Dalberg

Differences in mobile money and phone usage between men and women in Uganda

Case study

Contents

I. EXECUTIVE SUMMARY

Women are a powerful source of opportunity for growth, producing more than half of the world's food, and controlling $20 trillion in consumer spending.[1] They are crucial to a country's development, as persistent inequalities between men and women can reduce nationwide progress in health, education, and standard of living.

A key part of empowering women in their communities is making sure that they are financially included. The rapid spread of mobile money throughout Africa can play an important role in increasing financial inclusion among women. Yet, women today still face systemic, and often unmeasured barriers to closing the gender gap, including in mobile money access and usage.

To ensure that policy makers, private companies, and development organizations work towards bridging the gender gap, we need to increase the availability and use of sex-disaggregated mobile phone data. Detailed mobile money indicators that are disaggregated by sex provide critical insights that can be used by relevant actors to prioritize geographies, set up financial touch points to reach women, or identify needs to increase specific product usage and thus financial inclusion among women.

However, worldwide, sex disaggregation in mobile phone data is severely lacking. Even when data comes from the registration systems of the telecom operators, it is rare that it contains reliable information about the sex of subscribers or clients. Consequently, innovative solutions that combine traditional with new data sources can help understand how women behave differently. The insights can inform products and services design to better serve women.

The purpose of this study is to outline the differences in mobile phone and mobile money usage between men and women in Uganda. The activities conducted can be split into three groups: (i) Running a baseline survey which served as a ground-truth data for later sex-disaggregated analyses and for building a predictive machine learning model. For the purpose of the study, a representative sample of 10,500 mobile phone subscribers were interviewed.

(ii) Correlate the collected survey data with telecom data to understand phone usage patterns for men and women. This consisted in calculating 150+ indicators summarizing mobile phone usage for each subscriber, and feeding the indicators into a machine learning algorithm which predicts the sex of subscribers for whom this information is missing or inaccurate in existing Know-Your-Customer (KYC) data.

(iii) Analyzing the findings. In this use case, we describe our analysis of telecom data—call detail records, combined with sex labels—from a phone survey in Uganda. This provides valuable insights on sex-specific patterns in terms of both generic phone usage and mobile money usage.

When merging the survey and usage profiles, we found out that women tend to have fewer calls, with majority of their calls being incoming, and the average duration of incoming calls being longer for

[1] World Economic Forum. (2015). Why financial inclusion for women is critical to shared prosperity. [online] Available at: https://www.weforum.org/agenda/2015/04/why-financial-inclusion-for-women-is-critical-forshared-prosperity/

women. They top up their phones less frequently and with smaller value on average, have fewer contacts and travel less on average.

The case study was conducted in partnership with Data2X, a collaborative technical and advocacy platform dedicated to improving the quality, availability, and use of sex-disaggregated data in order to make a practical difference in the lives of women and girls worldwide.[2]

[2] https://www.data2x.org/

II. Methodology for understanding the digital gender gap using mobile phone data and call survey

Understanding the basics – a ground-truth call survey

Conducting a phone survey is an essential first step in understanding the male and female mobile phone usage behavior in Uganda. The phone survey provided basic demographic information of a random sample of 10,500 subscribers of one of the operators in Uganda. A random selection of the surveyed respondents from the full subscriber base should ensure that the distribution of males and females within the sample is the same as in the full base.

For each respondent, the following information was collected: sex, age, information on phone sharing, occupational status, sector of employment (for those employed), and consent to use this information for research purposes. This information was later combined with the phone usage indicators of the respondents which allowed us to understand sex differences in phone usage as well as train a machine learning model that can predict the sex of mobile phone subscribers with certain accuracy – see next section for more details.

The survey was designed provide accurate sex tags for a sample of the customer base that can be used to train the algorithm to recognize male and female usage patterns. This algorithm can then be run on the full subscriber base.

The survey was implemented by a local research company over a period of one month starting in midSeptember 2018. In order to ensure high response rate, the following measures were taken:

- Only subscribers active in the past 2 weeks were considered for the survey
- Female interviewers were used for this survey, to ensure a sufficiently high response rate from female subscribers who may be reluctant to speak to a male stranger on the phone.
- Call times were spread throughout the day, rather than concentrated during peak working hours when some respondents may be systematically excluded from the sample.
- At least three attempts were made to call each number in the database, at different times of day, before it was discarded, to avoid any skew to the results due to lower response rates among certain segments.

Nearly 30,000 calls were made and 10,500 respondents were successfully reached and interviewed. This accounts to a response rate of 35%.

Understanding the usage patterns

To understand how the subscribers use their mobile phones on a daily basis, we calculated 150+ indicators summarizing mobile phone usage for each subscriber, and fed the indicators into a machine learning algorithm which can predict the sex of subscribers based on their usage patterns. To provide additional insights, we further added a module calculating indicators focusing on mobile money usage.

Using this methodology, we calculated 201 indicators summarizing 60 days of phone usage for all the respondents of the ground-truth survey, with the observation period overlapping with the survey period (autumn 2018). The indicators provide insights on phone usage (e.g. number and duration of calls), mobility patterns (e.g. number of visited cell tower areas), social network (e.g. number of different contacts), mobile money patterns (number and value of transactions), etc. These summary indicators are used to differentiate the mobile usage of men and women.

The usage patterns are calculated from 60 days of anonymized call detail records (CDRs), which contain the logs of all phone activity during the period. The following data streams were used: voice and text CDRs, data CDRs, airtime credit recharges, mobile money transaction records, the overview of localization of cell towers.

The calculated indicators can be divided in six broader groups, as shown in Table 1: Indicator groups and sample indicators. Where possible, these indicators were calculated as an overall value, and separately for incoming and outgoing activity and for different weekdays and times.

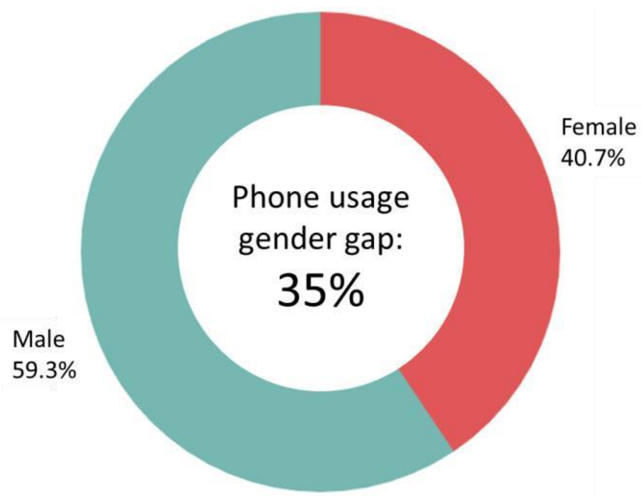
Indicator group	Example of indicators
Phone usage	overall number of data points, number and duration of calls, proportion of outgoing activity, proportion of text messages, volume of data used, average time needed to respond to a text message
Mobility patterns	number of visited cell tower areas and different levels of administrative regions, radius of gyration, home region
Social network	number of different contacts, number of different contacts for incoming/outgoing calls/texts, number of uniquely incoming/outgoing contacts
Top-up patterns	count and value of top-ups on an average day and overall
Mobile money usage	count and value of transactions on an average day and overall, proportion of incoming transactions, number of peers
Other characteristics	urban/rural status

Table 1: Indicator groups and sample indicators. Where possible, these indicators were calculated as an overall value, and separately for incoming and outgoing activity and for different weekdays and times.

III. FINDINGS

Findings from the call survey

The purpose of the survey was two-fold: 1) Gathering up-to-date reliable sex-disaggregated data on phone usage in Uganda, and 2) Validating the existing official and operator's statistics on digital gender gaps.



Up-to-date sex-disaggregated insights on phone usage in Uganda

Figure 1: Proportion of female and male phone users in Uganda as reported in the phone survey.

According to the survey, 40.7% of the subscribers are female, which corresponds to a phone usage gender gap of 35%, when using the gender gap formula used in a recent research paper investigating alternative ways to measure digital gender gap:[3]

Female to male sex ratio of people with characteristic / Female to Male sex ratio of the population[4]

The male-female subscriber proportions are fairly consistent across different age groups, but differ significantly across different regions. While among the rural population, the percentage of female subscribers drops to 38% (being as low as 22% in Karamoja sub-region), in the urban areas the average percentage is 42% with the Central sub-region (including the capital Kampala) showing nearly equal rates of female and male users. Refer to Figure 2 for more details about the geographical heterogeneities across Uganda's sub-regions.

[3] Fatehkia, M., Kashyap, R. and Weber, I., 2018. Using Facebook ad data to track the global digital gender gap. *World Development*, *107*, pp.189-209.
[4] The population gender ratio comes from https://www.ubos.org/onlinefiles/uploads/ubos/pdf%20documents/2017_UNHS_26092017Final_Presentation.pdf

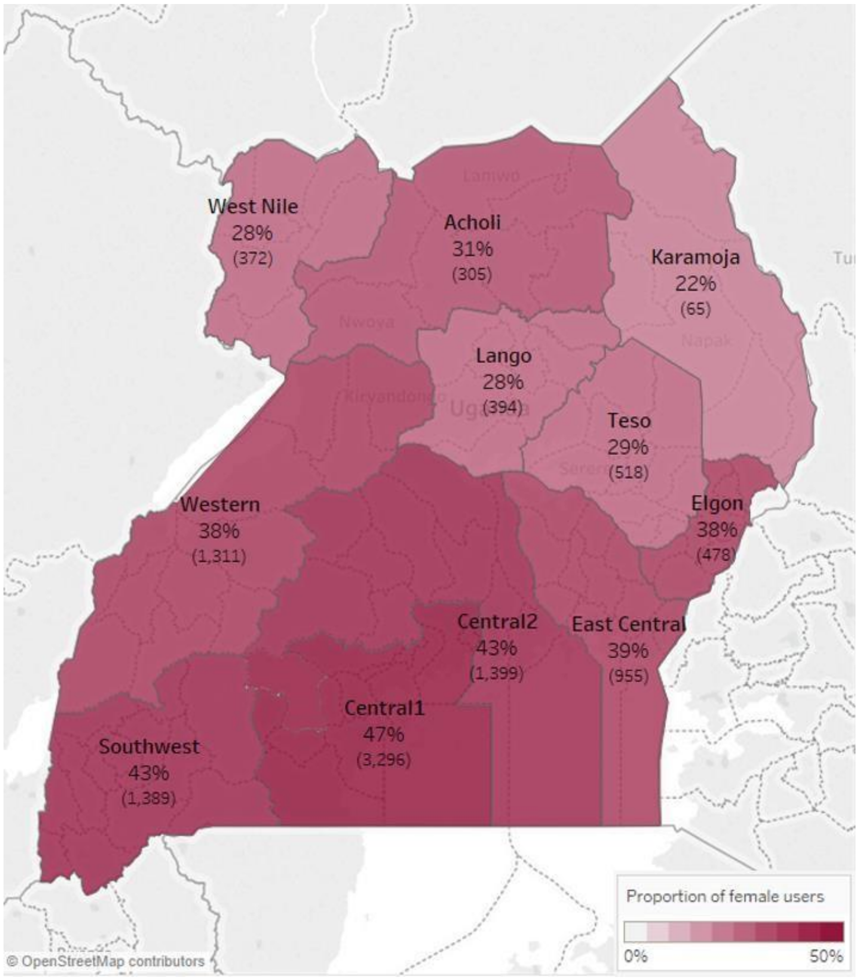


Figure 2: Proportion of female phone users across sub-regions in Uganda. The number in the parentheses corresponds to the sample size for a particular sub-region.

Additionally, the results from the survey suggest that both males and females are equally likely (ca 80%) to share their SIM card with other users.

Comparison of the survey gender and age statistics to such statistics among the full Ugandan population brings additional insights on the generic biases in phone ownership. As the unequal representation of men and women already suggests, males are more likely to own a phone than women (41% of phone users are female, while 51% Ugandans are female[5]). When comparing different age groups, we can see an overrepresentation of people aged 25 – 64 among phone users, while minors (below 15). Youth (aged 15 - 24) is represented in similar percentages among phone users as in the overall population. For detailed comparison, see Figure 3.

[5] https://www.ubos.org/wp-content/uploads/publications/06_2018women_and__men_in_uganda_FF2016.pdf

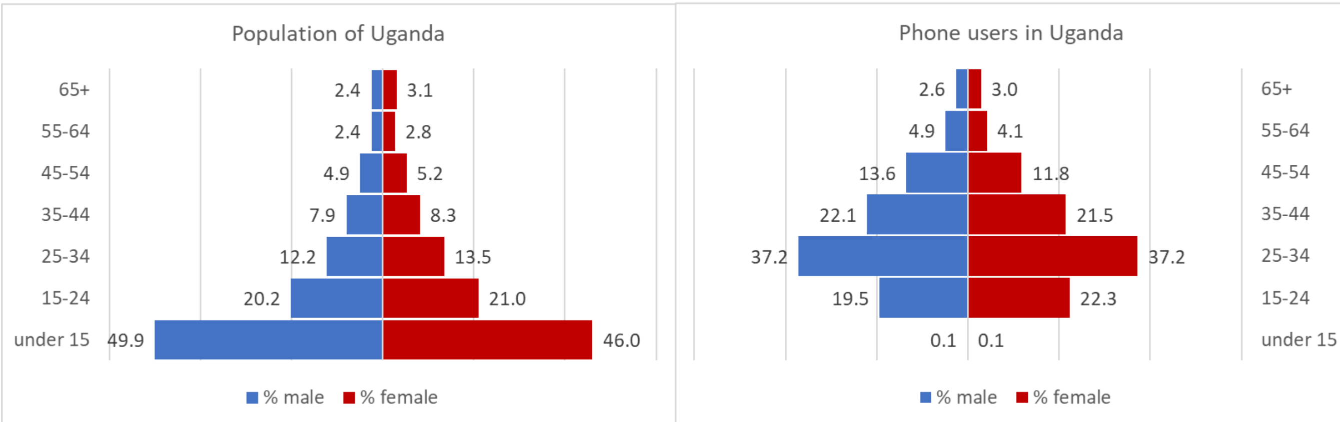


Figure 3: Population pyramids of Ugandan population (on the left) and phone users (on the right), showing a bias in phone ownership towards people aged 25-64.

Validation of existing sex-disaggregated data on phone usage in Uganda

As you can see in Table 2: Comparison of sex labels as listed in the operator's database and in the responses of the phone survey. in the operator's database, 22% users are listed to be female (31% if looking only at customers with known sex). Our survey suggests that the proportion of female phone users is in fact higher: 41% of our survey sample are female.

Sex label	**Operator's database**		**Phone survey**
	incl. unknown	**known only**	
Female	22%	31%	41%
Male	48%	69%	59%
Unknown	33%	-	-

Table 2: Comparison of sex labels as listed in the operator's database and in the responses of the phone survey.

Out of the 10 500 survey respondents, 33% did not have a correct sex label in the operator's database. Out of these, majority were listed as other/unknown sex, 3% males were listed as females in the database, and 9% females were listed as males in the database. While this means that only about 67% of the customers have the correct sex listed in the database, when looking at customers who have some sex specified there, the information is correct for 84% customers.

	Survey label		**Survey label**	
Operator label	**Female**	**Male**	**Female**	**Male**
Female	22.8%	2.9%	28.7%	3.7%
Male	9.4%	44.2%	11.9%	55.7%
Unknown	8.4%	12.2%		

Table 3: Proportions of correctly assigned labels in the operator's database. The table on the left shows the proportions including unknown labels, while the table on the right shows proportion with respect to assigned Sex labels only. The fields in green correspond to proportions of sex labels matching between the database and the survey, the fields in red highlight mismatching labels.

The sex information in the operator's database, if already present, can thus be considered fairly reliable, likely thanks to the recently introduced requirement to present a national ID upon SIM registration. Still, a significant fraction of the user base does not have any sex label at all, and female users are three times more likely to be not listed in the operator's database as the registered users of their sim card. This is in line with the commonly reported practice of men registering the SIM cards of their wives, sisters or other family members, even when their female counterparts are the actual main users of those SIM cards With regards to assessing digital gender gap in Uganda, this is rather a positive signal, as the gap likely is smaller than suggested by official numbers, which are based on the operator's database (as reported to the regulatory authority).

Findings from calculated usage patterns

Generic phone usage

When merging the survey and usage profiles, we could confirm the generic trends reported by previous similar studies form different countries:[8] women tend to have fewer calls, with majority of their calls being incoming, and the average duration of incoming calls being longer for women. They top up their phones less frequently and with smaller value on average, have fewer contacts and travel less on average. Examples of some of these distinctive indicators, along with their distributions are shown in Table 1.

On the other hand, there are no big differences between men and women on data usage, both in terms of volume and count. This finding is surprising in the context of GSMA's Mobile Gender Gap Report[9] which mentions the gender gap in mobile internet usage is systematically larger than the gender gap in mobile phone ownership and usage in most countries. The report does not include data for Uganda specifically, but its neighboring Tanzania conforms to the trend with 23% female vs 36% male subscribers using the internet. We could speculate that we see more balanced usage because our sample could be biased towards more active users – only people active in the two weeks prior to the survey were considered for the sample and people who are more active could also have higher

[8] Blumenstock, J. and Eagle, N., 2010, December. Mobile divides: gender, socioeconomic status, and mobile phone use in Rwanda. In *Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development* (p. 6). ACM;

GSMA Connected Women 2015: Bridging the gender gap: Mobile access and usage in low- and middle-income countries (https://www.gsma.com/mobilefordevelopment/programme/connected-women/bridging-gender-gapmobile-access-usage-low-middle-income-countries/);

Jahani, E., Sundsoy, P.R., Bjelland, J., Iqbal, A., Pentland, A. and de Montjoye, Y.A., 2015. Predicting gender from mobile phone metadata. In *NetMob Conference in Cambridge, MA*;

Stoica, A., Smoreda, Z., Prieur, C. and Guillaume, J.L., 2010. Age, gender and communication networks. *NetMob-Analysis of Mobile Phone Networks 2010 communication proposal*.

[9] https://www.gsma.com/mobilefordevelopment/wp-

content/uploads/2018/04/GSMA_The_Mobile_Gender_Gap_Report_2018_32pp_WEBv7.pdf

response rates to the survey calls. Such bias should however present itself to a similar extent in both internet and generic phone usage.

When comparing to other statistics on (mobile) internet usage in Uganda, it is hard to come to a clear conclusion, as different sources report very different numbers. Uganda's National IT Survey from 2017/2018[6] reports that smaller proportion of female subscribers use their phones for some mobileinternet-related services, like browsing the internet or writing and receiving e-mails, yet the usage rates are similar for other services like social networking or Skype. According to the same survey, lower proportion of the overall female population uses internet (9.5% females vs 15.8% males), yet a higher percentage of women uses the internet through their phone and cellular network, while higher percentage of men use alternative ways to access the internet such as desktop or laptop computers, mobile phones through wireless networks or tablets. However, the Uganda Communication
Commission in its sector performance report from mid-2018[11], based on the data from operators and internet providers, states an overall internet penetration of 47%, a number more consistent with the rate we see in our data. Their data is however not disaggregated by sex.

[6] https://www.nita.go.ug/sites/default/files/publications/National%20IT%20Survey%20April%2010th.pdf [11] https://www.ucc.co.ug/wp-content/uploads/2017/09/Communication-Sector-Performance-for-the-Quarterending-June-2018.pdf

Indicator	Gender dif. p value	Distribution
Avg duration of an incoming call (seconds)	3.61e-100	
Proportion of outgoing activity	2.46e-75	
Avg count of unique contacts per active day	1.00e-56	
Top-up count	1.40e-34	
Call count	3.03e-22	
Avg top-up value	3.25e-05	
Radius of gyration (km)	3.52e-03	
Number of active days	1.19e-02	

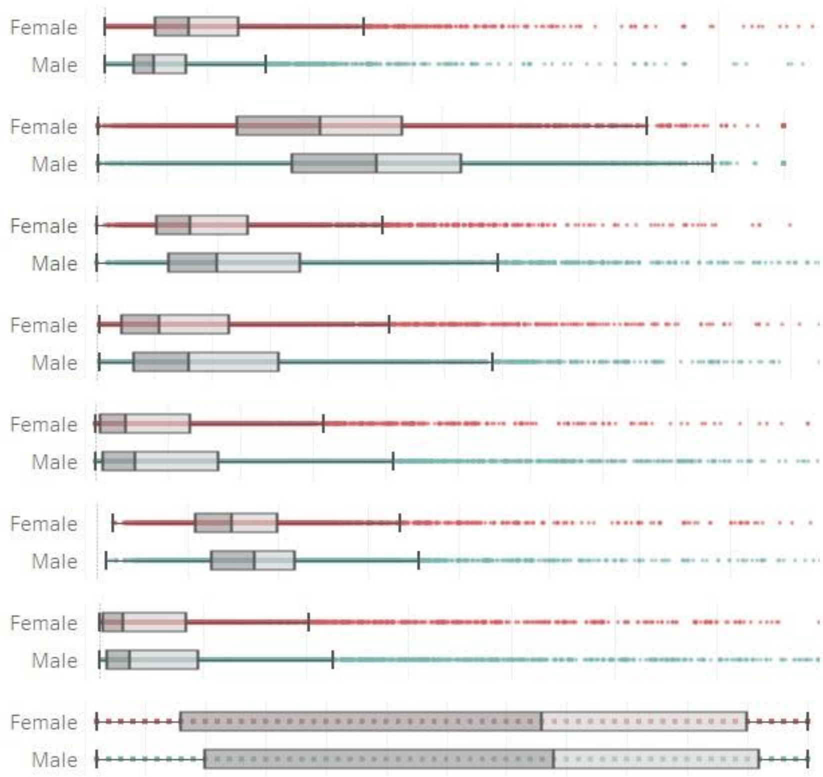


Table 4: Selected phone usage indicators distinctive for sex and their distributions. Note that these features were selected to illustrate the different components of distinct usage patterns, so do not correspond to the strictly most distinctive indicators (these would show for this use case unnecessary repetitions of similar indicators like average duration of an incoming call, but also average duration of an incoming call on a weekday morning etc.). The boxes in the visualization indicate the interquartile range (the middle 50%) of each distribution with the center line showing the median; the whiskers indicate values 1.5 times of the interquartile range (Note that some of the outliers were trimmed for visualization purposes). The indicators were calculated over a period of 60 days; the actual values are not displayed as they can be commercially sensitive for the operator, and the relative distributions are more relevant for the current study than the actual values.

When segmenting the survey respondents further, we found that in urban settings and for younger respondents, the differences between men and women in usage patterns tend to be less pronounced, yet they remain significant for most of the indicators. To pinpoint a few interesting findings from the segment analysis:

- Across all segments, women consistently have longer calls and smaller proportion of outgoing phone activity.
- When looking at social network, in this case represented by the daily average count of unique contacts, we see that across all segments, men connect with more people on a daily basis. However, age seems to be a factor here, as young females in both urban and rural settings have a bigger network of contacts than their elderly male counterparts.
- As for top-up indicators, consistently across all segments, men top-up more and more frequently, yet again, age is a decisive factor, as elderly men typically top up less frequently (though still higher value) than young and adult women. Similarly, urban residency plays a role, with urban women topping up as frequently as their rural male counterparts. *It* is *further*

interesting to note that the average total *top-up* value gets, consistently across all segments, higher with age and tends to be higher in urban settings.

- Average traveled distance, the radius of gyration, again shows the biggest differences for the elderly, with elderly men traveling longer distances in both urban and rural settings than women. Other than that, the traveled distance tends to be longer in rural areas.
- While overall across the user base, men tend to be active on more days, in the segmentation analysis we see that this trend is driven mainly by low activity of elderly rural women. In the remaining segments, the differences are less pronounced, with some segments (young and adult rural as well as young urban) having higher median scores for females than males.

Mobile money usage

For mobile money usage, we found similar, though in general much smaller, differences in sex-specific patterns, illustrated also in Table 5: Selected mobile money indicators and their distributions. The boxes in the visualization indicate the interquartile range (the middle 50%) of each distribution with the center line showing the median; the whiskers indicate values 1.5 times of the interquartile range. The indicators were calculated over a period of 60 days; the actual values are not displayed as they can be commercially sensitive for the operator, and the relative distributions are more relevant for the current study than the actual values.

- More women use mobile money than men, but for most indicators, either men use mobile money more or there is no sex difference. For example, men both deposit and withdraw higher amounts of money over 60 days and overall transact more. Interestingly, women pay on average a higher fee per transaction.

Differences in mobile money and phone usage between men and women in Uganda

- Another example to look at are peer-to-peer transactions, where there is no sex difference in the number of all transactions received over 60 days, yet men do initiate more peer-to-peer transactions, and the average transaction value both received and sent is higher for men.
- Interestingly, men and women don't differ in their average account balance, number of accounts they interact with (though men interact with a higher number of accounts of their peers), or the number of withdrawals they make.

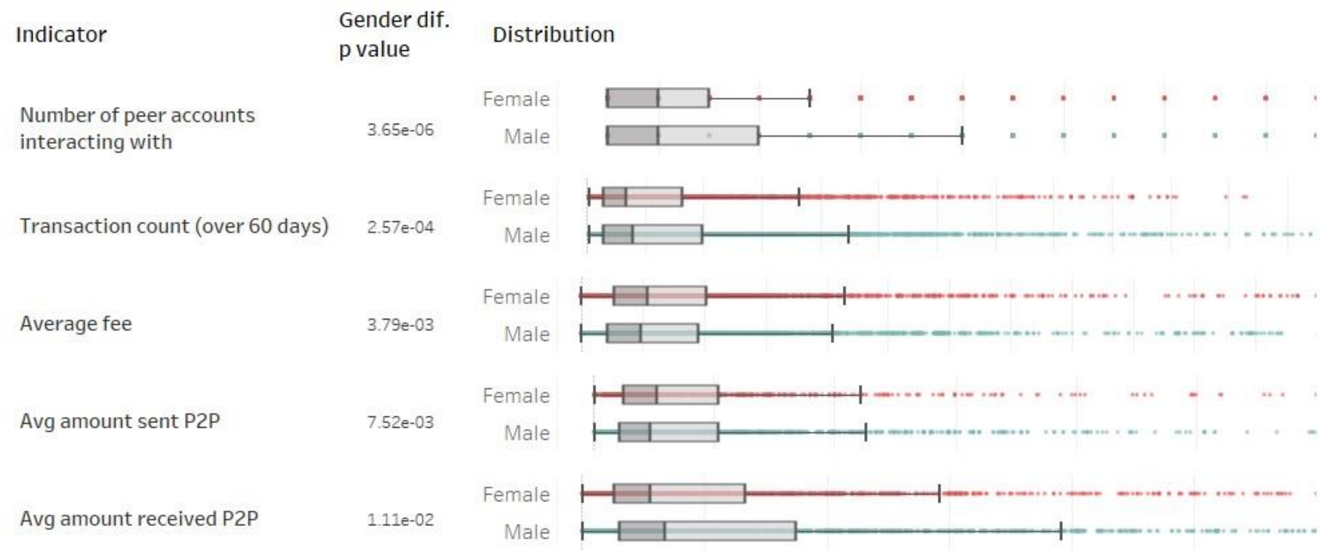


Table 5: Selected mobile money indicators and their distributions. The boxes in the visualization indicate the interquartile range (the middle 50%) of each distribution with the center line showing the median; the whiskers indicate values 1.5 times of the interquartile range. The indicators were calculated over a period of 60 days; the actual values are not displayed as they can be commercially sensitive for the operator, and the relative distributions are more relevant for the current study than the actual values.

We can further explore how these indicators differ across different population segments:

- The number of peers mobile money users interact with is very balanced across different segments. A major exception are adult males who tend to have a bigger peer network.
- As for the transaction count, men tend to transact more and that is consistent across different segments. Living in an urban area is important as well though - urban women across all ages do transact more than their rural male counterparts,
- Among our respondents, an average peer-to-peer transaction was higher when receiving rather than sending money. For both receiving and sending, the average transaction amount tends to be higher with higher age.

IV. Additional analyses

Sex prediction

Apart from calculating the descriptive set of phone and mobile money usage features, our analysis incorporated a predictive model. In the module, the calculated indicators and survey sex labels are used as an input for training several machine learning models to predict sex of the users based on their usage patterns. The best scoring model is then selected and can be used to predict the sex for the full user base, including subscribers that did not participate in the survey. This can provide valuable sex-disaggregation of telecom data for countries where reliable sex labels are not available and can server as a trigger for launching sex-related programs.

When running the predictive module in Uganda, we can reach accuracy of 72% in predicting subscribers' sex and identifying the following features as top contributors towards discriminating between females and males: average top-up value, count of unique contacts per active day, average call duration during the week, radius of gyration and the duration of incoming calls.[7] Consult Table 6 for an overview of the top features, their importance scores and the average value for men and women. These indicators are similar to the top predictive indicators from Bangladesh, though the feature importance scores in Bangladesh were about twice as high.

Indicator	Importance score	Female average	Male average
Average top-up value	0.021	913.56 (st.d. 1971)	1080.90 (st.d. 2060)
Count of unique contacts per active day	0.019	4.88 (st.d. 4)	6.39 (st.d. 5)
Average call duration during the week	0.019	339.75 (st.d. 334)	217.40 (st.d. 233)
Radius of gyration	0.018	1.58 (st.d. 4)	1.84 (st.d. 4)
Average duration of an incoming call	0.018	480.75 (st.d. 453)	304.60 (st.d. 382)

Table 6: Top sex-predicting features in Uganda, their importance scores and the mean values per sex.

While the 72% success rate is in a similar range as rates reported in several research papers focusing on predicting sex from call detail records,[8] it falls below the 84% success rate seen in an similar

[7] the most successful model was XGBoost. Note that the top features and their importances will slightly differ with each run of the model. The feature importance is calculated as an average feature importance across all decision trees. At the single tree level, feature importance is calculated as the weighted reduction in node impurity from a split at a node with the given feature.

[8] Felbo, B., Sundsøy, P., Lehmann, S. and de Montjoye, Y.A., 2017, September. Modeling the Temporal Nature of Human Behavior for Demographics Prediction. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases* (pp. 140-152). Springer, Cham.;

analysis in Bangladesh, run in partnership with the GSMA [9]in Bangladesh. We assume this difference is due mainly to the fact, that the phone usage gender gap is much smaller in Uganda, making it harder to distinguish between men and women based solely on their usage patterns.

Frias-Martinez, V., Frias-Martinez, E. and Oliver, N., 2010, March. A Gender-Centric Analysis of Calling Behavior in a Developing Economy Using Call Detail Records. In *AAAI spring symposium: artificial intelligence for development*.

Jahani, E., Sundsoy, P.R., Bjelland, J., Iqbal, A., Pentland, A. and de Montjoye, Y.A., 2015. Predicting gender from mobile phone metadata. In *NetMob Conference in Cambridge, MA*;

Stoica, A., Smoreda, Z., Prieur, C. and Guillaume, J.L., 2010. Age, gender and communication networks. *NetMob-Analysis of Mobile Phone Networks 2010 communication proposal*.

[9] https://www.google.com/search?client=safari&rls=en&q=sgsma+GAIT+dalberg&ie=UTF-8&oe=UTF-8

V. LIMITATIONS

When assessing the outputs of this study and its replicability, it is good to keep in mind several possible limitations.

First of all, for good ground truth data, it is important to run the phone survey, with all connected organizational and actual costs. While we sampled randomly from the operator's user base in order to ensure maximal representability, some biases can still be present in our sample. For example, subscribers from areas with poorer reception could be underrepresented as they simply would be less likely to answer when being called for the survey. We tried to account for this by redialing at varying calling times. Additionally, we validated that the representation of individual districts correlates highly between the survey respondents and the full user base.

For calculating the usage indicators, access to granular call detail records is needed, which is a proprietary data source including very sensitive information. This results in a lengthy process in order to gain access to the data, making sure all legal regulations are met. To ensure high security of the data, the most granular data is anonymized and stays at the premises of the operator and is only accessed through secured remote connection. Only aggregated and fully deidentified outputs can be exported out of the operator's premises.

As for potential biases in the survey sample, the call detail records are also prone to biases and imperfections, as people commonly share their SIM cards with other or use multiple SIM cards for different purposes. As a result, the usage patterns seen for a particular user in records of one operator might on one hand correspond to the usage of several people or, on the other, not be complete. Our survey included a question on shared usage and in general, 20% of the respondents (19% women, 22% men) share their phone with someone else. This implies that for some of the users, their individual activity is actually smaller than it appears from the data. For our analyses we considered only users who were the main users of each sim card, and thus contributed the most to its activity pattern. We do not have information on multi-simming for our respondents, yet the National IT Survey[10] implies that 45% phone users own more than one sim card. For these, their overall phone usage is in reality higher than what the data from one operator suggests.

[10] https://www.nita.go.ug/sites/default/files/publications/National%20IT%20Survey%20April%2010th.pdf

VI. Recommendations on way to further the financial inclusion for women

To increase financial inclusion among women, we first need to improve understanding of the underserved female population, allowing Financial Service Providers (FSPs) to better reach this market and offer suitable products.

Understanding differences in phone behaviors and mobile money usage is a good starting point. Mobile Network Operators, in particular, can greatly benefit from the findings. In addition to behavioral differences, to fill the sex data gap and fully understand women and strengthen current market approaches, FSPs need to grasp the psychometric and demographic differences, as well. That will help them understand not only how women behave but why women behave the way they do. And while mobile phone data can provide behavioral and demographic insights, the psychometric differences are best understood in small focus groups using human centric design techniques.

By identifying meaningful segments along psychometric, behavioral and demographic characteristics, FSPs can offer tailored products and services to both women and men, alike.

Mobile operators that wish to generate accurate estimates of their subscribers' gender can access the Gender Analysis and Identification Toolkit (GAIT), developed in partnership with the GSMA.[11]

[11] Further details can be found at https://www.gsma.com/mobilefordevelopment/resources/the-gsmas-gender-analysisand-identification-toolkit-gait/

VII. OPPORTUNITIES AND CONSTRAINTS TO INTEGRATING SIMILAR METHODOLOGIES INTO NATIONAL STATISTICAL SYSTEMS

Sex-disaggregated insights are essential in understanding differences across men and women and addressing inequalities.

National Statistics Offices (NSOs) worldwide have the mandate to collect and analyze data to report progress towards achieving the Sustainable Development Goals (SDGs). From a total of 230 monitored indicators, 53 are sex-disaggregated.[12] Apart from reporting on the SDGs, sexdisaggregated insights can inform inclusive policies and programming. While NSOs mainly rely on census and administrative data for national reporting purposes, these data sources are often obsolete and address a limited number of concerns that are of interest to analysis of differences between men and women.

As the findings from our research in Uganda suggest, combining traditional data sources, such as census and surveys (here as our custom phone survey), with new data sources, such as mobile phone data, can help fill data gaps in understanding differences between men and women. The sexdisaggregated insights can further support the design of more inclusive products and services by both, the private and public sectors.

And while the importance of new data sources is increasing, especially in data poor environments, National Statistics Offices are lagging behind in adopting new methods and data sources in national reporting, especially in less developed countries. The main reasons include the lack of standardization in methodologies, the biases associated with the new data sources and the lack of trust and understanding by the NSOs about the potential of new data sources to fill data gaps. Additional burdens to adopting new data sources are the general lack of capacity for working with alternative data sources, as well as the complexity of the big data ecosystem and the necessity to work in different ways with different actors.

To address these challenges, we need to share more stories about the positive impact of sexdisaggregated insights. Additionally, the data innovators and researchers need to work together with the NSOs to demonstrate the potential of new methods; they should be transparent about the biases and the efforts being made to address them. Finally, a focus should be put on capacity building within the NSOs themselves in order to build internal knowledge and resources to deal with alternative data sources in the most efficient and complementary way.

[12] https://www.data2x.org/wp-content/uploads/2017/05/UNWomenList_GenderSDGIndicators.pdf

VIII. ABOUT DALBERG DATA INSIGHTS

Dalberg Data Insights is the Big Data entity of Dalberg Group, active in developing data products and solutions that aim to support more evidence-based policies. International development actors and local governments must often work around data gaps when tackling important social challenges. At the same time, extensive data sources exist publicly and behind the firewalls of private companies, such as mobile phone operators, banks, digital platforms, and satellite operators. We, at Dalberg Data Insights, identify the data solutions to international development challenges. We access, analyse, and integrate data from different sources to design tools and specialized analytics. Using our data products, local and global communities can better target, implement, and evaluate their programs and initiatives.

For more information, please visit https://www.dalberg.com/what-we-do/dalberg-data-insights.