisable to upload these screenshots directly into the survey tool to ensure that each receipt is properly linked to the corresponding survey submission. Verification can then be performed manually or by using software to transcribe and match the information in the images with the survey data. We provide an example of such verification process in the online appendix.

vi) Daily team briefings

These daily meetings between enumerators and senior field staff are crucial for maintaining data quality assurance protocols. They not only ensure the accuracy and consistency of the data collected but also serve to motivate the field team and promptly address any problems or concerns. A key component of these briefings is the identification and resolution of potential issues arising from the HFCs, allowing for swift management of any emerging challenges.

vii) Data cleaning

After fieldwork concludes, the dataset requires a final check and 'cleaning' to prepare it for analysis. This process involves a detailed inspection of each variable to identify and rectify potential errors, often with the help of the enumerators who collected the data. It's also essential to establish consistent, unique identifiers for interviewees, geographical areas, and similar entities. Non-essential variables that clutter the dataset can be dropped, and categorical variables with numerous options but sparse data—like educational levels—can be condensed into fewer categories.

To enhance usability, particularly in collaborative settings, renaming variables for consistency and providing them with descriptive labels is advised. Creating a codebook that details each variable's name, a more elaborate description, and the corresponding survey question is a crucial step towards this end. Additionally, you may want to create new variables for analysis, such as indices or indicators. For instance, in our data preparation, we developed indicators of overcharging by comparing the fees reported by enumerators or consumers to the official fees listed on the providers' websites. We also calculated a poverty probability index to get a standardized measure of respondents' economic status.

viii) Data analysis

Prior to the launch of the survey, it is crucial to establish a clear plan for the key outcomes of interest. This plan guides decisions regarding which variables to include and how to segment your analysis. In our TCI fieldwork, we focused on both the monetary and non-monetary costs associated with mobile money transactions. The specific outcomes of interest included transaction charges, success rates, service quality, and pricing transparency. Furthermore, we aimed to understand how these outcomes varied between rural and urban areas, and among different consumer segments. Key questions we explored included whether poorer consumers face higher costs and how costs differ by gender, educational level, or age.

These questions can be addressed two ways. One method is regression analysis, which allows for the identification of significant correlations between variables while controlling for potential confounding factors. Alternatively, a descriptive analysis might suffice, involving summary statistics segmented by the desired groups, along with statistical tests to assess the significance of differences between these groups. In our experience, a descriptive approach often provided sufficient insights for our data. Ultimately, the choice of analysis technique and the level of detail necessary will depend on the intended audience for the research findings.

Online appendix

The online appendix can be found here.
Content and direct links are laid out below.

Appendix 1: Scraping of official prices

A. Converting online pricing data to excel format

B. Appending provider and country datasets to one masterfile

C. Compile Codebook

Appendix 2: Transaction verification resources

A. TCI Transaction Verification Tool.R

B. TCI Transaction Verification Tool Documentation.pdf

Appendix 3: Fieldwork questionnaires

A. Agent census survey.pdf

B. Consumer intercept survey.pdf

C. Professional enumerator mystery shopping survey.pdf

D. Local consumer recruitment survey.pdf

E. Local consumer mystery shopping survey.pdf

F. Local consumer mystery shopping training guide.pdf


ipa
Innovations for
Poverty Action