Understanding Key Mobile Money Users

A Public Report

February 2018

BY

Carolina Mattsson, PhD Student, Network Science, Northeastern University

Guy Stuart, Ph.D., Executive Director, Microfinance Opportunities

CONTACT

guystuart@mfopps.org

This work would not have been possible without the generous support of the Gates Foundation, and the collaboration of the IFC and Airtel, Uganda. We are grateful for thoughtful comments by Soren Heitmann and Christian Ruckteschler at the IFC. All Airtel data used in this analysis was cleaned and structured by Cignifi, Inc.

INTRODUCTION

Mobile money platforms are networked systems, and understanding *key mobile money users* is essential for mobile money providers, policy makers, and promoters of financial inclusion who seek to impact their future development. Are there especially active mobile money users with an outsize effect on the system? If so, who are these *key mobile money users*, what drives them to use the service extensively, and what is the nature of their influence?

Mobile money platforms are networked systems, and two features of networked systems conspire to set them apart. First, networked systems often follow the Pareto principle (colloquially, the 80/20 rule) where a small group of users is responsible for a majority of the activity happening on the system.[1] Second, adding more users makes a networked system more valuable to other users and potential users.[2] For instance, a merchant willing to accept payments in mobile money creates an additional opportunity to use the service and increases the value of the network as a whole. These two features interact strongly. By using a networked system extensively, the small group of very active users exerts substantial influence over the character of that system as a whole.[3] The members of this small active group help define how the system can be used and how valuable the system becomes for everyone else. We will refer to members of this small group as *key users*.

In this report we begin by identifying *key users* from the individual-level transaction data of a mobile money provider in one East African country. We then identify a roughly analogous set of self-reported mobile money users from a cross-country panel survey. The survey data and the billing records of the provider offer two different lenses through which to understand *key users*.

Through the lens of the survey data, the diversity of *key users* is most striking. For example, although a *key user* is much more likely to live in an urban area than another user, 45% of all *key users* live in rural areas. To the extent that providers can identify and support *key users* in underserved populations, we see an opportunity for growth. We also find that distinguishing *key users* from *users* is substantially different from distinguishing *users* from *non-users*. Factors like non-farm employment and bank account ownership strongly encourage *key users* even when the

[1] The equivalent in the network literature are "hubs", see: Jackson, M. O. (2010). An overview of social networks and economic applications. The Handbook of Social Economics, 1, pp, 14-16.
[2] Katz, Michael L., and Carl Shapiro. "Network Externalities, Competition, and Compatibility." *The American Economic Review*, vol. 75, no. 3, 1985, pp. 424–40.
[3] Szabo, G., and A. L. Barabasi. "Network Effects in Service Usage." *arXiv:physics/0611177*, Nov. 2006. *arXiv.org*, http://arxiv.org/abs/physics/0611177.

related spatial and demographic characteristics are taken into account. Less encouraging is that network considerations like local adoption rates and access to agents fall in importance.

Through the lens of the provider data, we see that *key mobile money users* are a relatively stable group that underlies a substantial fraction of all transactions. User features from billing records, like social network size and mobile calling habits, strongly distinguish *key users* even when we take demographic inferences into account. This raises the possibility of targeting potential *key users* from underserved populations by emphasizing the behavioral signal in provider data.

The provider data confirms that *network* variables are more subtle in the context of *key users*. Network effects primarily accentuate underlying demographic and behavioral trends, and mobile money users create opportunities for others not by having an account, but by using it.

DATA SOURCES ON MOBILE MONEY USERS

The data used in this report comes from two very different sources: the transaction and billing records of a mobile money provider, and a geocoded survey on financial inclusion. These two datasets provide complementary lenses through which to understand *key mobile money users*, and we draw on their respective strengths throughout this report.

Our first dataset consists of fine-grained behavioral data from the transaction and mobile billing records of Airtel Uganda for the six-month period November 1, 2014 to April 30, 2015. Our analysis draws on a one percent random sample of Airtel Money users—about 25,000 customers in all. This is fine-grained behavioral data, which lets us identify *key mobile money users* based on their transaction history, and characterize them based on their mobile calling behavior. We can also follow users over the six-month period, and directly observe the behavior of users' mobile phone contacts to gain insight into network effects.[4]

Our second dataset consists of detailed survey data from Waves 2 and 3 of the Financial Inclusion Insights Survey for Uganda, Kenya, and Tanzania, with which we incorporated additional geospatial data. These surveys cover about 18,000 respondents in total. While detailed transaction histories are not available, respondents were asked to self-report information on their mobile money use. This lets us identify a roughly analogous set of *key mobile money users* and complement our understanding of this group with rich socio-demographic information. We can also directly measure

[4] See Appendix I for details.

the impact of specific access constraints and look beyond a single network to draw conclusions about mobile money users in East Africa more generally.[5]

[5] See Appendix II for details.

ARE THERE KEY MOBILE MONEY USERS?

There is indeed a small group of especially active mobile money users – *key mobile money users*.

We identify *key users* in the mobile money transaction records of Airtel Uganda using algorithmic segmentation. We base the segmentation on the number of times users cash-in, cash-out, send a person-to-person (P2P) transfer and receive a P2P transfer over a sequence of three two-month periods.[6] Algorithmic segmentation uncovers a small group of *key users* in each period who represent about 8% of our sample of Airtel Money users and are much more active than other users in every way. These *key mobile money users* use the service more often, conduct more transactions of every kind, and maintain larger balances. This exceptional activity gives *key users* a substantial presence in the system as a whole: *key users* make 47% of all cash-ins.

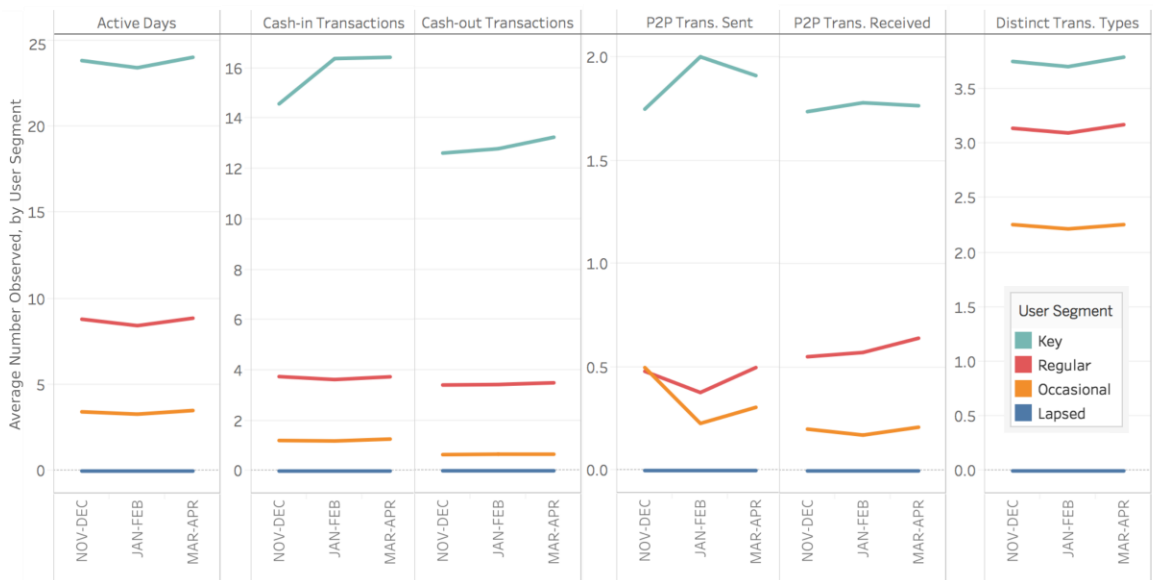


Figure 1: The group means of *key mobile money users* in comparison to three additional customer segments that emerged from algorithmic clustering. *Key users* are ~8% of users, while the other segments each represent about 30%. *Key users* are more active than other users in every way.

We identify roughly analogous *key users* in the FII survey data using the available self-reported information on mobile money use.[7] The survey asked about recent use and about specific use cases, framed as: "Have you ever used a mobile money account to do the following?" We consider a

[6] Note that we only observe as P2P those person-to-person transfers sent officially. Unofficial workarounds evading fees were prevalent among users at the time of data collection, creating some data artifacts. See Appendix II details.

[7] While we intend to capture similar users in both datasets, inherent data limitations mean that the segmentations are not entirely equivalent. This would require directly linking survey respondents to their behavioral data, via their phone number, and is a promising avenue for future research that would require strict anonymity provisions.

respondent to be a *key user* if they used mobile money in the past week and report two or more "intentional" use cases such as transferring money to another person, paying a utility bill, saving, or conducting business activities.[8] This results in small groups of *key users* who represent 13%, 14%, and 27% of mobile money users in Uganda, Tanzania, and Kenya, respectively. Like those identified in the Airtel Uganda data, these *key users* use mobile money services more heavily and more often than others. In Table 1 we include, for comparison, the means for two additional customer segments with one or no "intentional" uses. Any respondent who reports no use in the past 90 days is placed in the passive segment.

Table 1: Key Users

Country	Segment	Share of Respondents	Share of Users	Average Total Uses	Monthly Top-ups (est)	Active Yesterday	Report A Business Use	Registered MM Users
Kenya	Key User	21.5%	27.3%	7.3	5.38	38.9%	35.5%	94.8%
	Intentional User	31.9%	40.5%	5.0	2.80	11.6%	10.8%	89.4%
	Passive User	25.4%	32.2%	2.6	1.21	9.9%	0.9%	78.5%
	Non User	21.1%	0.0%	0.0	0.00	0.0%	0.0%	0.0%
Tanzania	Key User	7.4%	13.9%	6.6	5.16	38.1%	30.9%	95.3%
	Intentional User	22.2%	41.9%	4.6	3.78	10.1%	7.5%	92.8%
	Passive User	23.4%	44.2%	2.4	1.77	12.0%	1.0%	84.0%
	Non User	47.0%	0.0%	0.0	0.00	0.0%	0.0%	0.0%
Uganda	Key User	5.7%	13.1%	7.0	4.00	33.3%	42.3%	97.1%
	Intentional User	18.8%	42.9%	4.3	2.24	10.1%	11.2%	82.1%
	Passive User	19.4%	44.0%	2.3	1.09	7.6%	1.5%	61.9%
	Non User	56.0%	0.0%	0.0	0.00	0.0%	0.0%	0.0%

Table 1: The relative sizes and group means of *key mobile money users* as defined for the FII Survey data. Two additional customer segments and non-users are included for comparison. This segmentation is intended to define *key users* who are roughly analogous to those identified in the Airtel Uganda transaction data, see Appendix II for details.

[8] "Intentional" uses are: sending transfers to other persons, paying utility bills, school fees, or medical fees, making purchases, saving, investing, making or taking loans, receiving wages, and conducting business activities. Excluded from our count are particularly common and/or passive uses: depositing or withdrawing cash, purchasing airtime, receiving a transfer from another person, and receiving a payment from the government. See Appendix II for details.

ARE KEY USERS A STABLE GROUP?

While the group of *key users* in the Airtel Uganda data looks similar over the six-month period, there is considerable movement of individuals in and out of this group. Around 44% of the *key users* as of November and December were still in the top category in March and April. Still, over 80% of these *key users* remained in the top two categories and only 10% performed no transactions at all. There is more stability in use from month to month among key users than among other users.

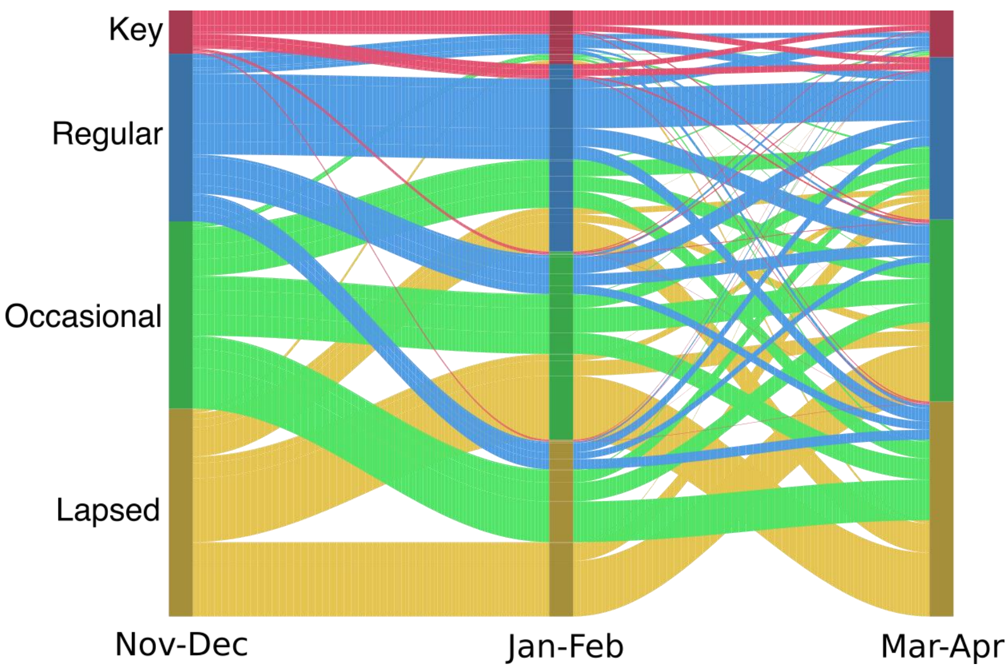


Figure 2. This plot tracks individuals, highlighting movement between segments over time. We see more stability in use from month to month among key users and regular users.

WHO ARE KEY MOBILE MONEY USERS?

The FII data provides a wealth of information on the characteristics of different types of mobile money users. We have used these data to help us understand who *key users* of mobile money are. In absolute terms, the diversity of *key mobile money users* is most striking. *Key users* reflect the breadth of individual characteristics in East Africa, although there are relatively fewer *key users* from disadvantaged populations. *Key mobile money users* are more educated, more urban, less likely to be poor, more likely to be male, less likely to be teenagers, more likely to be employed in non-farm occupations, and more financially sophisticated than other users. *Key users* overwhelmingly own mobile phones.[9] Nevertheless, there are many *key users* who are women, low-

[9] See Appendix IV for variable tables and complete mixed effect logistic regression results.

income, less-educated, rural, employed in farming, or otherwise financially excluded. These key users from underserved demographic groups could be especially interesting to providers.

As an illustration of the interplay between absolute and relative trends, Figure 3 conveys the number of survey respondents in each category by the size of its tile. The orange tiles correspond to rural respondents, while the blue tiles correspond to urban respondents. Note that the orange tiles are in sum larger than the blue, reflecting the fact that there are more rural than urban respondents in total. The lightest tiles correspond to non-users of mobile money, while darker tiles correspond to more active users. We can see that a larger *proportion* of the urban users are *key users*, yet in total there are about as many *key users* in rural areas as there are in urban areas, as indicated by the comparable sizes of the dark-orange and dark-blue tiles.

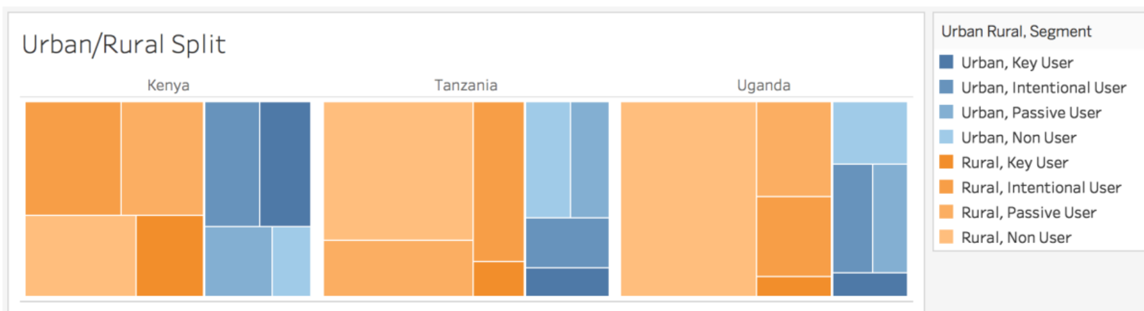


Figure 3. In this illustration the size of the tiles reflect the number of respondents. The orange tiles represent rural respondents and the blue tiles urban respondents. Darker colors represent more active users. Note the interplay between relative trends and absolute numbers that results in the dark-orange and dark-blue tiles being of comparable size.

Within the FII data, other user characteristics reinforce this picture:

- *Key users* are more educated than *users* who are more educated than *non-users*. However, fully 58%, 61%, and 45% of *key users* in Uganda, Tanzania, and Kenya, respectively, had not completed secondary school.
- *Key users* are better off than *users* who are better off than *non-users*. However, an estimated 40%, 72%, and 32% of *key users* in Uganda, Tanzania, and Kenya, respectively, fall below the $2.50/day 2005 PPP poverty line.[10]
- Farmers and persons who are not employed are underrepresented among *key users*. Laborers, professionals, and those who run a business (including self-employed) make up

[10] Specifically, we use the Poverty Probability Index that a user falls below $2.50/day 2005 PPP. This is included in the FII Survey data, and was calculated using the standard PPI survey questions. $1.25/day 2005 PPP was the World Bank standard of extreme poverty at the time the surveys were conducted.

63% of *key users*, 39% of other *users*, and 22% of *non-users*. Nevertheless, in all countries there are as many or more *key users* who are farmers as there are who are professionals. See Appendix IV for a full table.

- Having a bank account, which we interpret as an indicator of financial sophistication, raises the odds of being a key user by 75% or more in all three countries. However, having a bank account is still rather uncommon in both Uganda and Tanzania, so in total there are more key users who do not own bank accounts than who do in those countries.
- We employ phone access/ownership and distance to a mobile money agent[11] as quantitative measures of access to mobile money. We find that access constraints are dramatically important in distinguishing *non-users* from *users*, and continue to have a smaller impact on whether a *user* is a *key user*. *Key users* overwhelmingly own mobile phones.

The Airtel Uganda data allows us to infer a handful of characteristics for Airtel Money users, and we observe a similar pattern. Again, the diversity of *key users* is most striking:

- Airtel Money *users* from Kampala[12] are nearly twice as likely to be *key users*. However, most of the *users* in our sample live outside Kampala. With respect to absolute numbers, about a third of *key Airtel Money users* are in Kampala.
- Using a simple proxy for income[13], we infer that *key Airtel Money users* are better off than other *users* on average. However, 26% of *key Airtel Money users* fall lower on this measure than the median *lapsed* or *occasional user*.

The interplay between the absolute diversity and relative advantage of *key mobile money users* means that there are *key users* to be found in all but the most disadvantaged populations. The existence of these users hints at opportunities for growth in underserved populations to the extent that providers can identify and support their use.

[11] This is calculated using geodesic distance to the nearest mobile money agent catalogued in a separate dataset of financial service providers in these countries, published by FSP Maps. See Appendix II for details.
[12] We use the geographic location of a user's most-called cell tower to infer home location. See Appendix I for details.
[13] We use the average size of cash-in transactions over the full six-month period, in UGX, as a simple proxy for income. We expect this to be a noisy inference that may nonetheless be useful.

WHAT SETS KEY USERS APART?

Although the diversity of *key users* is most striking, overarching trends among *key users* contain actionable information as well. In particular, since many of the characteristics described above go hand-in-hand, we want to understand whether a characteristic is still important when all the other characteristics are taken into account – its *independent* effect. We use mixed-effects logistic regression to compare characteristics in this way, and provide full tables of regression results in Appendix IV for the interested reader. Here we summarize, combining specific variables together into four broad categories that suggest different avenues for action:

- *Fixed variables*[14] are demographic and spatial characteristics of users like education, poverty, and the population density of the area in which they live. These variables capture societal constraints affecting adoption and use of mobile money.
- *Behavioral variables*[15] are characteristics of users that directly affect their day-to-day financial activity, such as occupation and bank account ownership. Together, these variables get at a customers' individual use case for mobile money.
- *Network variables*[16] are characteristics of the mobile money rollout in a user's local area. These variables capture the importance of ease of access and network effects.
- *Personal access to or ownership of a mobile phone* is included separately, since it has a direct causal effect on a person's individual ability to access mobile money services.

We build composite indices for each of these categories so that we can compare their relative impact on distinguishing *users* from *non-users* and *key users* from *users*. In Figure 4, we plot the relative impact of these indices for each of the three countries in the FII survey data.

Having *personal access* to mobile money via access to and ownership of a mobile phone is by far the most distinguishing factor between *users* and *non-users*. *Fixed, behavioral,* and *network* variables all play a supporting role in mobile money adoption. Distinguishing *key users* from *users* appears to be substantially different – *fixed* and *behavioral* variables have the strongest independent effects. Personal access is less distinguishing as most of those without mobile phones are not *users* to begin with, and *network* variables appear to have a secondary effect.

[14] Includes: age, under-20 binary variable, gender, PPI probability, education, and population density
[15] Includes: occupation, marital status, bank account ownership, and survey reported loan behavior
[16] Includes: distance to nearest mobile money agent and the local penetration rate of mobile money

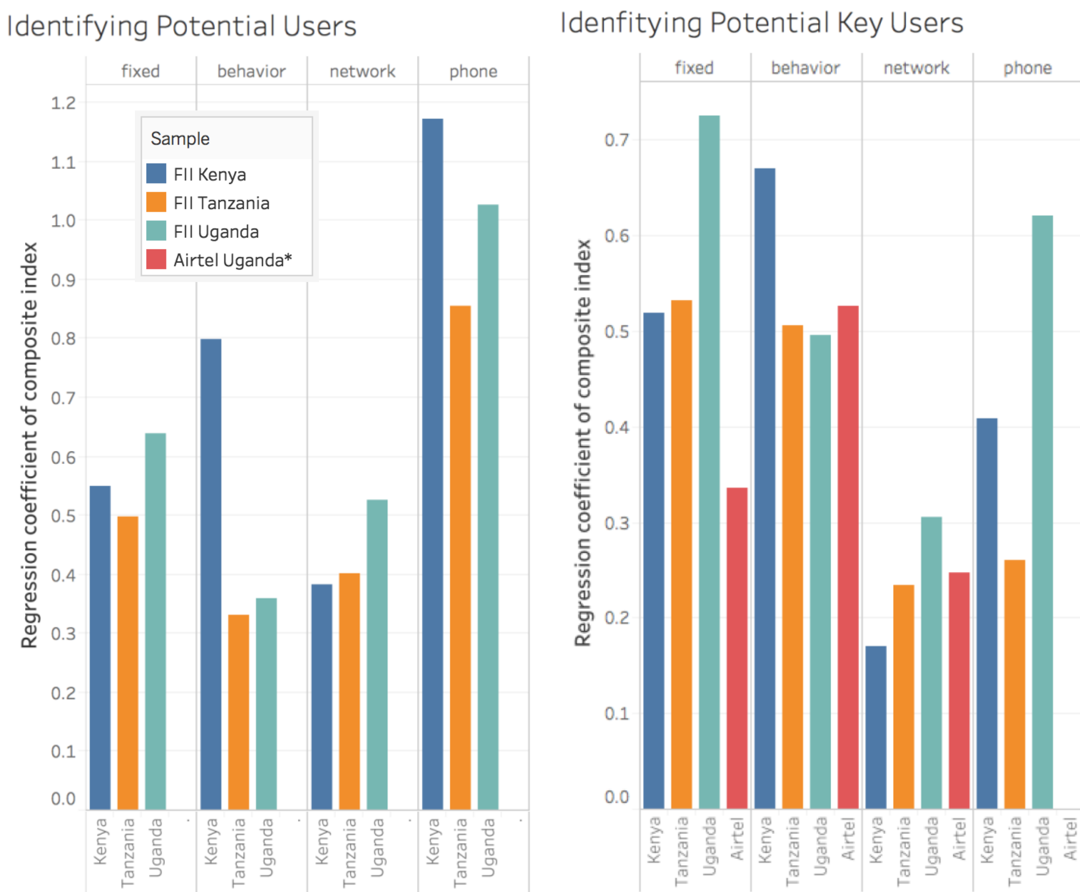


Figure 4: Plotted is the relative importance of composite indices in distinguishing *non-users* from *users* and *users* from *key users*. The variables underlying the composite indices for Airtel Uganda are substantially different, and discussed in the next section. See Appendix III and IV for details.

The increased importance of *behavioral* variables in distinguishing *key users* from those who are already *users* is especially interesting and actionable. Factors like non-farm employment and bank account ownership strongly encourage *key users* even when the related spatial and demographic characteristics are taken into account. This raises the possibility of targeting potential *key users* from underrepresented demographic groups by emphasizing *behavior*.

The decreased importance of *network* variables in distinguishing *key users* from those who are already *users* poses a potentially serious problem for mobile money providers and proponents. Networked systems rely on network effects – that adding more users makes the system more valuable to other users – to flourish, but it seems that higher adoption rates are not encouraging more active use on their own. The Airtel Uganda data allow more detailed analysis of this question, which we explore in our last section.

WHAT SETS KEY USERS APART IN PROVIDER DATA?

Transaction histories and mobile billing records provide a second lens through which to view overarching trends among *key Airtel Money users*. Airtel Uganda billing records contain fine-grained information on customers' mobile phone use. These records capture whom users talked to, how often, when, and approximately where. These records provide insights into mobile money users and their social networks, and can be used to understand what sets *key users* apart.

The variables with the strongest *independent* effect on *key users* are: the size of a customers' social network[17], their average number of calls per contact, and a simple proxy for income[18]. These variables have a substantial effect on the likelihood that a *user* is a *key user*, although they are not precise since the overall population of *key users* is small. Having 50 additional voice contacts raises the odds of being a *key user* by 52%, an additional 2.24 average calls per contact raises the odds by 45%, and the equivalent[19] increase in the income proxy raises the odds by 19%. To a lesser extent, indicators capturing the population density of users' inferred home locations, calling habits, geographic diversity of contacts, the fraction of users' social networks who have mobile money, and the mobile money activity of users' social networks are also significant.[20]

Although the variables themselves are not directly comparable to those in the FII data, they reflect underlying *fixed*, *behavioral*, and *network* characteristics that we can relate back to on a general level. As in the FII survey analysis, *fixed* variables like inferred income and *behavioral* variables like average calls per contact are independently important. *Network* variables detailing the average characteristics of a user's social network have a secondary effect (see Figure 4).

It is important to note that many observed variables, like social network size, will pick up on both *behavioral* and *fixed* differences between users. We use inferred income and population density in an effort to take this into account, but we expect that our *fixed* composite index is weaker than it could be. If more detailed demographic information is available, such as through surveys or Know Your Customer data, providers can better control for *fixed* differences and better isolate *behavioral* differences.[21] This raises the possibility of targeting potential *key users* from underrepresented demographic groups by emphasizing the *behavior* seen in provider data.

[17] This is the number of unique voice contacts over the two-month period.
[18] As a rough indicator of income, we calculate the average size of all cash-in transactions made by an Airtel Money user over the full six-month period. We use the log value of this in our analysis.
[19] These are all continuous measures. The interval presented is the observed difference in averages between super users and lapsed or occasional users.
[20] See Appendix III for details, including a list of all variables included and full result tables.
[21] The relevant question becomes: is this user more likely to be a potential key user *than similarly well-off users*?

HOW IMPORTANT ARE NETWORK EFFECTS?

Our initial findings suggest that users' networks have a lesser impact on their mobile money use than expected: a higher local adoption rate and easier access to agents have a subdued influence on whether *users* are *key users*. However, the Airtel data allows us to take a more granular look, which reveals two nuanced yet important network effects.

First, we see direct evidence of network effects at the individual level: having voice contacts who use mobile money more often is correlated with being a *key user* oneself. This effect is stronger and more consistent than having a higher local adoption rate (a larger fraction of voice contacts who have adopted mobile money). In other words, when individual *fixed* and *behavioral* effects are taken into account, it is the level of mobile money activity that a user's social network maintains that has the greatest correlation with being a *key user*. This suggests that mobile money users create opportunities for others to use the service not by having an account, but by using it.

Second, network effects accentuate underlying trends. To illustrate, we use an example. Figure 5 shows in orange the average difference in social network size across user categories. *Key Airtel Money users* have roughly twice as many voice contacts as lapsed and occasional users; their social networks are larger. In addition to being larger, the social networks of *key users* have higher mobile money adoption rates and make more cash-ins on average. Figure 5 shows in blue the average total number of cash-ins made by the social networks of users in those categories (not including themselves). *Key Airtel Money users* have roughly twice the social networks of lapsed and occasional users, but these social networks make three times as many cash-ins in total. Network effects are accentuating the underlying trend in social network size.

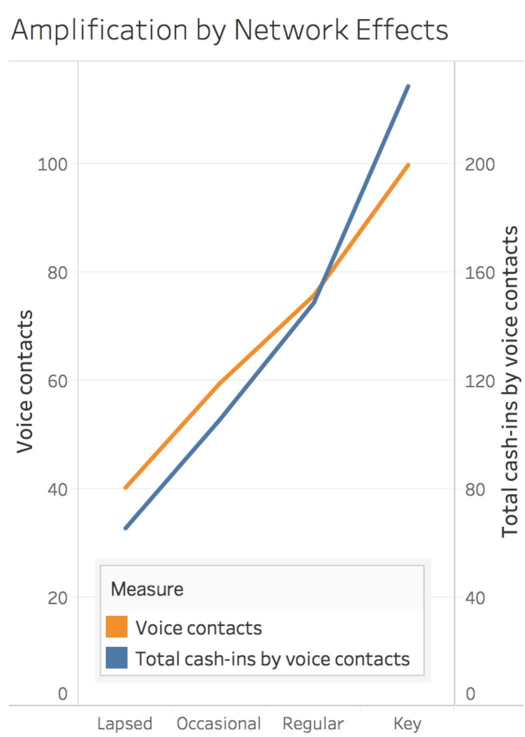


Figure 5. Small but consistent trends in the mobile money activity levels of *users' social networks* magnify personal differences in network size.

CONCLUSIONS

In this report we identify and characterize a subset of users, in two complementary datasets, whose use of mobile money services is particularly extensive. We find that *key mobile money users* underlie a substantial fraction of all transactions and are a relatively stable group. These *key users* help determine the value of the system itself and are of particular interest.

Although *key users* are relatively more educated, urban, male, employed, financially sophisticated and less likely to be poor, teenagers, or farmers, the population of *key users* reflects the diversity of East Africa. For example, though a *key user* is more likely to live in an urban area than a regular user, 45% of all *key users* live in rural areas. More generally, the data suggest that there are *key users* in all but the most disadvantaged populations. To the extent that providers can identify and support *key users* in underserved populations, we see an opportunity for growth into these populations. Furthermore, given the large number of people in these underserved populations, it is here where the most growth can occur.

Distinguishing *key users* from *users* appears to be substantially different from distinguishing *users* from *non-users*. Especially actionable is that *behavioral* variables are of increased importance in distinguishing *key users* from those who are already *users*. Factors like non-farm employment and bank account ownership strongly encourage *key users* even when the related spatial and demographic characteristics are taken into account. This raises the possibility of targeting potential *key users* from underrepresented demographic groups by emphasizing behavior. Our analysis suggests that billing records of mobile money providers can be used for this purpose – social network size and mobile calling habits are strong distinguishing factors of *key users* even when we take demographic inferences into account. Providers could better isolate *behavioral* differences where more detailed demographic information is available.

Network effects are important, but secondary, in the context of *key users*. Compared to mobile money adoption, factors like the local adoption rate and ease of access to agents are subdued. Primarily, network effects accentuate underlying *fixed* and *behavioral* trends. When these individual factors are taken into account, it is the mobile money activity level that a user's contacts maintain that has the greatest correlation with being a *key user*. Mobile money users create opportunities for others to use the service not by having an account, but by using it.

RECOMMENDATIONS

For Mobile Money Providers

Providers should focus on (1) fostering and (2) expanding *key users* from across their user base.

First, identify and find ways to support existing *key mobile money users,* especially those from otherwise underserved populations. *Key users* are a double benefit – they answer for a large fraction of existing activity themselves and encourage other users to be more active.

Second, identify and target potential *key users* by behavior, rather than demographics, to reach a broader pool. Mobile money providers can identify potential *key users* from behavioral signatures in their own billing records. Providers can better isolate behavioral differences by controlling for demographic information where it is available or can be inferred.

For Promoters of Financial Inclusion

First, encourage mobile money providers to pursue the recommendations listed above.

Second, while demographic characteristics remain resoundingly important – and this is an issue – absolute numbers show space for optimism. There are substantial numbers of more rural, poorer, and less educated people using mobile money services already, and thus a clear opportunity upon which to build. Proponents ought to study the experience of statistically less likely users and encourage providers to cater to them and solve the problems they are facing.

Lastly, we must continue to emphasize access and adoption while acknowledging the limits of this approach. Network effects are helping mobile money spread, but higher adoption rates are not enough to encourage more extensive use of mobile money services. There needs to be activity, and activity breeds activity. Especially in terms of encouraging use of mobile money among underserved populations, it may be prudent to move towards promoting this use directly. Promoters can encourage the cultivation of communities of key users in underserved populations and begin to exploit network externalities more effectively.

APPENDIX I

AIRTEL: DATA AND CUSTOMER SEGMENTATION

Through the Gates Foundation and IFC, we received cleaned and anonymized Airtel Uganda data files for each of the six months between November 2014 and April 2015. We use the files containing Airtel Money transactions, Airtel Uganda voice calls, and a mapping from cellular tower IDs to geographic coordinates that can be cross-referenced with the voice files. Individual phone numbers are hashed in the same way across files, meaning that we can compare the social and financial networks of the anonymized individual accounts in the data.

Data irregularities

Preliminary data checks identified several irregularities that must be kept in mind. Especially vexing is an apparent under-reporting of P2P transactions due to a practice we term over-the-counter person-to-person transfers or OTC-P2P transfers. These are transfers from one person to another whereby the sender deposits money directly onto the phone of the intended recipient, rather than depositing the money on their own phone and then transferring the money to the recipient. An analysis of the one-percent sample found that 30 percent of the amount cashed in on a phone was cashed out within four hours from the same phone, and that 45 percent of the amount cashed in was cashed out within 24 hours. These data are suggestive of the scale of the OTC-P2P transfer phenomenon. This phenomenon also limits the data we have on the network context of Airtel mobile money users, because these types of transaction only contain a record of the recipient of the transfer, not the originator. As a result, we have relied on voice call data for the bulk of our network analysis.

We noticed several issues with the geographic data we received. First, the recipient of a call is rarely assigned a cell tower and so we receive geographically keyed information about mobile money users mainly when they initiate a call. Furthermore, some cell tower IDs are missing geographic coordinates, or missing data entirely. In particular, these data are missing cell tower IDs (and their corresponding locations) for several regions of the country beginning sometime in February. This adds sizeable uncertainty to our geographic embedding.

Data processing

We reformatted the mobile money transaction data from an unwieldy list of transactions into chronological histories indexed by individual account. First, we combined the files for the 6 months and added the inverse of all transactions so they are acknowledged as transactions for both the sender and receiver. Topup transactions were not inverted. We then brought together all transactions involving one individual together chronologically. For each transaction we store the unique ID of the individual, the unique ID of their counterparty in the transaction, the transaction 'type', whether the focal ID was the 'source' or the 'target', the timestamp, and the amount.

Using a similar process, we created meaningful aggregations of voice variables for each mobile money user. We noted a user's total calls, outward calls, daytime calls, call duration, and number of days communicated with each of their voice alters in each month. In order to do geographically embedded analysis, we record the number of hits to each cell tower during daylight hours (6am-6pm), night hours (6pm-6am), and overall.

As mentioned above, we received a file linking cellular tower IDs with geographic locations. We used QGIS and the administrative areas from GDAL to place each tower location within a level 2 geographic area, corresponding to Ugandan

Counties. Within Kampala we used the City Council boundaries from the third GDAL level, as there was no difference between the first and second levels. The handful of cell towers who were placed just outside the country or in water bodies were manually assigned to the closest administrative area. The handful of cell towers where we could not impute the geographic information were given a location ID of 998 and cell towers with no additional information provided were given a location ID of 999.

Variable extraction

1. Mobile money activity

For all mobile money accounts we calculated: the number of financial contacts, number of active days, number of transaction types, mean of account balance, total number of each transaction type, total amount transacted for each transaction type, and average size of transactions for each transaction type. These values are calculated for each of the three two-month periods and for the full six-month period.

2. Voice call activity

For all mobile money accounts we calculated: the number of voice contacts, number of calls, average calls per contact, average call duration, number of outgoing calls, number of daytime calls, and a user's 'home location'. A 'home tower' is the tower with the most calls routed through it, and a 'home location' is the administrative area that contains that 'home tower'. These values are calculated for each of the three two-month periods and for the full six-month period.

3. Social network averages

For all mobile money accounts we calculated measures of mobile money activity across their voice contacts. We first calculate the fraction of voice contacts who were Airtel Money customers.[22] These are, in a direct sense, potential transaction partners with which the user already has a social connection. We then calculate *averages* over these voice contacts who were Airtel Money customers: the number of voice contacts, number of active days, mean of account balance, total number of cash-in transactions, total amount cashed in.

An interesting observation

Most customers have many social ties who are active mobile money customers but very few financial transaction partners. The customer-to-customer financial network is very thin. This contrast is intriguing, and we suggest three potential explanations:

- These social contacts have no reason to send each-other money
- These social contacts send money to one another in some non-digital way
- These social contacts do send money to one another, but do so using informal OTC

Creating a 1% sample

We simply looped through the accounts and selected customers[23] with a probability of 0.01.

[22] Specifically, the fraction of their voice contacts who appear in the Airtel Money data. This means they used mobile money at any point in the six-month observation period.
[23] We used the recorded transaction types to distinguish between customers and other accounts, including agents. An account is defined to be an agent account if they, at any point in the 6 months, initiate a CASHIN or receive a CASHOUT.

Customer segmentation

We used Tableau's built-in clustering algorithm to perform a k-means clustering of the 1% sample of the Airtel mobile money customers. As our dimensions we used the count of transactions for cash in, cash out, P2P transfer sent, and P2P transfer received. We chose to focus on counts of transactions (rather than amounts) to avoid conflating mobile money behavior with characteristics related to the resources available to the customers.

Initially we focused on the November-December period, with the intent of building a baseline profile of the sample and following the clusters through time. Tableau offers the option of letting the algorithm determine the optimal number of clusters based on the Calinski-Harabasz criterion and this was the starting point for our analysis. The initial clustering identified two clusters, one of which was about 8 percent of the sample – these are our *key mobile money users*. In order to compare these users to other groups of users, we conducted a separate clustering analysis of the larger group. The second clustering identified three further clusters, representing 11 percent, 16 percent, and 65 percent of the total. A closer look at the data identified that about half of the larger group (the 65 percent) performed no transactions in November and December—they were in the data set because they performed transactions in later months. We use these findings to divide the 92 percent group into roughly thirds.

We divide the 1% sample into four basic segments that emerged from this analysis:

- Key Users —the top ~8% of the sample, corresponding to the group that was identified as being distinct from the vast majority of users in the first round of Tableau's k-means clustering
- Regular Users —the upper middle ~27% of the sample, this is a combination of the two smaller groups emerging out of the second round of Tableau's k-means clustering
- Occasional Users —the lower middle ~30% of the sample, these are the customers in the large group resulting from the second round of clustering that performed at least one transaction
- Lapsed Users —those that performed no transactions in November and December

We again used Tableau's k-means clustering to find analogous segments in the January-February and March-April periods. We used the exact same definitions of segments, specifying two groups in the first round of clustering and three in the second, to examine the extent to which segments remained consistent over time and whether individual mobile money users moved between segments over time.

Appendix I, Table 1: Behavioral Segments, Airtel Money Data

A non-agent account is defined to be a customer account if, at any point in the 6 months, they are on the expected side of a CASHIN, CASHOUT, TOPUP, BILLPAY, BULKPAY, or P2P transactions.

Current Time Period Segment Means

User Segment	Time Period	Segment Size	Active Days	# Txn Types	Mean Balance	Cash-in Txns	Cash-in Amt/Txn	Cash-out Txns	Cash-out Amt/Txn	P2P Txns Received	P2P Txns Sent	Topup Txns	Billpay Txns
Key	NOV-DEC	1,806	23.82	3.75	67,267	14.58	54,338	12.62	50,604	1.74	1.75	16.23	1.66
	JAN-FEB	1,662	23.41	3.70	73,510	16.38	59,402	12.79	53,490	1.78	2.00	14.34	1.07
	MAR-APR	2,063	24.02	3.79	68,726	16.43	53,974	13.25	47,446	1.77	1.91	15.58	0.75
Regular	NOV-DEC	6,266	8.81	3.14	40,100	3.75	50,297	3.41	49,080	0.55	0.48	5.42	0.43
	JAN-FEB	6,473	8.44	3.10	38,904	3.63	49,830	3.43	47,364	0.57	0.38	4.95	0.31
	MAR-APR	7,239	8.87	3.17	38,098	3.74	48,183	3.50	46,818	0.64	0.50	5.15	0.19
Occasional	NOV-DEC	7,035	3.44	2.26	37,088	1.21	38,904	0.64	43,190	0.20	0.50	2.30	0.25
	JAN-FEB	7,271	3.30	2.22	38,579	1.19	35,915	0.66	43,345	0.17	0.23	2.27	0.20
	MAR-APR	7,292	3.51	2.26	38,692	1.27	38,907	0.66	43,777	0.21	0.31	2.40	0.14
Lapsed	NOV-DEC	8,267	0.00	0.00		0.00		0.00		0.00	0.00	0.00	0.00
	JAN-FEB	7,968	0.00	0.00		0.00		0.00		0.00	0.00	0.00	0.00
	MAR-APR	6,780	0.00	0.00		0.00		0.00		0.00	0.00	0.00	0.00

November-December Segment Means

Nov-Dec Segments	Time Period	Segment Size	Active Days	# Txn Types	Mean Balance	Cash-in Txns	Cash-in Amt/Txn	Cash-out Txns	Cash-out Amt/Txn	P2P Txns Received	P2P Txns Sent	Topup Txns	Billpay Txns
Key	NOV-DEC	1,806	23.82	3.75	67,267	14.58	54,338	12.62	50,604	1.74	1.75	16.23	1.66
	JAN-FEB	1,806	17.33	3.37	53,746	11.51	53,521	8.65	51,887	1.38	1.35	11.65	1.04
	MAR-APR	1,806	17.19	3.29	53,950	11.13	50,676	8.48	49,399	1.33	1.55	12.02	0.83
Regular	NOV-DEC	6,266	8.81	3.14	40,100	3.75	50,297	3.41	49,080	0.55	0.48	5.42	0.43
	JAN-FEB	6,266	6.98	2.40	38,246	3.18	50,156	2.78	50,084	0.45	0.39	4.32	0.29
	MAR-APR	6,266	7.37	2.35	41,043	3.53	49,574	2.95	50,630	0.52	0.50	4.57	0.18
Occasional	NOV-DEC	7,035	3.44	2.26	37,088	1.21	38,904	0.64	43,190	0.20	0.50	2.30	0.25
	JAN-FEB	7,035	3.69	1.49	67,126	1.40	50,639	1.09	53,645	0.21	0.28	2.45	0.18
	MAR-APR	7,035	4.11	1.55	65,947	1.63	52,678	1.26	54,251	0.27	0.34	2.66	0.12
Lapsed	NOV-DEC	8,267	0.00	0.00		0.00		0.00		0.00	0.00	0.00	0.00
	JAN-FEB	8,267	2.00	1.30	21,510	1.07	32,608	0.91	34,226	0.14	0.07	0.84	0.03
	MAR-APR	8,267	4.03	1.89	25,273	2.01	36,550	1.78	35,838	0.28	0.18	2.16	0.06

January-February Segment Means

Jan-Feb Segments	Time Period	Segment Size	Active Days	# Txn Types	Mean Balance	Cash-in Txns	Cash-in Amt/Txn	Cash-out Txns	Cash-out Amt/Txn	P2P Txns Received	P2P Txns Sent	Topup Txns	Billpay Txns
Key	NOV-DEC	1,662	18.11	3.28	71,132	12.31	56,426	8.98	54,545	1.21	1.70	11.82	1.32
	JAN-FEB	1,662	23.41	3.70	73,510	16.38	59,402	12.79	53,490	1.78	2.00	14.34	1.07
	MAR-APR	1,662	19.73	3.51	65,516	14.49	55,273	10.05	49,772	1.51	1.98	13.19	0.84
Regular	NOV-DEC	6,473	7.55	2.40	42,691	3.29	47,025	2.96	47,217	0.50	0.44	4.81	0.45
	JAN-FEB	6,473	8.44	3.10	38,904	3.63	49,830	3.43	47,364	0.57	0.38	4.95	0.31
	MAR-APR	6,473	8.22	2.59	38,458	3.53	47,590	3.33	48,084	0.57	0.49	5.15	0.17
Occasional	NOV-DEC	7,271	3.98	1.58	52,109	1.40	49,094	1.14	51,851	0.24	0.48	2.88	0.27
	JAN-FEB	7,271	3.30	2.22	38,579	1.19	35,915	0.66	43,345	0.17	0.23	2.27	0.20
	MAR-APR	7,271	4.51	1.71	58,022	1.67	48,752	1.37	51,493	0.27	0.34	3.04	0.17
Lapsed	NOV-DEC	7,968	1.82	1.23	20,630	0.81	37,862	0.78	39,365	0.13	0.07	0.98	0.05
	JAN-FEB	7,968	0.00	0.00		0.00		0.00		0.00	0.00	0.00	0.00
	MAR-APR	7,968	2.59	1.54	25,046	1.41	37,277	1.16	36,933	0.21	0.11	1.20	0.03

March-April Segment Means

Mar-Apr Segments	Time Period	Segment Size	Active Days	# Txn Types	Mean Balance	Cash-in Txns	Cash-in Amt/Txn	Cash-out Txns	Cash-out Amt/Txn	P2P Txns Received	P2P Txns Sent	Topup Txns	Billpay Txns
Key	NOV-DEC	2,063	14.30	2.79	60,195	9.29	54,926	7.13	53,880	0.95	1.14	9.38	1.16
	JAN-FEB	2,063	15.82	3.15	61,753	11.44	56,591	8.22	51,806	1.21	1.18	9.57	0.81
	MAR-APR	2,063	24.02	3.79	68,726	16.43	53,974	13.25	47,446	1.77	1.91	15.58	0.75
Regular	NOV-DEC	7,239	6.28	2.04	43,484	2.81	48,531	2.41	49,743	0.41	0.42	4.02	0.35
	JAN-FEB	7,239	6.44	2.25	40,328	2.89	48,362	2.51	50,270	0.43	0.38	3.85	0.26
	MAR-APR	7,239	8.87	3.17	38,098	3.74	48,183	3.50	46,818	0.64	0.50	5.15	0.19
Occasional	NOV-DEC	7,292	3.61	1.40	57,609	1.33	52,309	1.06	53,351	0.24	0.46	2.54	0.26
	JAN-FEB	7,292	3.47	1.46	61,010	1.26	46,004	0.99	49,773	0.20	0.24	2.39	0.20
	MAR-APR	7,292	3.51	2.26	38,692	1.27	38,907	0.66	43,777	0.21	0.31	2.40	0.14
Lapsed	NOV-DEC	6,780	3.11	1.71	22,097	1.36	35,506	1.28	36,708	0.20	0.13	1.84	0.09
	JAN-FEB	6,780	1.91	1.31	17,395	0.85	32,806	0.86	34,745	0.13	0.07	1.07	0.03
	MAR-APR	6,780	0.00	0.00		0.00		0.00		0.00	0.00	0.00	0.00

APPENDIX II

FII SURVEY: DATA AND CUSTOMER SEGMENTATION

Data processing

We received the survey data from The Financial Inclusion Insights Program by InterMedia. We received the survey for Uganda, Tanzania, and Kenya along with geo-coded locations of respondents for Waves II and III.[24] We conducted extensive restructuring of the FII data to enable an analysis of the data across the three countries covered in this report and across time. We identified questions that were common across waves and countries and recoded variable names to enable cross-wave and cross-country analysis. The three waves asked many of the same questions but they were given different variable names in the SPSS output requiring extensive recoding, and certain crucial questions were re-framed between waves requiring clever fixes. All in all we integrated the data sets across waves to create three data sets of ~9000 respondents each and integrated the geo-coded data across countries to create a data set of ~18,000 respondents.

With the FII Survey data, we incorporated geographic data from other sources. We integrated geographic information on the administrative region of the geo-coded respondents from the Geospatial Data Abstraction Library (www.gdal.org). We integrated geographic information on population density at the respondent's location using the Gridded Population of the World, v4: UN-Adjusted Population Density 2015, which we accessed through the Socioeconomic Data and Applications Center (sedac.ciesin.columbia.edu). Finally, we received plain text versions of the FSP Maps database of GPS coordinates for mobile money agents in Uganda, Tanzania, and Kenya from Insight2Impact (www.fspmaps.com, i2ifacility.com). Using this we calculated the distance from each respondent to the nearest mobile money agent recorded in the database.

Customer segmentation

The FII Survey data includes detailed survey questions about specific use cases, framed as: "Have you ever used a mobile money account to do the following?" We standardized the possible answers across waves and countries. It also asked respondents "Apart from today, when was the last time you conducted any financial activity with this mobile money service?" for each of the mobile money services the respondent reported to have used, where the answers could be: "Yesterday", "In the past 7 days", "In the past 30 days", "In the past 90 days", or "More than 90 days ago". We use the most recent of these reported intervals.

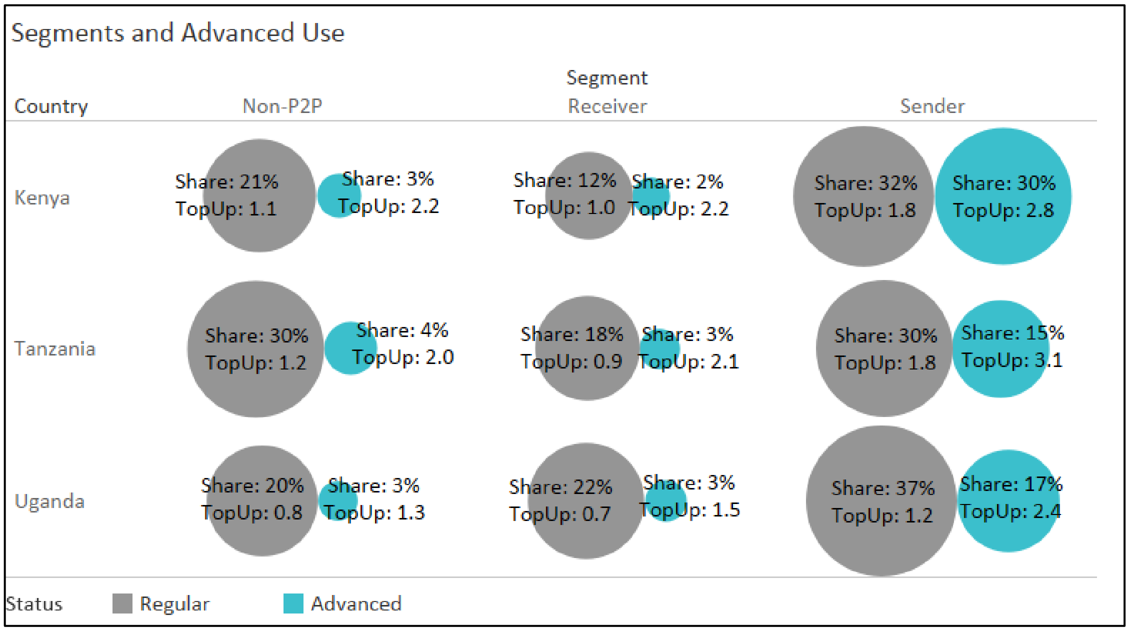
Across all three countries and all years, most of the reported activity was concentrated in five types of use: deposits (cash in), withdrawals (cash out), sending money to or receiving money from friends, family, and associates (P2P), and airtime top ups.

Appendix II, Table 1: Share of Respondents Reporting Mobile Money Uses, by Country

Use	Kenya	Tanzania	Uganda
Withdraw	93%	94%	89%
Deposit	81%	71%	59%
Receive	71%	57%	65%
Top-up	66%	54%	38%
Send	62%	46%	53%
Save	16%	6%	9%
Utility	10%	11%	5%
School Fee	8%	4%	6%
Wages	7%	2%	4%
Loan	5%	1%	1%
Goods	4%	1%	1%

[24] GPS coordinates were anonymized by adding a jitter in the range of (.001 – .002) to both the latitude and longitude.

Use of more specific services: paying utility bills, school fees, or medical fees, making purchases, saving, investing, making or taking loans, receiving wages, and conducting various business activities were far less commonly reported. We find a clear relationship between these less common ("advanced") uses and sending P2P transfers, see figure below. Cognizant that our intent is to identify key mobile money users, we employ this insight to create our segmentation, grouping sending P2P and "advanced" uses together as "intentional" uses. Excluded from this count are particularly common and/or passive uses: depositing or withdrawing cash, purchasing airtime, receiving a transfer from another person, and receiving a payment from the government.



We segment survey respondents based on the number of distinct "intentional" mobile money uses reported and their reported most recent use of mobile money. We consider a respondent to be a *key user* if they report using mobile money "In the past 7 days" and report two or more "intentional" use cases. We consider a respondent to be an *intentional user* if they report using mobile money "In the past 90 days" and report one or more "intentional" use cases. Users who report no "intentional" use cases or have not used mobile money in the past 90 days are considered to be *passive users*. *Non-users* report no users at all. This results in small groups of *key users* who represent 13%, 14%, and 27% of mobile money users in Uganda, Tanzania, and Kenya, respectively.

Survey Segmentaion

Apart from today, when was the last time you used a mobile money service?

Country	Intentional Uses	Never	More than 90 days ago	In the past 90 days	In the past 30 days	In the past 7 days	Yesterday
Kenya	3 or more		22	31	210	426	341
	2		49	62	245	361	161
	1		152	151	452	541	221
	0	1,264	199	153	437	360	151
Tanzania	3 or more		8	11	74	111	105
	2		23	26	121	163	64
	1		116	132	432	401	135
	0	2,820	201	174	438	277	169
Uganda	3 or more		8	24	81	109	60
	2		27	31	100	121	55
	1		143	136	358	287	114
	0	3,363	199	176	297	224	88

Segment

Key User | Intentional User | Passive User | Non User

While we intend to capture similar users with this segmentation in this dataset as algorithmic clustering did in our first dataset, there are inherent data limitations and the segmentations are not entirely equivalent. The FII Survey data is based on recall and does not ask questions that can be used to estimate total volume of usage. We hope that our combination of reported "intentional" use diversity and reported recency of use reflects this adequately, and the segment means suggest that this is the case. Of course, frequent users who make use of many different services are also interesting in their own right! Note that it is possible to conduct a study where the behavioral and survey information are truly linked. This would require directly linking survey respondents to their behavioral data, via their phone number, and is a promising avenue for future research given strict anonymity provisions.

APPENDIX III

AIRTEL: ANALYSIS

Independent variables

From the variables extracted from the Airtel Uganda transaction and voice calling data, we select a group of variables that are individually interpretable. When possible, we define variables orthogonally to one another. As an example, we include the *fraction* of voice contacts who are Airtel Money users rather than the total number as that would be strongly collinear with the total number of voice contacts.

- Inferred demographic or spatial variables (*fixed*)
    a. Average size of a user's cash-in transactions*#
    b. Population density in the user's 'home location'*#
    c. A dummy variable for if the user's 'home location' is in Kampala*
- Observed user behavior (*behavior*)
    a. Number of voice contacts
    b. Average number of calls per voice contact
    c. Number of distinct counties with voice contacts (who use Airtel Money)
    d. Percent of calls made by the user that are routed through a user's 'home tower'
    e. Percent of calls during daylight hours (6am-6pm)
    f. Percent of calls initiated by the user
- Local environment for mobile money (*network*)
    a. Percent of voice contacts who are Airtel Money users
    b. Percent of Airtel Money voice contacts who share user's 'home location'
    c. Average cash-in transactions by Airtel Money voice contacts
    d. Average cash-in amount by Airtel Money voice contacts#

* Denotes a variable derived from the full six-month dataset
# Denotes that a log transformation was done to this variable prior to inclusion in regressions

Note that our reliance on voice contacts means that any Airtel Money user who is not an Airtel Uganda mobile phone user, or who made no calls in a time period, is excluded from our analysis.

We conducted mixed effect logistic regression analysis using these independent variables on the propensity of a *user* to be identified as a *key user* in each of the time periods. We used Stata's xtmelogit regression model to include a random effect at the level of the individual and at the level of Ugandan counties to control for multiple observations per individual and the geographic sample clustering by cell tower. To evaluate consistency, we also ran each two-month period separately.

```
Mixed-effects logistic regression               Number of obs      =     66015

--------------------------------------------------------------------------
                |   No. of       Observations per Group       Integration
 Group Variable |   Groups    Minimum    Average    Maximum      Points
----------------+---------------------------------------------------------
      L2adminID |      158          1      417.8       9133           7
             ID |    33910          1        1.9          3           7
--------------------------------------------------------------------------

                                                Wald chi2(13)      =   1582.79
Log likelihood = -15464.508                     Prob > chi2        =    0.0000

-----------------------------------------------------------------------------------------------
                     key_user |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
------------------------------+----------------------------------------------------------------
         logAvgAmtTransCASHIN |    1.14714   .0777412    14.76   0.000     .9947695     1.29951
               logPopDensity |   .6406827   .0876498     7.31   0.000     .4688922    .8124732
                      Kampala |  -.5477266   .1295343    -4.23   0.000    -.8016091   -.2938441
               NVoiceContacts |   .0083462   .0007356    11.35   0.000     .0069044     .009788
              CallsPerContact |   .1658383   .0078032    21.25   0.000     .1505443    .1811323
    NcountiesMMVoiceContacts |   .0485778   .0101456     4.79   0.000     .0286927    .0684628
            PercHitsHomeTower |  -.0026983   .0014092    -1.91   0.056    -.0054602    .0000636
                 PercCallsDay |  -.0085681   .0022816    -3.76   0.000     -.01304   -.0040963
                 PercCallsOut |   -.033804   .1489163    -0.23   0.820    -.3256747    .2580667
      PercentMMVoiceContacts |   .0112343   .0024013     4.68   0.000     .0065278    .0159407
                 Percentlocal |   .0021509   .0014126     1.52   0.128    -.0006177    .0049196
  MMVoiceContactsAvgTransCASHIN |   .0597997   .0087799     6.81   0.000     .0425915    .0770079
logMMVoiceContactsAvgAmtCASHIN |   .1704078   .0988882     1.72   0.085    -.0234094    .3642251
                        _cons |  -14.32947    .621183   -23.07   0.000    -15.54696   -13.11197
-----------------------------------------------------------------------------------------------
```

November-December 2014

```
Mixed-effects logistic regression               Number of obs      =     21437
Group variable: L2adminID                       Number of groups   =       155

                                                Obs per group: min =         1
                                                               avg =     138.3
                                                               max =      3025

Integration points =   7                        Wald chi2(13)      =    873.16
Log likelihood = -5582.5975                     Prob > chi2        =    0.0000

-----------------------------------------------------------------------------------------------
                     key_user |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
------------------------------+----------------------------------------------------------------
         logAvgAmtTransCASHIN |    .563863   .0647829     8.70   0.000      .436891    .6908351
               logPopDensity |   .3057007   .1040818     2.94   0.003     .1017042    .5096973
                      Kampala |   -.231745   .1597188    -1.45   0.147    -.5447881    .0812981
               NVoiceContacts |   .0041278   .0006049     6.82   0.000     .0029421    .0053134
              CallsPerContact |   .0812785   .0070515    11.53   0.000     .0674579    .0950991
    NcountiesMMVoiceContacts |    .043908   .0094904     4.63   0.000     .0253071     .062509
            PercHitsHomeTower |  -.0026079   .0013985    -1.86   0.062     -.005349    .0001331
                 PercCallsDay |  -.0063819   .0023595    -2.70   0.007    -.0110065   -.0017574
                 PercCallsOut |  -.4039159   .1399603    -2.89   0.004     -.678233   -.1295988
      PercentMMVoiceContacts |   .0066079   .0025957     2.55   0.011     .0015204    .0116955
                 Percentlocal |   .0012541   .0014296     0.88   0.380    -.0015478    .0040561
  MMVoiceContactsAvgTransCASHIN |   .0781316    .011072     7.06   0.000     .0564308    .0998324
logMMVoiceContactsAvgAmtCASHIN |  -.0029902   .1124585    -0.03   0.979    -.2234048    .2174243
                        _cons |  -6.932445   .6134796   -11.30   0.000    -8.134843   -5.730047
-----------------------------------------------------------------------------------------------
```

January-February 2015

```
Mixed-effects logistic regression               Number of obs      =     22446
Group variable: L2adminID                       Number of groups   =       156

                                                Obs per group: min =         1
                                                               avg =     143.9
                                                               max =      3193

Integration points =   7                        Wald chi2(13)      =    962.48
Log likelihood = -5289.9182                     Prob > chi2        =    0.0000
```

key_user	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
logAvgAmtTransCASHIN	.6648689	.0657588	10.11	0.000	.5359839	.7937538
logPopDensity	.3134279	.0944896	3.32	0.001	.1282318	.4986241
Kampala	-.3110656	.1475702	-2.11	0.035	-.600298	-.0218333
NVoiceContacts	.0035122	.0005783	6.07	0.000	.0023787	.0046457
CallsPerContact	.1048094	.0068766	15.24	0.000	.0913315	.1182872
NcountiesMMVoiceContacts	.032292	.0093611	3.45	0.001	.0139447	.0506394
PercHitsHomeTower	-.0025456	.0014502	-1.76	0.079	-.0053879	.0002967
PercCallsDay	-.0064608	.0023074	-2.80	0.005	-.0109832	-.0019385
PercCallsOut	.0565192	.1450451	0.39	0.697	-.227764	.3408023
PercentMMVoiceContacts	.0064829	.0025761	2.52	0.012	.0014338	.011532
Percentlocal	.0032498	.0014531	2.24	0.025	.0004018	.0060978
MMVoiceContactsAvgTransCASHIN	.0551954	.0101256	5.45	0.000	.0353496	.0750411
logMMVoiceContactsAvgAmtCASHIN	.2143791	.1092951	1.96	0.050	.0001647	.4285935
_cons	-8.722235	.6034731	-14.45	0.000	-9.90502	-7.539449

March-April 2015

```
Mixed-effects logistic regression               Number of obs      =     22132
Group variable: L2adminID                       Number of groups   =       136

                                                Obs per group: min =         1
                                                               avg =     162.7
                                                               max =      2915

Integration points =   7                        Wald chi2(13)      =    927.07
Log likelihood = -6210.4958                     Prob > chi2        =    0.0000
```

key_user	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
logAvgAmtTransCASHIN	.7149747	.0592007	12.08	0.000	.5989434	.831006
logPopDensity	.3666272	.0941905	3.89	0.000	.1820172	.5512371
Kampala	-.2621368	.1447693	-1.81	0.070	-.5458794	.0216059
NVoiceContacts	.003405	.000556	6.12	0.000	.0023154	.0044947
CallsPerContact	.0817869	.006164	13.27	0.000	.0697056	.0938681
NcountiesMMVoiceContacts	.0381813	.0093951	4.06	0.000	.0197673	.0565953
PercHitsHomeTower	-.0014848	.0012704	-1.17	0.243	-.0039747	.0010052
PercCallsDay	-.0060339	.0019954	-3.02	0.002	-.0099449	-.0021229
PercCallsOut	-.0471642	.1300878	-0.36	0.717	-.3021316	.2078033
PercentMMVoiceContacts	.0044347	.0022772	1.95	0.051	-.0000284	.0088979
Percentlocal	.0009068	.0012729	0.71	0.476	-.001588	.0034016
MMVoiceContactsAvgTransCASHIN	.0268641	.008779	3.06	0.002	.0096576	.0440705
logMMVoiceContactsAvgAmtCASHIN	.1166496	.1005988	1.16	0.246	-.0805205	.3138197
_cons	-8.078445	.5601935	-14.42	0.000	-9.176404	-6.980486

APPENDIX IV

FII SURVEY: ANALYSIS

<u>Independent variables</u>

We grouped the restructured survey data to help us categorize and understand these user segments. Some explanation is required because these groupings are not what you traditionally see in social science survey data descriptions–traditionally the first two groupings would all be considered respondent characteristics. Broadly this is true, but we have classified a number of variables as "financial network" behaviors to highlight their direct relationship with personal use cases of mobile money.

Occupation and marital status are both "financial network" characteristics of the respondents because what is important about these variables, from our perspective, is how they identify people with differing financial relationships with other people. This is clearly the case for marital status – being married puts an individual in an intimate relationship with another person with corresponding obligations, including financial obligations, which might entail the digital movement of money. An individual's occupation is a product of their education and determines their poverty level, but these characteristics are taken into account by other variables included in the survey data. So what is important about occupation itself is how it shapes one's day-to-day work interactions and the need, or not, to move money digitally. For example, a person who owns a business may need to move money digitally on a regular basis as part of their buying and selling of goods and services, while a farmer may only need to use mobile money once in a while to pay for seasonal inputs or receive harvest payments. Similar logic applies to the bank account and loan variables, which establish whether an individual has a formal relationship with the banking system already.

We identify a respondent's distance to a mobile money agent and the local penetration rate of mobile money as variables that convey whether the local environment for mobile money is favorable.

The phone access/ownership variable has a clear independent causal effect on a respondent's ability to use mobile money, and is thus included separately.

- Fixed demographic and spatial characteristics
    a. Age
    b. Under 20 binary variable
    c. Gender
    d. Progress out of Poverty Index (PPI probability)
    e. Education#
    f. Population density at the respondent geo-location@
- Behaviors affecting personal financial network
    a. Marital status#
    b. Occupation*#
    c. Ownership of a bank account
    d. Survey reported loan behavior
- Local environment for mobile money, including community access considerations

a. Distance in kilometers to the nearest mobile money agent from household GPS location

b. Fraction of *other respondents* within the highest administrative area who are mobile money *users*

- Individual access considerations

  a. Access and ownership of a cell phone

# Denotes a variable where multiple survey categories were combined
* Denotes a category requiring considerable standardization to be comparable across country and wave
@ Denotes that a log transformation was done to this variable prior to inclusion in regressions

Descriptive tables

Education Split

Country	Segment	Education: Post Secondary	Completed secondary	Some secondary	Completed primary	Some primary	No formal education	Other
Kenya	Key User	314	389	161	257	143	21	4
	Intentional User	206	453	262	510	368	105	9
	Passive User	83	260	235	391	365	176	13
	Non User	15	80	187	208	410	338	26
Tanzania	Key User	35	127	57	208	9	3	4
	Intentional User	38	278	107	804	63	40	2
	Passive User	27	220	125	836	107	89	2
	Non User	47	244	237	1,634	286	365	7
Uganda	Key User	90	58	117	37	34	7	2
	Intentional User	124	159	380	165	241	56	6
	Passive User	69	121	365	189	340	73	5
	Non User	26	130	509	485	1,424	774	15

Gender Split

Segment	Gender	Country: Kenya	Tanzania	Uganda
Key User	Female	643	184	142
	Male	646	259	203
Intentional User	Female	1,144	587	594
	Male	769	745	537
Passive User	Female	981	769	719
	Male	542	637	443
Non User	Female	838	1,567	2,096
	Male	426	1,253	1,267

Urban/Rural Split

Segment	Urban Rural	Kenya	Tanzania	Uganda
Key User	Urban	700	254	186
	Rural	589	189	159
Intentional User	Urban	744	466	489
	Rural	1,169	866	642
Passive User	Urban	505	493	405
	Rural	1,018	913	757
Non User	Urban	299	572	500
	Rural	965	2,248	2,863

Est. Poverty Rate

Segment	Country: Kenya	Tanzania	Uganda
Key User	31.8%	72.5%	40.4%
Intentional User	44.0%	77.3%	56.1%
Passive User	55.4%	81.0%	62.6%
Non User	71.8%	88.6%	79.9%

Bank Account Split

Segment	Bank Account	Country: Kenya	Tanzania	Uganda
Key User	Bank account holder	766	138	162
	No bank account	523	305	183
Intentional User	Bank account holder	595	173	278
	No bank account	1,318	1,159	853
Passive User	Bank account holder	294	142	179
	No bank account	1,229	1,264	983
Non User	Bank account holder	56	385	123
	No bank account	1,208	2,435	3,240

Occupation Split

Segment	Occupation	Kenya	Tanzania	Uganda
		Country		
Key User	Professional	127	48	53
	Business	256	89	97
	Labor	440	104	89
	Farmer	214	113	51
	Other/Unknown	12	4	2
	Not employed	240	85	53
Intentional User	Professional	94	68	106
	Business	284	179	166
	Labor	501	265	319
	Farmer	452	576	251
	Other/Unknown	17	15	10
	Not employed	565	229	279
Passive User	Professional	38	30	49
	Business	149	117	129
	Labor	319	187	293
	Farmer	397	704	320
	Other/Unknown	7	14	6
	Not employed	613	354	365
Non User	Professional	9	39	40
	Business	85	131	192
	Labor	195	257	694
	Farmer	262	1,801	1,234
	Other/Unknown	4	27	9
	Not employed	709	565	1,194

Phone Split

Segment	Phone	Kenya	Tanzania	Uganda
		Country		
Key User	Own phone	1,268	428	338
	Access phone	21	11	6
	No phone		4	1
Intentional User	Own phone	1,746	1,274	1,012
	Access phone	159	54	107
	No phone	8	4	12
Passive User	Own phone	1,237	1,200	873
	Access phone	263	183	244
	No phone	23	23	45
Non User	Own phone	395	1,576	1,096
	Access phone	480	842	1,150
	No phone	389	402	1,117

We conducted mixed effect logistic regression analysis using these independent variables on the propensity of a *respondent* to be a *user*. We used Stata's xtmelogit regression model to include a random effect at the level of the smallest administrative areas (www.gdal.org) to control for geographic sample clustering observed in the data.

UGANDA – 6001 respondents, 2638 mobile money users:

```
Mixed-effects logistic regression               Number of obs      =      6001
Group variable: _smAdmin_ID                     Number of groups   =       998

                                                Obs per group: min =         1
                                                               avg =       6.0
                                                               max =        37

Integration points =   7                        Wald chi2(26)      =   1203.24
Log likelihood = -2688.4542                     Prob > chi2        =    0.0000
```

mm_user	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Year						
2015	.4453988	.0871074	5.11	0.000	.2746715	.6161262
age	-.00003	.0002262	-0.13	0.895	-.0004732	.0004133
age_U20	-.3494979	.1397031	-2.50	0.012	-.6233109	-.0756849
2.gender	-.1281712	.0805661	-1.59	0.112	-.2860779	.0297356
ppi_prob	-.0092563	.0016464	-5.62	0.000	-.0124831	-.0060295
education						
2	.439399	.1307397	3.36	0.001	.1831539	.695644
3	.6761507	.1455709	4.64	0.000	.3908371	.9614644
4	1.12787	.1422249	7.93	0.000	.8491147	1.406626
5	1.152216	.180752	6.37	0.000	.7979486	1.506483
6	1.738309	.2737646	6.35	0.000	1.20174	2.274878
7	.4097664	.5091383	0.80	0.421	-.5881263	1.407659
marital_status						
2	.0600114	.1053807	0.57	0.569	-.1465311	.2665538
3	-.0368094	.1427657	-0.26	0.797	-.3166251	.2430062
4	.1586235	.5830323	0.27	0.786	-.9840988	1.301346
log_popdensity	.1805924	.1100159	1.64	0.101	-.0350349	.3962196
occupation						
1	.0955934	.1023466	0.93	0.350	-.1050022	.296189
2	.509222	.1365779	3.73	0.000	.2415343	.7769097
3	.1158106	.1049582	1.10	0.270	-.0899036	.3215248
4	.6161016	.2296838	2.68	0.007	.1659297	1.066273
5	.9429884	.5436291	1.73	0.083	-.122505	2.008482
registered_BANK	.850756	.1308605	6.50	0.000	.5942742	1.107238
loan_imp	.1593788	.0869705	1.83	0.067	-.0110802	.3298378
phone						
1	1.415256	.161842	8.74	0.000	1.098051	1.73246
2	2.814428	.1565481	17.98	0.000	2.507599	3.121257
log_dist_MM	-.3843491	.1161024	-3.31	0.001	-.6119056	-.1567925
L1Admin_userfrac	1.972002	.285849	6.90	0.000	1.411748	2.532256
_cons	-4.084453	.4008485	-10.19	0.000	-4.870101	-3.298804

TANZANIA – 5993 respondents, 3181 mobile money users:

```
Mixed-effects logistic regression               Number of obs     =      5993
Group variable: _smAdmin_ID                     Number of groups  =       847

                                                Obs per group: min =        1
                                                               avg =      7.1
                                                               max =       49

Integration points =   7                        Wald chi2(25)     =    920.90
Log likelihood = -2951.8377                     Prob > chi2       =    0.0000
```

mm_user	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Year						
2015	1.031462	.1114841	9.25	0.000	.8129571	1.249967
age	-.0003601	.0004355	-0.83	0.408	-.0012136	.0004934
age_U20	-.5214609	.1542328	-3.38	0.001	-.8237515	-.2191702
2.gender	-.2964056	.077201	-3.84	0.000	-.4477169	-.1450944
ppi_prob	-.0139837	.002474	-5.65	0.000	-.0188328	-.0091347
education						
2	.4203137	.1854088	2.27	0.023	.0569191	.7837083
3	.9487527	.1436059	6.61	0.000	.6672903	1.230215
4	1.118887	.1983052	5.64	0.000	.7302155	1.507558
5	1.56693	.1816913	8.62	0.000	1.210821	1.923038
6	1.333646	.3059593	4.36	0.000	.7339769	1.933315
7	1.175516	.7372178	1.59	0.111	-.2694041	2.620437
marital_status						
2	.0289885	.103823	0.28	0.780	-.1745008	.2324778
3	.0641203	.1465612	0.44	0.662	-.2231345	.351375
log_popdensity	-.0689327	.0929708	-0.74	0.458	-.2511521	.1132866
occupation						
1	-.3888457	.1133013	-3.43	0.001	-.6109122	-.1667792
2	.4912142	.1613552	3.04	0.002	.1749639	.8074645
3	.1401042	.1393184	1.01	0.315	-.1329549	.4131632
4	.5606098	.2700893	2.08	0.038	.0312445	1.089975
5	.22909	.3659235	0.63	0.531	-.4881068	.9462868
registered_BANK	-.2160657	.120759	-1.79	0.074	-.4527491	.0206177
loan_imp	-.0203299	.1176694	-0.17	0.863	-.2509577	.210298
phone						
1	1.137876	.2312638	4.92	0.000	.6846076	1.591145
2	3.031528	.2212735	13.70	0.000	2.59784	3.465216
log_dist_MM	-.1904105	.0946454	-2.01	0.044	-.3759122	-.0049089
L1Admin_userfrac	3.399822	.4884243	6.96	0.000	2.442527	4.357116
_cons	-3.829876	.5070329	-7.55	0.000	-4.823643	-2.83611

KENYA – 5989 respondents, 4725 mobile money users:

```
Mixed-effects logistic regression               Number of obs      =      5989
Group variable: _smAdmin_ID                     Number of groups   =       706

                                                Obs per group: min =         1
                                                               avg =       8.5
                                                               max =        31

Integration points =   7                        Wald chi2(26)      =   1044.85
Log likelihood = -1819.3894                     Prob > chi2        =    0.0000

------------------------------------------------------------------------------
     mm_user |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        Year |
        2015 |   .3360937   .1060548     3.17   0.002     .1282302    .5439572
             |
         age |  -.0075475   .0037979    -1.99   0.047    -.0149912   -.0001037
     age_U20 |  -.8180432   .1717357    -4.76   0.000    -1.154639   -.4814474
    2.gender |  -.1030623    .105413    -0.98   0.328    -.3096681    .1035435
    ppi_prob |  -.0043041   .0017947    -2.40   0.016    -.0078217   -.0007865
             |
   education |
          2  |   .5315888   .1510956     3.52   0.000     .2354469    .8277308
          3  |   .9007046   .1667809     5.40   0.000       .57382    1.227589
          4  |   .8882235   .1904242     4.66   0.000     .5149989    1.261448
          5  |   1.296458   .2029931     6.39   0.000     .8985987    1.694317
          6  |    1.48995   .3381274     4.41   0.000     .8272329    2.152668
          7  |  -.7092167    .384902    -1.84   0.065    -1.463611    .0451773
             |
marital_status |
          2  |   .3310926   .1338257     2.47   0.013     .0687991    .5933861
          3  |   .3637983   .2025935     1.80   0.073    -.0332778    .7608743
          4  |   15.88556   824.0848     0.02   0.985    -1599.291    1631.062
             |
log_popdensity | -.0407697   .0926653    -0.44   0.660    -.2223904     .140851
             |
  occupation |
          1  |   .2229221   .1278949     1.74   0.081    -.0277472    .4735914
          2  |   .4523706   .1652963     2.74   0.006     .1283959    .7763453
          3  |   .3868204   .1311347     2.95   0.003     .1298011    .6438397
          4  |   .6588648   .4372376     1.51   0.132    -.1981051    1.515835
          5  |   .2478002   .6739895     0.37   0.713    -1.073195    1.568795
             |
registered_BANK | 1.122697  .1674421     6.70   0.000     .7945164    1.450877
    loan_imp |   .1200297   .1183028     1.01   0.310    -.1118395     .351899
             |
       phone |
          1  |   2.178728   .2180972     9.99   0.000     1.751266    2.606191
          2  |   4.201709   .2177885    19.29   0.000     3.774852    4.628567
             |
 log_dist_MM |  -.2516638   .1145491    -2.20   0.028    -.4761759   -.0271516
L1Admin_userfrac | 2.092675 .4435803     4.72   0.000     1.223274    2.962077
       _cons |  -4.259209   .4874387    -8.74   0.000    -5.214571   -3.303847
------------------------------------------------------------------------------
```

We conducted mixed effect logistic regression analysis using these independent variables on the propensity of a *mobile money user* to be a *key mobile money user*. We used Stata's xtmelogit regression model to include a random effect at the level of the smallest administrative areas (www.gdal.org) to control for geographic sample clustering observed in the data. Cases with marital status 'Other', mobile access 'None', and occupation 'Unknown' are dropped to avoid small, over-defined categories.

UGANDA – 2553 users, 345 key mobile money users:

```
Mixed-effects logistic regression               Number of obs      =      2553
Group variable: _smAdmin_ID                     Number of groups   =       724

                                                Obs per group: min =         1
                                                               avg =       3.5
                                                               max =        32

Integration points =   7                        Wald chi2(23)      =    234.54
Log likelihood = -831.91079                     Prob > chi2        =    0.0000

------------------------------------------------------------------------------
        key_user |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------+------------------------------------------------------------
            Year |
            2015 |   .4906549   .1450869     3.38   0.001     .2062898    .7750199
                 |
             age |  -.0003091   .0005465    -0.57   0.572    -.0013801     .000762
         age_U20 |  -.9508337   .4168416    -2.28   0.023    -1.767828   -.1338392
        2.gender |  -.6748105   .1460084    -4.62   0.000    -.9609817   -.3886393
        ppi_prob |  -.0139132   .0026383    -5.27   0.000    -.0190842   -.0087422
                 |
       education |
               2 |  -.1393204   .4512997    -0.31   0.758    -1.023852    .7452106
               3 |   .2848695    .451641     0.63   0.528    -.6003306     1.17007
               4 |   .4953782   .4305205     1.15   0.250    -.3484264    1.339183
               5 |   .5071353   .4506582     1.13   0.260    -.3761385    1.390409
               6 |   .9122787   .4626266     1.97   0.049     .0055473     1.81901
               7 |    .802961   .9292555     0.86   0.388    -1.018346    2.624268
                 |
  marital_status |
               2 |   .0528113   .1592236     0.33   0.740    -.2592613    .3648839
               3 |  -.2028942   .2668847    -0.76   0.447    -.7259785    .3201902
                 |
  log_popdensity |  -.2790485   .1654875    -1.69   0.092     -.603398     .045301
                 |
      occupation |
               1 |   .5258779   .2427306     2.17   0.030     .0501347    1.001621
               2 |   1.385717   .2132461     6.50   0.000     .9677627    1.803672
               3 |   .4081522   .2052189     1.99   0.047     .0059306    .8103738
               4 |   .6168623   .2488228     2.48   0.013     .1291785    1.104546
                 |
 registered_BANK |   .5666757   .1503955     3.77   0.000     .2719058    .8614455
        loan_imp |   .1200421   .1566566     0.77   0.444    -.1869991    .4270834
         2.phone |   1.859893   .4650064     4.00   0.000     .9484977    2.771289
     log_dist_MM |  -.2890675   .1958671    -1.48   0.140    -.6729599    .0948248
L1Admin_userfrac |   .9299954   .4444974     2.09   0.036     .0587965    1.801194
           _cons |  -3.799692   .8262293    -4.60   0.000    -5.419071   -2.180312
------------------------------------------------------------------------------
```

TANZANIA – 3116 users, 443 key mobile money users:

```
Mixed-effects logistic regression               Number of obs      =      3116
Group variable: _smAdmin_ID                     Number of groups   =       690

                                                Obs per group: min =         1
                                                               avg =       4.5
                                                               max =        32

Integration points =   7                        Wald chi2(23)      =    217.18
Log likelihood = -1095.1363                     Prob > chi2        =    0.0000
```

key_user	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Year						
2015	-.3628726	.1406308	-2.58	0.010	-.638504	-.0872413
age	-.0006069	.000745	-0.81	0.415	-.0020671	.0008532
age_U20	.2514169	.2660675	0.94	0.345	-.2700659	.7728996
2.gender	-.3634449	.1279063	-2.84	0.004	-.6141366	-.1127531
ppi_prob	-.0115895	.0033347	-3.48	0.001	-.0181253	-.0050536
education						
2	.5528515	.7034374	0.79	0.432	-.8258605	1.931564
3	1.219624	.6084091	2.00	0.045	.0271642	2.412084
4	1.660422	.6337031	2.62	0.009	.4183865	2.902457
5	1.543032	.6189137	2.49	0.013	.3299831	2.75608
6	2.063324	.6671452	3.09	0.002	.7557436	3.370905
7	3.044631	1.069357	2.85	0.004	.9487294	5.140533
marital_status						
2	.078558	.1467921	0.54	0.593	-.2091493	.3662653
3	-.2966716	.2553518	-1.16	0.245	-.7971519	.2038087
log_popdensity	.1060108	.1121862	0.94	0.345	-.11387	.3258916
occupation						
1	.0689279	.2046538	0.34	0.736	-.3321861	.4700419
2	.9797257	.2076703	4.72	0.000	.5726993	1.386752
3	.6037275	.1969609	3.07	0.002	.2176912	.9897639
4	.5597563	.2706877	2.07	0.039	.0292182	1.090294
registered_BANK	.8959895	.156971	5.71	0.000	.588332	1.203647
loan_imp	.6227805	.1711811	3.64	0.000	.2872717	.9582893
2.phone	1.036743	.3510844	2.95	0.003	.3486306	1.724856
log_dist_MM	-.2124799	.1231566	-1.73	0.084	-.4538624	.0289027
L1Admin_userfrac	.8596682	.6073742	1.42	0.157	-.3307633	2.0501
_cons	-4.453447	.8908844	-5.00	0.000	-6.199549	-2.707346

KENYA – 4651 users, 1289 key mobile money users:

```
Mixed-effects logistic regression               Number of obs     =       4651
Group variable: _smAdmin_ID                     Number of groups  =        680

                                                Obs per group: min =         1
                                                               avg =       6.8
                                                               max =        26

Integration points =   7                        Wald chi2(23)     =     674.24
Log likelihood = -2207.8173                     Prob > chi2       =     0.0000
```

key_user	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Year						
2015	.8563364	.0904116	9.47	0.000	.6791329	1.03354
age	-.0079688	.0035757	-2.23	0.026	-.0149771	-.0009605
age_U20	-.3278424	.2260192	-1.45	0.147	-.7708319	.115147
2.gender	-.3241217	.0838068	-3.87	0.000	-.4883799	-.1598635
ppi_prob	-.006345	.0015237	-4.16	0.000	-.0093314	-.0033586
education						
2	.6428134	.2653957	2.42	0.015	.1226474	1.162979
3	.7952316	.2607585	3.05	0.002	.2841543	1.306309
4	.9164577	.2718415	3.37	0.001	.3836583	1.449257
5	1.119283	.2658836	4.21	0.000	.5981609	1.640406
6	1.484176	.2839205	5.23	0.000	.9277017	2.04065
7	.4001996	.6603224	0.61	0.544	-.8940086	1.694408
marital_status						
2	.3542891	.1051384	3.37	0.001	.1482217	.5603566
3	.0698622	.1777042	0.39	0.694	-.2784317	.4181561
log_popdensity	.0862404	.0777276	1.11	0.267	-.066103	.2385838
occupation						
1	.2347972	.130249	1.80	0.071	-.0204863	.4900806
2	.8343468	.1266212	6.59	0.000	.5861739	1.08252
3	.6045115	.1119405	5.40	0.000	.3851122	.8239108
4	.5034776	.1766242	2.85	0.004	.1573005	.8496547
registered_BANK	.904913	.0870327	10.40	0.000	.7343321	1.075494
loan_imp	.4533806	.1013127	4.48	0.000	.2548114	.6519498
2.phone	1.532967	.2395482	6.40	0.000	1.063461	2.002473
log_dist_MM	-.1043096	.101921	-1.02	0.306	-.304071	.0954518
L1Admin_userfrac	1.183479	.451203	2.62	0.009	.2991373	2.067821
_cons	-5.608589	.5404099	-10.38	0.000	-6.667773	-4.549406

To consolidate these results into an easily interpretable form, we created a composite index for each of these regressions along each of the dimensions in our conceptual framework: (1) fixed demographic and spatial characteristics, (2) personal use case for mobile money, and (3) local environment. We add a fourth indicator that accounts for personal mobile phone access/ownership, an independently strong driver. Composite indices are based on a linear combination of the variables in that category using the logistic regression coefficients as weight, creating a variable that is then standardized. The specific definition thus differs for each regression presented.