The M-PESA Household Survey- Data Notes

Compiled from related research papers and data notes by Tavneet Suri and William Jack

Contents

A. **M-PESA Overview**
B. **Project Motivation and Partners**
C. **Details on Enclosed Datasets**
D. **Survey Design and Instruments**
E. **Survey and Sampling Methodology**
F. **Related Research**

A. M-PESA Overview

Mobile money is an innovation that allows individuals to store, send, and receive money on their mobile phone via text message (SMS)[1]. Launched in 2007 by Safaricom, the dominant mobile network operator in Kenya, M-PESA[2] is the most widely adopted mobile phone-based financial service in the world. This explosive growth has opened up basic financial services to many who were previously excluded.

Mobile money systems have fundamentally altered the extent to which resources can be transferred among households and individuals in Kenya. M-PESA facilitates safe storage and transfer of money, and makes it easier for people to pay and receive payment for goods and services while saving time and minimizing risks of carrying large sums of money. It also allows for more efficient risk sharing through social networks and enables support to keep negative shocks manageable through timely money transfers. Based on these surveys, our research finds that better access to M-PESA:

- Dramatically lowers transaction costs;
- Increases savings;
- Increases consumption and reduces poverty;
- Does not increase (physical) assets;
- Increases use of a bank account;
- Changes occupational choice (less likely to be a farmer, more likely to be in a business), especially for women in the household

Mobile payment systems have also been developed in the Philippines, Afghanistan, Sudan, Ghana, and in a number of countries in Latin America and the Middle East[3]. M-PESA itself has been started in Tanzania and South Africa.

[1] http://www.pnas.org/content/109/26/10257.full.pdf
[2] "M" is for mobile, and "PESA" means money in Swahili
[3] Mas, Ignacio. 2009. "The Economics of Branchless Banking." Innovations 4 (2): 57–75.C Ivatury, Gautam, and
[2] "M" is for mobile, and "PESA" means money in Swahili
[3] Mas, Ignacio. 2009. "The Economics of Branchless Banking." Innovations 4 (2): 57–75.C Ivatury, Gautam, and Mark Pickens. 2006. Mobile Phone Banking and Low-Income Customers: Evidence from South Africa. Washington, DC: Consultative Group to Assist the Poor.

B. Project Motivation and Partners

The key motivation of this project was to understand the impact and influence of M-PESA adoption, and what drives take up. Given the low penetration of banking services in Kenya, and the high costs of money transfers from existing non-bank providers, M-PESA promised to deliver a broad range of financial and money management and transfer services.

The household surveys were thus designed to quantify and track the short and long term impacts of M-PESA on household transactions and welfare. They enable a deeper understanding of the characteristics of interpersonal transactions including consumption trends, investments, remittances, and response to shocks. The surveys also facilitate the study of financial interactions and risk sharing within networks and the effects of facilitated remittances on household migration decisions. Thanks to our partners, we were able to administer 5 survey rounds between 2008 and 2014.

This survey data has been used in several research pieces. A full list of related research can be found in section F below.

C. Details on Enclosed Datasets

The household level data files from all five survey rounds are enclosed in Stata format. Section D includes a table with survey contents, corresponding sections, and data file names across rounds.

In accordance with IRB requirements, we dropped all variables that contain personally identifiable information (PII). These include variables on demographics, location, bank information, and a few string variables that fall under the category of "other, specify" that contained personal information. Variables dropped include: subject's name, phone number, village, district, location, sub-location, town, and bank name. In sections on remittances, the surveys contain questions on the closest town to the recipient / sender, and the distance separating them. For these questions, we drop the town name and keep the distance.

To conserve the information on the sample's demographics (section 2 in all rounds), we constructed a household level demographics section. This section contains the share of household members who can read and write, the size of the household, and the share of male and female members of different age brackets. This section also has information on the household head, such as the head's age, gender, marital status, and highest level of education. Across the 5 rounds, there were two occurrences where a household had more than one head listed. In these cases, we report as head the older member. This section also

contains the number of family members who have relocated within each household for round 5, the only round where this question was asked.

The round 1 questionnaire has a section on demographics of family members not residing in the household. We therefore construct a household level demographics section for non-residing members for round 1.

D. Survey Design and Instruments

Contents

The interviews collected information on basic household composition and demographic information, data on household wealth and assets, consumption, positive and negative shocks, and remittances (sent/received). They also covered questions on the use of financial services, savings, and collected data on cell phone use and general knowledge as well as M-PESA use in particular. They were administered to heads of households or their spouses in case they were absent after three visits.

This *dataverse* includes the instruments (questionnaires) used throughout the five rounds of the household survey. A separate document: "Survey Contents, Sections, & File Names" will help you link the survey contents to questionnaire sections and their respective data files. The *dataverse* also includes a codebook for each survey with all variable names and definitions for your reference.

E. Survey and Sampling Methodology[4]

Overview

The surveys were conducted in August-October of 2008, October 2009 -January of 2010, May-August of 2010 March-June of 2011 and a fifth survey to explore longer term impact was conducted in June-September of 2014. Several partnering organizations participated in the survey administration including Financial Sector Deepening (FSD Kenya), Steadman Group (R1), Synovate (R2), Kimetrica (R3), Financial Sector Deepening in Nairobi and the Consortium on Financial Services and Poverty in the US (R4-5).

Table 1: Survey Timeframe and Key Indicators across Rounds

	Round 1	Round 2	Round 3	Round 4	Round 5
Timeframe	Aug-Oct 2008	Oct-2009 Jan 2010	May-Aug 2010	Mar-June 2011	June-Sep 2014
Sample Households	3,000	2,016	1,531	1,649	1,688
Number of Accounts (million users)	4	8	13	14	25.4
Number of Agents	4,000	16,000	20,000	28,000	125,000

Sampling

With the exception of 8% of the national population in areas of Northern Kenya, which had to be dropped for logistical reasons, the surveys were nationally representative.

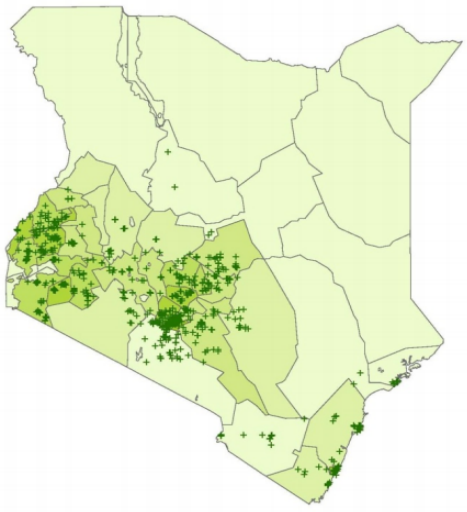
In 2008, we undertook a survey of **3,000** randomly selected households across a large part of Kenya. At the time, both cell phone tower and M-PESA agent coverage were very limited in the sparsely populated northern and northeastern parts of the country, so these districts (covering 8 percent of the population) were excluded from the sampling frame. From the remaining districts, we randomly selected 118 locations [5] (the second-smallest administrative unit), with at least one M-PESA agent. In these locations, were 300 enumeration areas routinely visited by the Kenyan National Bureau of Statistics. In order to

[4] http://fsdkenya.org/dataset/m-pesa-panel-survey-kenya-2014/ and http://fsdkenya.org/wp-content/uploads/2015/11/FSD-Research-Brief-Nov-2015.pdf

[5] Kenya is divided into districts, then divisions, then about 2,400 locations and further about 6,600 sub-locations. The average population of each location is about 3,000 households

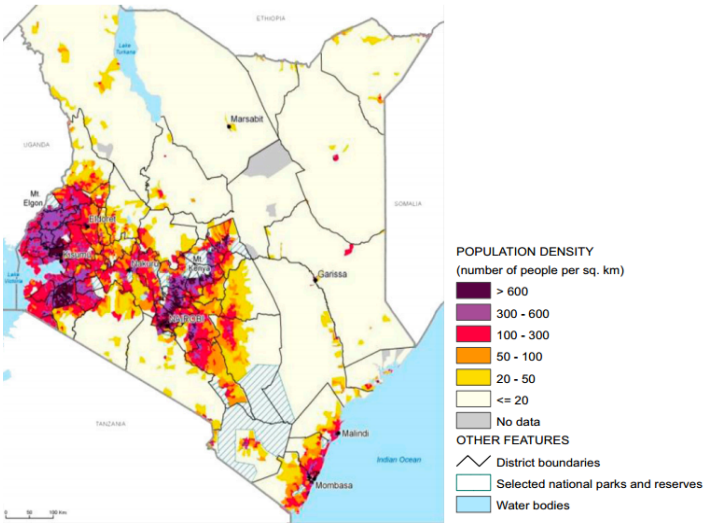
increase our chances of interviewing households with M-PESA users, we oversampled locations on the basis of the number of M-PESA agents present in that location. We sampled ten households randomly from each of these enumeration areas to take part in the survey. The GPS-recorded locations of the households are shown in figure 1 below, and you can find the population density in Kenya in figure 2. Follow-up surveys of the same households were administered in December 2009 and June 2010. Attrition rates were high but relatively comparable to similar surveys. We also designed the interview strategy for the third round with an eye toward finding households missed in the second round. In 2009 we re-interviewed 2,017 households, and in 2010 we were able to find 1,595 of the original sample, 265 of whom were not interviewed in 2009. In 2011, we interviewed 1,649 and were able to secure 1,688 household interviews for the final round in 2014. Data collection for this final round was done electronically.

Fig. 1 Interviewed Households*



" lighter areas have higher poverty rates

Fig. 2 Population Density in Kenya[6]



There was an unanticipated increase in the use of M-PESA over the period spanned by the two rounds of the survey. The number of users is measured by the total number of SIM cards registered to M-PESA. Additionally, this expansion was consistent with an expansion in the number of M-PESA agents. At the time of the first round of the survey, there were just over 4,000 agents across the country. By the fifth round in 2014, which we specifically undertook to measure longer-term impacts of M-PESA, agents across the

6 http://www.pnas.org/content/suppl/2012/06/11/1115843109.DCSupplemental/pnas.201115843SI.pdf#nameddest=STXT

country reached 125,000 and the growth in users reached more than 25 million (60% of the population).

Agent Mapping

In addition to these rounds, M-PESA agents across the country were surveyed and their GPS locations recorded. They were also asked the dates on which they first conducted M-PESA business. This sample covered the entire population of agents in each of the administrative locations from which our household sample was drawn. This allowed us to construct detailed rollout data on the agents, and to determine when our households first got easy access to M-PESA. At the national level, the agent network grew from about 4,000 agents at the time of round one of the survey to close to 15,000 by round three. Between 2008 and 2010, there was therefore a four- to fivefold increase in the number of agents across the country, a period over which bank branches grew by 20 percent (from 887 to 1,063).

F. Related Research

Table 3 below includes information on papers and briefs published using these datasets.

Table 3: Publications using this Data

Papers			
Title	**Authors**	**Date of Publication**	**Rounds Used**
The Economics of M-PESA	Tavneet Suri William Jack	2009	1,2
Documenting the birth of a financial economy	Tavneet Suri William Jack Thomas M. Stoker	2011	1,2
Transaction Networks: Evidence from Mobile Money in Kenya	William Jack Adam Ray Tavneet Suri	2013	1,2,3
Risk Sharing and Transactions Costs: Evidence from Kenya's Mobile Money Revolution	Tavneet Suri William Jack	2014	1, 2,3

Briefs			
Title	**Authors**	**Date of Publication**	**Rounds Used**
The M-PESA Effect: Are Financial Transaction Costs a Barrier to More Effective Insurance for Families in Kenya?	FSD Kenya	2015	1 to 5
An Overview of M-PESA	FSD Kenya Tamara Cook	2015	1 to 5
The Mobile Money Revolution in Kenya: Can the Promise be Fulfilled?	Tavneet Suri	2015	1 to 5