Field Note 12

WHO WILL CHURN? LEVERAGING PREDICTIVE MODELING FOR INSIGHTS AND ACTION ON DFS CUSTOMER INACTIVITY

By: Fabian Reitzug | Contributing Authors: Soren Heitmann, John Irungu Ngahu

mastercard
foundation

IFC | **International Finance Corporation** | **WORLD BANK GROUP**

Creating Markets, Creating Opportunities

TABLE OF CONTENTS

FIGURES AND TABLES

EXECUTIVE SUMMARY

To expand their customer base and to provide access to financial services, mobile network operators (MNOs) and financial institutions (FIs) have invested in digital channels and digital financial services (DFS) product offerings to reach and better serve new customer segments. In Sub-Saharan Africa, growth in DFS adoption and mobile money access has been a success story, with 286 million registered accounts in the region.[1] However, this emphasis on adoption overlooks significant numbers of inactive accounts—where customers signed up for accounts that quickly became dormant, or perhaps the accounts were never used to make transactions. Only 35 percent of all registered accounts in Sub-Saharan Africa were active in 2017.[2] Reducing customer inactivity is important to build a vibrant network of regular users in emerging markets, as well as in maturing DFS markets where targeting new customers or converting customers from other providers becomes increasingly difficult and costly.

To successfully tackle inactivity, it is important to understand who these inactive customers are, their behavior patterns, and why they ceased to be active. Research conducted as part of the Mastercard Foundation-IFC Partnership for Financial Inclusion leverages surveys, segmentation analysis, and predictive modeling to provide deep insights that can drive actions for service providers to boost activity among account holders. IFC-led surveys in Uganda, Zambia, Ghana, and Côte d'Ivoire identify relevant factors for inactivity, such as age, gender, and education. To an extent, inactivity patterns are specific to different markets and providers, but general patterns can also be observed across contexts. For example, educated users are less likely to be inactive in Zambia and Uganda; younger users are more likely to be inactive in Ghana; and women tend to use DFS less actively in Zambia and Ghana.

IFC data analytics segmentation and network analysis techniques uncover the behavioral patterns of inactivity. For instance, social network size, as measured by voice contacts, sets apart active from inactive and non-users. Linking demographic information, qualitative research on why customers became inactive, and information from surveys makes for rich insights on ways to address root causes of inactivity. For example, irregular income, no perceived need to use DFS, and high service prices emerged as key reasons for inactivity.[3] Therefore, educating customers on the cost-benefits of DFS may address some of these (mis) perceptions. Likewise, ensuring products and marketing messages are relevant to different customer demographics, and revisiting service quality and pricing, provide potential pathways to improved customer retention and reactivation.

However, these strategies are only meaningful before the individual becomes inactive. Certainly, retrospective insights provided through surveys allow DFS providers to understand reasons for inactivity, churn, and customer attrition, and can drive longer-term service changes that improve the value proposition and appeal of DFS to inactive customers. However, these adjustments often come too late to retain customers at risk of churning today. Predictive churn modeling offers a proactive approach to inactivity and provides early warning signs of customers who may be leaving the service. This allows providers to implement marketing and outreach strategies to retain customers.

Churn modeling can be applied using both phone usage and DFS data and is thus available to both financial institutions and MNOs providing digital financial services. IFC research with Tigo Ghana showcases how big data mobile phone usage can be leveraged to identify predictors of churn, using supervised machine learning methods. IFC delivers proof of concept that DFS churners can be identified based on phone usage behavior that sets them apart from non-churners. Likely churners show consistently lower usage in terms of SMS sent and calls made (30 percent and 27 percent lower rates respectively). They also exhibit less variation in the location from where they make and receive calls. The value-add of IFC's research is that it can quantify how much more phone activity differentiates a soon-to-be churner from a dedicated customer. Applying machine learning techniques, the most predictive features are used to compute the probability of each customer to leave the service. These 'churn scores' allow for the identification and targeting of customers at highest churn risk within a provider's portfolio.

One of the applications of churn modeling is to test and implement effective outreach measures for likely churners. A|B testing is an experimentation technique to improve marketing by benchmarking different product offerings and contact strategies, thereby identifying the most effective churn prevention approach.

Based on projections by IFC, campaigns that deploy churn modeling to target customers at risk of churning can have a substantial impact in terms of customer activity, retention, and provider revenue. In addition to informing on churn, predictive models can be modified to deliver insights on DFS uptake and activity. This illustrates that data analytics techniques such as predictive modeling have wide-ranging business applications and can support forward-looking decision-making of both MNOs and FIs providing DFS.

This paper shows that while reasons for churn and inactivity are complex, inactivity can be reversed, and churn ameliorated. Neither surveys nor data analytics provide a complete remedy on their own. However, providers are likely to see benefits if they invest in surveys that provide an in-depth understanding of the DFS customer perspective and then complement this with analytic models, especially predictive churn modeling. These models can more precisely identify potential churners and improve the cost, efficiency, and reach of marketing campaigns aimed at preventing inactivity. The benefit of this approach is that it enables providers to deploy proactive re-engagement and re-activation strategies—finely targeting at-risk customers and preventing inactivity and churn before it occurs.

1 Global Findex, 2017.
2 Ibid.
3 IFC, 2018a. The Mobile Banking Customer that Isn't: Drivers of Digital Financial Services Inactivity in Cote d'Ivoire

INTRODUCTION

Digital financial services have dramatically expanded financial access to previously unbanked populations in Sub-Saharan Africa. While the share of adults with financial institution accounts has remained flat, the share with a mobile money account has almost doubled since 2014.[4] In 2017, there were 286 million registered mobile money accounts in the region[5] and 21 percent of adults in the region had a mobile money account.[6] Yet despite widespread uptake, only 100 million of these accounts are in use, yielding an activity rate of just 35 percent. As Figure 1 demonstrates, high inactivity rates are a phenomenon observed in almost all markets. Only the most mature countries, such as Kenya, appear to be resilient to these dynamics.

Figure 1: DFS Inactivity in Sub-Saharan Africa (IMF Financial Access Survey, 2017)
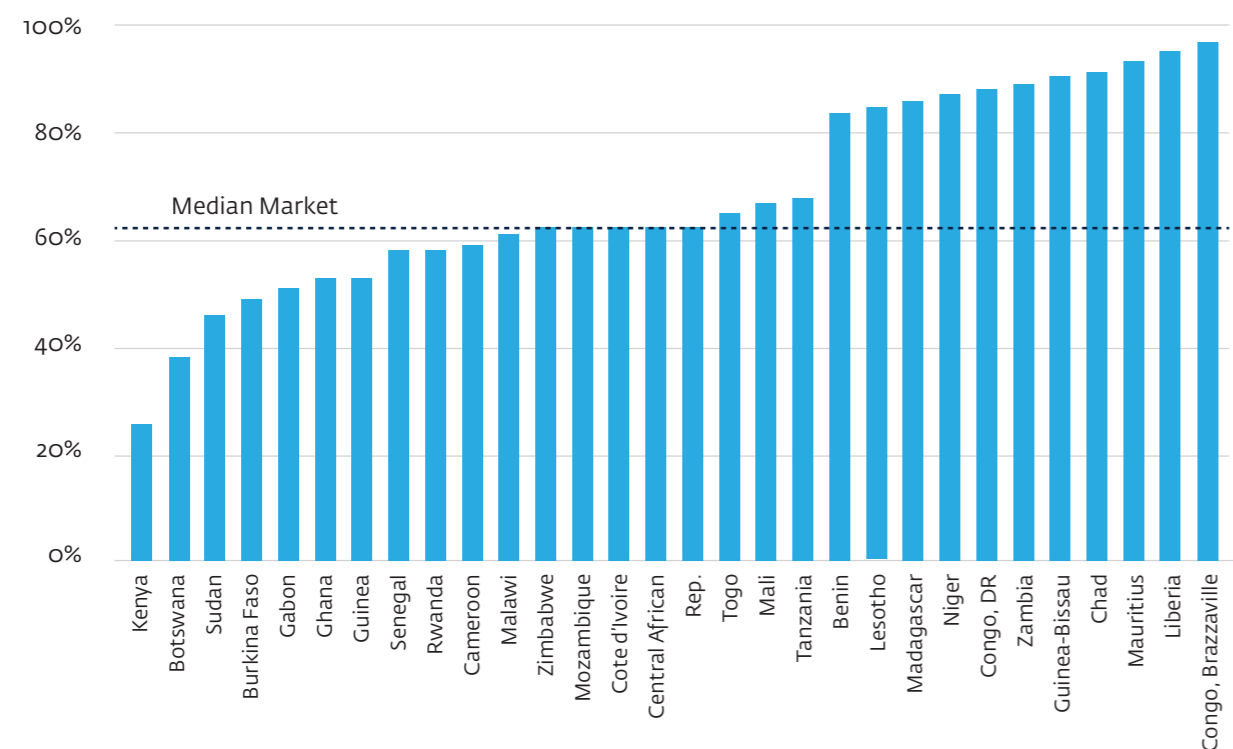
Inactivity (account non-usage), churn (customer attrition), and dormancy (temporary inactivity) are important issues in the industry, both for MNOs and financial institutions providing DFS. As shown in Figure 1, in the median African market, more than 60 percent of customers are inactive. To tackle this challenge, providers should seek to better understand this important segment in their portfolios. While often framed as a problem, addressing these issues represents a significant opportunity for providers who manage to prevent customers from leaving their service or those that succeed in reactivating dormant customers. Globally, mobile money providers with a 'highly active' customer base—meaning a large proportion of their customers conduct at least one revenue generating transaction in 90 days—earn 42 percent higher monthly revenue per active customer (USD 1.7) than the average across all providers (USD 1.2).[7] Thus, converting customers into more active users represents an important business opportunity, especially given that institutions often have many such customers in their portfolios.

While DFS churn has emerged as an important issue, it is interesting to note that MNOs often introduce mobile money in order to reduce churn on the Global Services of Mobile Communications (GSM) channel. According to the GSM Association (GSMA), average annual churn of MNO customers in developing countries, so-called GSM churn, was 26 percent in 2018.[8] Even the highly successful provider Safaricom (the parent company of M-Pesa) experienced substantial churn of 21 percent per year as of September 2017, representing 5.5 million customers[9]. For financial institutions, churn of bank customers is also a concern. The Boston Consulting Group estimates that globally, attrition affects between 30 to 50 percent of a bank's client base on an annual basis,[10] further highlighting that customer churn is a prevalent phenomenon with negative impacts in DFS, GSM, and banking services.

Estimates from different industries show that it can cost five to 25 times more to acquire a new customer than to retain an existing one.[11] Customer attrition causes MNOs in the Middle East and Africa region to lose an estimated 25 to 35 percent of annual revenues.[12] In the financial institution space, banks globally are estimated to lose 10 to 15 percent of gross revenues to customer attrition.[13] Tackling customer inactivity is key for providers in emerging markets to build a vibrant network of active customers and reach critical mass. It is also important in maturing DFS markets, where targeting new customers or converting customers from other providers becomes increasingly difficult and costly.

Understanding the causes of inactivity is an efficient and cost-effective way to identify activity and retention strategies based on in-depth research insights. Retrospective inactivity identification techniques, particularly segmentation and surveys of inactive customers, contribute in important ways to knowledge of churn and inactivity. This paper will show that state-of-the-art predictive analytics offers additional insights and a forward-looking perspective by proactively identifying potential churners so they can be prevented and dissuaded from leaving the service.

WHAT DRIVES CUSTOMER INACTIVITY

Reasons for customer inactivity are multifaceted and can be situated at different levels: the customer level (for instance, a lack of understanding of the service); the provider level (for example, low investment in and poor management of the service); or the market level (a lack of competition or positive network effects). Identifying salient analytic drivers of inactivity through surveys, customer segmentation, and qualitative research provides insight into who inactive customers are, why they became inactive, and on which channels.

Demographic characteristics of inactive users

IFC conducted surveys of over three thousand inactive customers in Uganda, Zambia, Ghana, and Côte d'Ivoire to identify demographic characteristics and common reasons for inactivity.[14]

While inactivity patterns are to a certain extent market- and provider-specific, these cross-country surveys found commonalities among churners in terms of age, gender, education, and household size.

- **Age:** While younger people in Ghana are the most active voice users, they also have the largest share of registered, yet inactive, DFS accounts.[15] In Ghana and Zambia, insufficient funds and high costs emerged as two key reasons for inactivity among young users.[16] Different age groups face different barriers to active usage. In Uganda, adults cited inaccessible services and blocked lines as primary reasons for mobile money inactivity, while youth frequently mentioned unreliable networks, high transaction costs, and lost SIM cards.[17]

4 Global Findex, 2017.
5 IMF Financial Access Survey 2017.
6 Global Findex, 2017.
7 GSMA 2017. State of the Industry Report on Mobile Money.

8   GSMA Intelligence Data, 2018. (No data for Africa specifically available.)
9   Safaricom Press Release, 2017. News Release: Un-Audited Results Half Year Ended 30 September 2017.
10 Boston Consulting Group. 2017. How Banks Can Close the Back Door on Attrition.
11 Harvard Business Review, 2014. The Value of Keeping the Right Customers.
12 McKinsey, 2016. Middle East and Africa: Telecommunications industry at cliff's edge.
13 BCG, 2017. Ibid.
14 IFC, 2016. Find the Gap: Can Big Data help to increase Digital Financial Services adoption?
15 IFC. 2016. Find the Gap: Can Big Data help to increase Digital Financial Services adoption?
16 IFC. 2018b. Banking on the Future: Youth and Digital Financial Services in Sub-Saharan Africa" World Bank Group, Washington, D.C.
17 Ibid.

- **Gender:** Women are more likely to be inactive according to IFC surveys in Zambia and Ghana. In addition, women make fewer transactions on average than men.[18] For instance, female DFS customers in DR Congo make two transactions per year (median), while men make four. However, the median value per transaction made by women is 33 percent higher than that by men. Women also have different service needs and transaction relationship preferences. For instance, female clients may prefer to transact with female agents[19] or to use products tailored to their needs, and the absence of these features may drive their inactivity.

- **Education:** In Zambia and Uganda, levels of education are inversely related to inactivity. Education can make people more open and trusting of new technology and thus foster activity.[20]

Identifying the demographics of active and inactive users enables providers to segment the customer base and develop profiles of active and inactive users.

BEHAVIORAL PATTERNS OF ACTIVE VERSUS INACTIVE USERS

The understanding of inactivity is enriched by the analysis of customer usage patterns. IFC surveys of DFS customers in Ghana and Uganda uncovered consistent trends in which roughly 45 percent of respondents received money, 40 percent sent money, and 25 percent deposited money within the previous 30 days. Lower activity was demonstrated for more 'active' transactions (sending and depositing) compared to more passive transactions (receiving money). A big data analysis with Airtel Uganda also demonstrated the positive impact of the size of social networks. Highly active users (so-called "key" users) had roughly twice as many voice contacts as lapsed and occasional users.[21] Another analysis technique, vintage analysis, segments a provider's customers based on activity. In an IFC operational project, a vintage analysis of 4.5 million accounts of an East African financial institution was used to determine the time elapsed since the last transaction. The analysis segmented accounts with a balance greater than USD 1 based on their activity levels. It showed that very few of these customers were inactive for 30 days or more (Figure 2). Such analyses can be conducted on a per-customer or per-product level, depending on the insights sought.

Figure 2: Vintage analysis – days since the last transaction
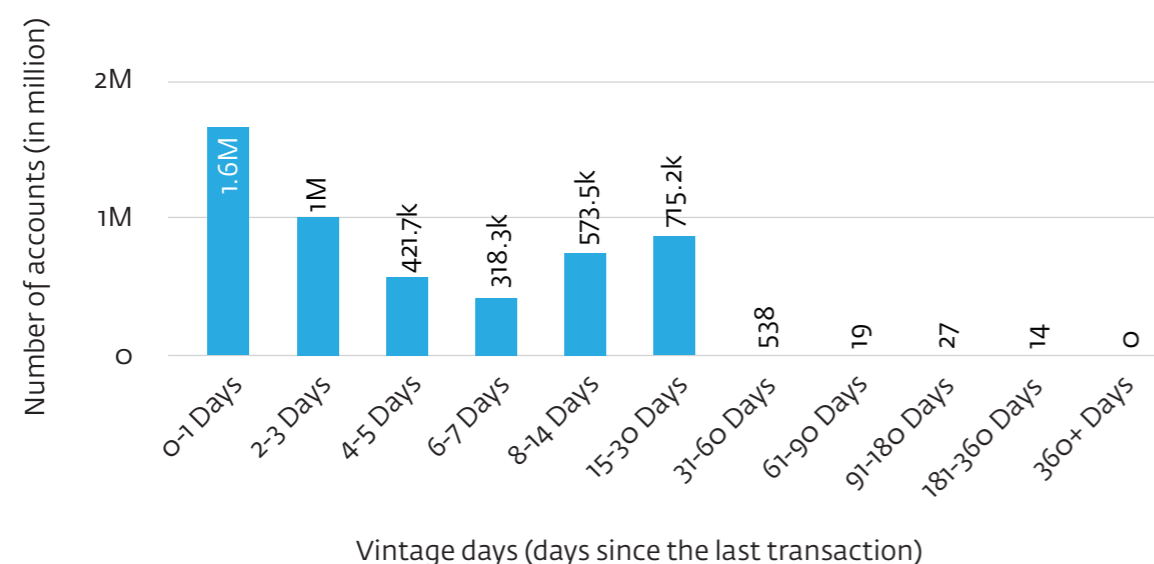
Reasons for inactivity from the user perspective

Demographics and usage patterns tell part of the story of inactivity, yet research that gives voice to inactive users adds in-depth insights, as it allows customers themselves to explain why they ceased to use a service. In Côte d'Ivoire, an IFC survey of one thousand inactive DFS customers found the top three reasons for inactivity

were irregular income (43.6 percent of respondents); no need to use DFS (27 percent); and the service being too expensive (15.5 percent).[22] An ethnographic study by IFC in Cameroon, DR Congo, Senegal, and Zambia, which examined attitudes towards DFS from the perspective of users and nonusers, highlighted a perceived lack of service reliability as a factor creating mistrust and non-usage of mobile money.[23] Customers who cited transfers not going through, or money allegedly 'disappearing' in the system without effective recourse, may decide to stop using the service. For customers, perceptions are reality—providers must take misconceptions seriously and rectify them in order to address barriers to activity.

From inactivity identification to churn prediction

The identification of inactive customers—their demographics, usage patterns, and reasons for inactivity—provides important descriptive information. Overlaying demographics with inactivity reasons can help to address relevant causes for different customer groups. For instance, by educating customers on how to use their service, providers can tackle a root cause of inactivity, especially among older users or women with less exposure to mobile technology. Meanwhile, addressing perceptions that the service is too expensive may be pertinent for younger and more educated people, who are more price-sensitive.[24]

While information from churn identification can be used to refocus a provider's value proposition and outreach to inactive customers, and to finetune products, churn prediction goes a step further. By pointing to customers at risk of leaving the service (future churners), churn prediction generates information that can drive proactive strategies to target at-risk customers and prevent churn.

HOW DATA ANALYTICS ADD ACTIONABLE INSIGHTS

Data analytics expands existing knowledge by providing a more granular understanding of inactivity. While surveys provide *ex post* information on why a customer became inactive, data analytics can predict which customers are likely to become inactive in the future.[25] IFC projects in Ghana and Uganda provide concrete use cases for data analytics in understanding and addressing inactivity and churn. The team used churn modeling to identify and segment inactive customers. This predictive approach uses machine learning to identify features that are associated with future churn and makes

projections about which customers are more likely to leave the service based on the characteristics they display today. The insights generated by churn modeling can inform changes in product design, pricing, incentives, or outreach measures to reduce churn and enhance activity.

OPERATIONALIZING CHURN AND INACTIVITY IN DATA ANALYTICS PROJECTS

In industry parlance, churn and inactivity are sometimes treated as synonyms—and while they are similar concepts, there are subtle differences. *Inactivity* is commonly operationalized as 90-days of non-activity.[26] A customer is thus called 'inactive' if she fails to make at least one financial transaction in this time period. Since passive users who make no cash-ins but receive money and directly cash-out would be considered active under this definition, a useful approach is to restrict the 90-day definition solely to cash-ins. For providers with automatic transactions, such as loan installment or interest payments, a suitable operationalization would be to limit analyses to customer-generated financial transactions.

Churn describes customer attrition and is commonly used in connection with subscriber-based business services, where churn is defined as the expiration of a subscription. For the analytic purposes of predictive modeling, churn occurs when customers have stopped transacting on the digital platform for a period of time long enough to warrant the assumption that they will not transact again.

IFC research operationalizes churn in two ways. Firstly, a period of X days of inactivity is used to make forward-looking predictions of whether customers will remain inactive (IFC used 30 and 90 days). Using this approach in Uganda, IFC determined churn by first identifying phone usage patterns that separate churners from non-churners over a 30-day period. These features were then used to predict churn (i.e. whether inactive customers would remain inactive for a further 60-days). Assuming customers who are inactive for a total of 90 days are unlikely to make any further transactions, this approach gives providers a critical two-month window to re-engage these individuals.

18 IFC. 2018c. *Women and Digital Financial Services in Sub-Saharan Africa: Understanding the Challenges and Harnessing the Opportunities.*
19 Chamboko, Richard; Reitzug, Fabian; Cull, Robert; Heitmann, Soren; van der Westhuizen, Morne. 2019. *Does Gender Matter in Digital Financial Services? Evidence from a highly underdeveloped financial market in DR Congo.* Forthcoming Publication.
20 IFC, 2017a. *A Sense of Inclusion: An Ethnographic Study of the Perceptions and Attitudes to Digital Financial Services in Sub-Saharan Africa.*
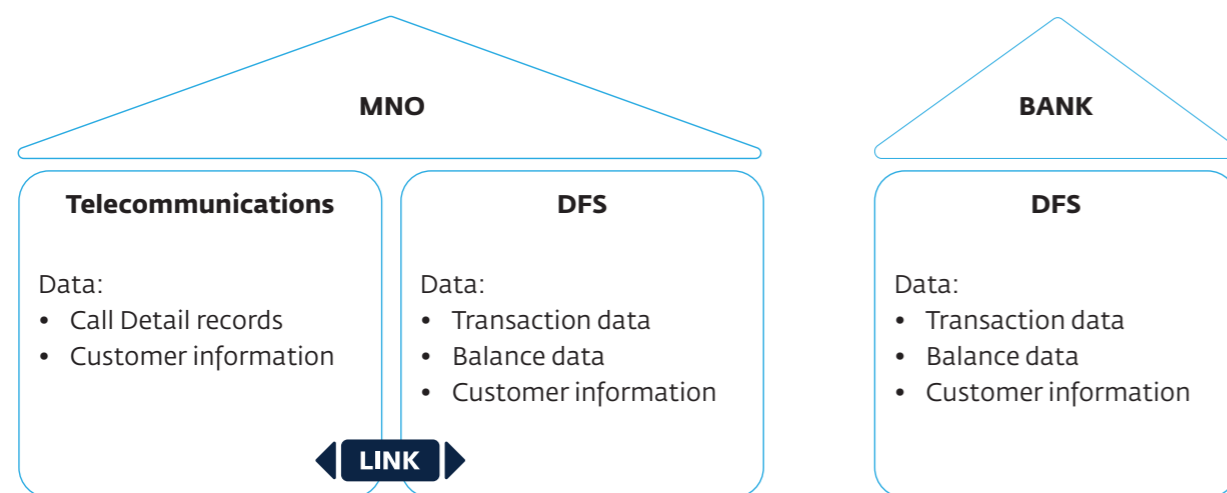21 Mattsson, Carolina; Stuart, Guy. *Understanding Key Mobile Money Users.*

22 IFC, 2018a. *The Mobile Banking Customer that Isn't: Drivers of Digital Financial Services Inactivity in Cote d'Ivoire.*
23 IFC, 2017. *A Sense of Inclusion: An Ethnographic Study of the Perceptions and Attitudes to Digital Financial Services in Sub-Saharan Africa.*
24 IFC. 2018b. *Banking on the Future: Youth and Digital Financial Services in Sub-Saharan Africa"* World Bank Group, Washington, D.C.
25 Beyond churn, IFC has also applied predictive analytics to mobile money adoption and identify potential high activity users. See for instance: IFC, 2017b. *Data Analytics and Digital Financial Services: Handbook.*
26 GSMA, 2017; IFC, 2018; CGAP, 2015.

A second approach particularly relevant to financial institution churn is to operationalize churners as customers who have a positive balance and then show a zero balance for X days thereafter, suggesting they have emptied their accounts. Both these methods illustrate the close link between churn and inactivity. The crucial difference between the two concepts is that churn is a permanent loss of a customer, while inactive customers could be dormant and become active again. Thus, inactivity precedes churn, but does not necessarily lead to churn.

Inactivity and churn are concepts that can be flexibly adapted based on project goals. For instance, inactivity can be defined as 30, 60, 90, or any other number of non-active days. There are possibilities to look at churn and inactivity on a per-product or per-geographic-area basis, using different types of transactions, and to conduct analyses of a single provider or for a whole market, data permitting. One can differentiate between *channel churn* (when a customer becomes inactive on the agent channel but remains active at branches) and *provider churn* (a customer becomes inactive on all the provider's channels). *Churn identification* is a retrospective analysis technique that can, for instance, establish who churned based on whether customers emptied their accounts, while *churn prediction* is prospective, identifying customers who will likely churn.

Applying suitable metrics to operationalize churn and inactivity is a key challenge and no approach is infallible. Customers who are predicted to churn may become active again (the assumption of churn may thus be unwarranted), and customers with zero balance could still deposit money again or seek a new loan in the future. For this reason, churn prediction risks a high level of false-positives. Still, incorporating providers' on-the-ground knowledge can help to operationalize inactivity and churn in ways that are meaningful to the specific context in which they are applied and superior to above-the-line engagement.

Figure 3: Data available in banks and MNOs to model and predict inactivity and churn

MNO

BANK

Telecommunications

Data:
- Call Detail records
- Customer information

DFS

Data:
- Transaction data
- Balance data
- Customer information

LINK

DFS

Data:
- Transaction data
- Balance data
- Customer information

DATA AVAILABLE TO ANALYZE INACTIVITY DIFFER BETWEEN MNOS AND BANKS

Both MNOs and banks have transaction and balance data at their disposal, as well as information on customers. However, MNOs possess additional data from complimentary products, most notably, telephony data. This includes information on phone calls made, SMS sent and received, and mobile data usage. This difference is relevant when it comes to churn prediction where Call Detail Records (CDR) are an important source of input information.

Linking CDR and DFS data can pick up 'early warning signs' of churn. CDR data is rich in information, such as numbers of outgoing and incoming calls and SMS, size of network and distance between calls, and the number of different locations from where calls originate. Information from phone usage contains more events than those available through DFS data, where transactions occur less frequently. Hence, CDR data provides 'high-resolution' information that can be leveraged to make churn predictions.

Usage patterns on the DFS channel itself may also vary between MNOs and banks. For instance, after repaying a loan, bank customers may not necessarily take out another loan. IFC also found that access to bank agents impacts customer activity, with customers who use agents being more active than those who rely on branches.[27,28] The differences in products, channels, and services between MNOs and banks delivering DFS influences the type and characteristics of the data available. For these reasons, IFC has been better able to predict churn for MNO-led DFS and succeeded in identifying churn for banks.

PREDICTING CUSTOMER INACTIVITY WITH CHURN MODELING: A DATA ANALYTICS PROJECT IN GHANA

Predictive churn modeling is a 'big data' analytics technique that identifies features that predict which customers may cease to be active on a service. Most commonly used in the MNO space, IFC collaborated with researchers from the University of California Berkeley to apply this technique to DFS. This novel operationalization provides a basis for the development of measures that tackle inactivity and churn.

Figure 4. Mapping of voice data in Ghana
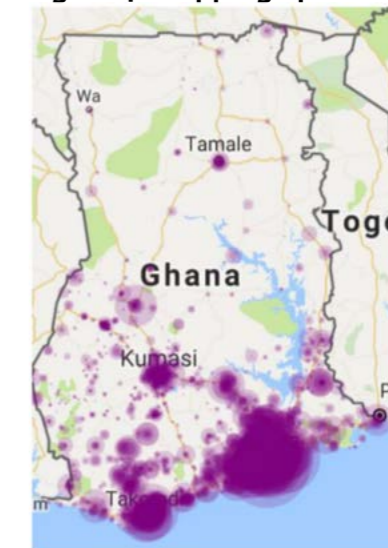
Deploying a multi-step process to transform data into actionable insights, the team worked with Tigo Ghana to analyze CDRs. A sample of 20,000 non-churners (active voice and DFS users) and 20,000 DFS churners was drawn for the analysis. The team defined DFS churners as those who used the service in the first three-month period, making at least one transaction, and then became inactive on DFS in the subsequent three-month period.

From this sample of 40,000 users, the team sought to identify the features most predictive of churn by applying two selection processes: a classical statistical approach (t-test) and a machine learning approach using gradient boosting.

Table 1 details the most predictive features selected through gradient boosting and through t-tests. It shows that predicted DFS churners are less active across all top 10 of the selected GSM metrics than non-churners (see last column). In other words, these top characteristics have a reliable statistical relationship with customer churn.

The top three features from both gradient boosting and t-test reveal that fewer *outgoing* calls, fewer locations, and smaller network size of this *outgoing* traffic are all associated with churn. On average, churners send 30 percent fewer SMS and make 27 percent fewer calls than non-churners. Realizing that reduced telephone service activity correlates with reduced activity on the digital channel is an obvious point—the value of the model is to quantify *how much less* GSM activity constitutes a trigger point for churn and, importantly, to use this as an engagement point for the provider. Regarding the top 10 features, GSM activity levels of 16 to 30 percent below that of DFS non-churners indicated elevated risk of DFS churn, in this specific model.

27 Buri, Sinja; Cull, Robert; Gine, Xavier; Harten, Sven; Heitmann, Soren. 2018. *Banking with Agents: Experimental Evidence from Senegal.*
28 IFC, 2019. *Do Agent Networks Boost Savings? Effects on institutional deposit mobilization and customer saving behavior.*

Features were selected by splitting the sample using a subsample to train the model, and then applying it to the remainder of the data to test how well the features identify churners. The output of the churn model are scores bounded between zero and one, indicating the probability that a customer who used the service in the first three months of the period of analysis will stop using the service in the three following months. Computing churn scores for the entire customer-base allows for the identification and targeting of customers at highest risk, thereby providing opportunities to proactively prevent churn.

Table 1: Most predictive features of churn selected through gradient boosting and t-test

Rank	Gradient Boosting	T-Test			
	Top 10 features selected through gradient boosting	Top 10 features selected through t-test (sorted by percent difference)	Characteristics of churners and non-churners compared		
			DFS churners	DFS non-churners	Absolute mean difference between churners and non-churners1	Percent difference between churners and non-churners
1	Number of outgoing calls per day	Total number of outgoing SMS	45.928	65.976	-20.048***	-30%
2	Total duration of outgoing calls	Size of the outgoing SMS network	6.987	9.811	-2.824***	-29%
3	Total number of distinct outgoing call locations	Total number of outgoing calls	578.939	794.361	-215.422***	-27%
4	Size of the outgoing call network	Size of the incoming SMS network	1.821	2.417	-0.596***	-25%
5	Number of days with incoming calls	Total number of incoming SMS	21.829	28.411	-6.582***	-23%
6	Number of distinct locations from which calls are received	Number of distinct locations from which data network is used	20.799	26.731	-5.932***	-22%
7	Number of distinct locations from which data network is used	Size of the outgoing calls network	72.259	91.822	-19.563***	-21%
8	Variance per day in the number of calls received from different locations	Total number of incoming calls	239.486	301.368	-61.882***	-21%
9	Variance per day in the average duration of calls	Size of the incoming calls network	21.119	26.333	-5.214***	-20%
10	Variance per day in the geographic distance for incoming SMS	Total number of data network transactions	1004.985	1193.192	-188.207***	-16%

1 Statistical significance: * for <.05; ** for <.01; *** for <.001

A comparison of predictive accuracy demonstrates that the model using features selected through gradient boosting performs slightly better and forecasts who will churn with 58 percent accuracy. The logistic regression using features selected through t-tests performs marginally worse, while models using call, SMS, mobile data, and transactions provide the lowest accuracy (see Figure 5).

DFS churn modeling is still a nascent field and can benefit from lessons learned in the telecommunications space, where churn models using CDR data are more established.[29] Learning what can be gained from GSM churn modeling, training models with larger datasets, and adding DFS usage data, as IFC did in Uganda, expands the potential applications and the utility of churn modeling. In Ghana, IFC also applied predictive modeling to inform customer acquisition, and one-time model use resulted in 70,000 new active mobile money users.[30]

Figure 5: Comparing accuracy of churn prediction models
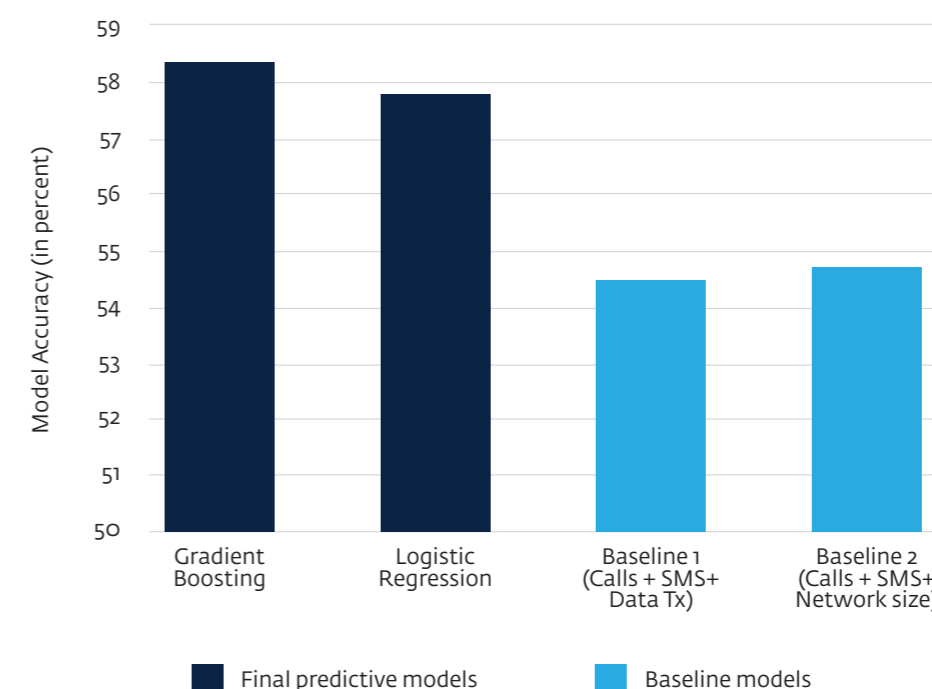
USING PREDICTIVE MODELING TO SEGMENT AND TARGET LIKELY CHURNERS: A USE CASE FROM UGANDA

Traditional segmentation methods deliver information on inactive users within a portfolio, but their operational value is limited as they inform on inactive users only ex post, or after the fact. Reactivating users who have stopped using a service is more difficult and costlier than preventing them from becoming inactive in the first place. Recognizing this, IFC developed a predictive churn model for an MNO in Uganda. This model produced churn scores that indicated the likelihood of customers leaving the service. This allowed for the segmentation of those at risk of churning, and the targeting of incentives encouraging them stay active on the platform.

IFC analyzed customer behavior data of the MNO from November 2016 to February 2017. Potential churners were identified as those who became inactive between November and December 2016. CDR and DFS data was used to predict and validate whether those individuals remained inactive throughout the 90-day period, and met the criteria for churners, unlikely to return to the service. The predictive model classified the MNO's customers correctly with 76 percent accuracy, using both CDR and DFS data. The Tigo model, which used only CDR data, achieved a much lower accuracy rating of 58 percent.

When arranged by decile, the customers at highest risk for churn had a 40 percent probability of churn, while the lowest decile had a churn probability of close to zero percent. Segmenting customers based on churn modeling identifies the customer group with the highest propensity to churn. Targeting customers in the top 20 percent
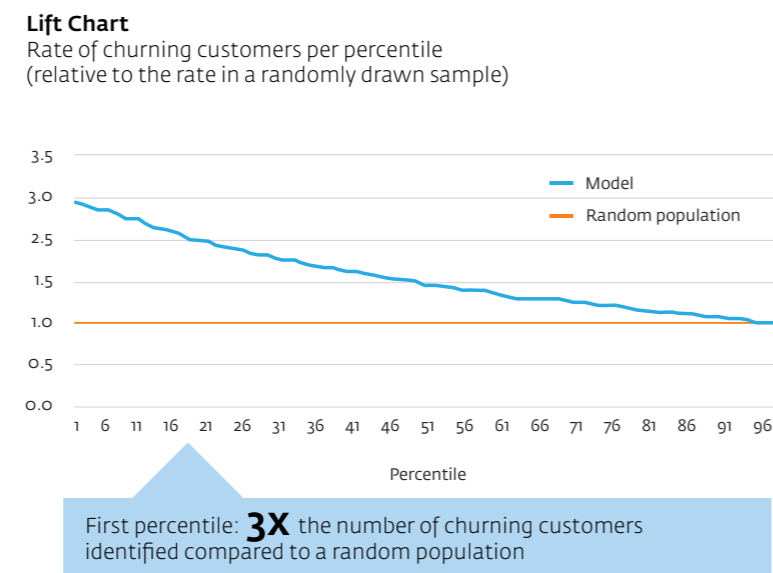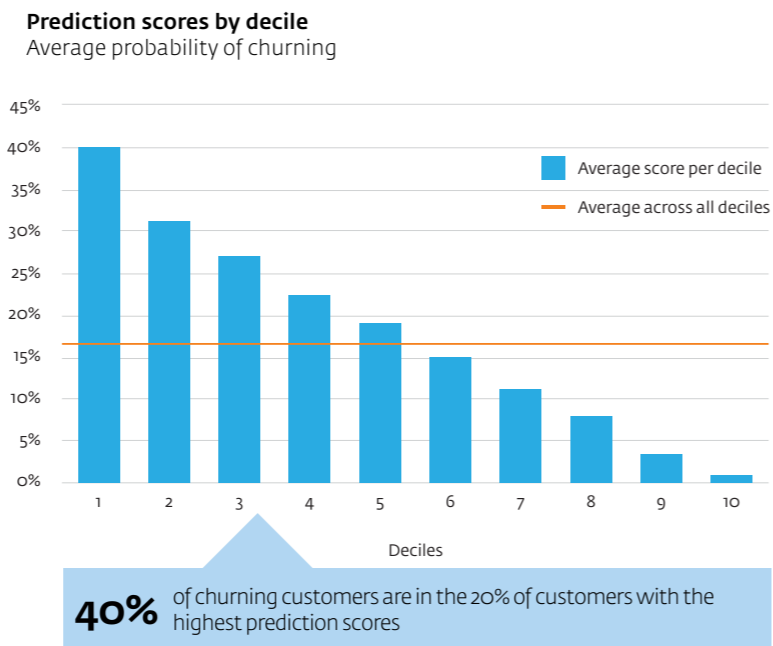
29 Blumenstock et al., 2015. *Behavioral Modeling for Churn Prediction: Early Indicators and Accurate Predictors of Custom Defection and Loyalty.*
30 IFC, 2017b. *Data Analytics and Digital Financial Services: Handbook.*

of highest churn scores can identify 2.5 to 3 times the number of churners compared to drawing randomly from the customer base. When targeting all customers in the two deciles with highest churn risk (blue columns, Figure 6), the provider can identify three times as many churning customers compared to choosing a random sample.

In order to maintain activity, providers need to target predicted churners with effective outreach methods and offers. While this research project did not explore anti-churn incentives, A|B testing could be applied to determine the most effective ways to prevent churn. For instance, a small sample of churners (e.g., 50,000 customers split up into five groups) could be presented with an offer such as a 20 percent transaction fee discount or a one-time-cash bonus, either through SMS or call. Churn rates could then be compared against a control group that did not receive any outreach or incentive, and the results used to benchmark the effectiveness of different incentive campaigns.

This highlights a key area of open research—while likely churners can be meaningfully identified, it remains unclear how this group would respond to incentives compared to new or existing active customer segments. Experimenting with a variety of retainment strategies and understanding whether these are distinct from customer activation measures may yield fruitful insights to optimize the impact of marketing campaigns aimed at churn reduction.

Figure 6. Churn campaign

Prediction scores by decile
Average probability of churning

40% of churning customers are in the 20% of customers with the highest prediction scores

Lift Chart
Rate of churning customers per percentile
(relative to the rate in a randomly drawn sample)

First percentile: **3X** the number of churning customers identified compared to a random population

CONCLUSION AND IMPLICATIONS

Growing to scale is crucial for mobile money providers to achieve profitability,[31] and accruing a large customer base is an important part of this endeavor. Yet it is increasingly clear that this is not a mere numbers game. Customers must also be active. Some past approaches to grow the customer base, such as auto-registration of MNO customers for mobile money, or low-KYC (know your customer) products for digital channels, may have boosted subscribers, but often left the provider with large numbers of inactive accounts down the road. Reducing churn and reactivating customers is important for MNOs to increase the profitability of their DFS business and may also create positive spillovers for GSM usage. For banks who usually offer fewer products and have only the banking (and not the GSM) channel, customer churn reduction may be of even greater importance. While MNO customers who churned on DFS may still be active phone users, a banking customer who ceased to be active on DFS and at branches is lost to the provider.

Churn identification techniques have made important contributions to understanding churn, by clarifying the profiles of inactive users and factors contributing to churn. Based on IFC-led surveys in Uganda, Zambia, Ghana, and Côte d'Ivoire, variables such as age, gender, education, and household-size emerged as relevant factors. Segmentation and network analysis techniques uncovered behavioral patterns of inactivity. In particular, social network size as measured by voice contacts set apart active users from inactive and non-users. Linking demographic information with information from surveys and qualitative research on why customer became inactive yielded rich insights to address root causes of inactivity. Key reasons for inactivity were identified as irregular income, no need to use DFS, and high prices. Hence, educating customers about the benefits of DFS, ensuring products are relevant to different customer demographics, and revisiting service quality and pricing were identified as pathways to improved customer retention and reactivation.

Churn prediction adds a forward-looking perspective to addressing inactivity and churn. Retaining customers tends to be less costly than acquiring new ones, and insights from churn prediction can be leveraged to grow activity from within an existing customer base. With statistical methods and machine learning techniques, phone usage and DFS data can be processed to obtain a list with the most predictive variables. IFC analyses provide proof-of-concept that churners and non-churners can be identified based on their different phone usage patterns.

Using the predictive power of these features, churn scores can be computed, and at-risk customer identified and targeted with marketing or outreach campaigns to prevent them from leaving the service. IFC data analytics projects in Ghana and Uganda provide proof of concept that this can be done. In Uganda, IFC modeling identified churners correctly in 76 percent of cases when using data drawn from both telephony and mobile money channels. In Ghana, using gradient boosting techniques with CDR data alone yielded 58 percent accuracy rates. Such accuracy rates may not seem overly compelling. However, in the context of millions of individual account holders, identifying churners with more than 50 percent accuracy is a notable improvement, equipping providers with powerful and actionable insight.

Churn modeling in DFS is still a nascent field with great potential for expansion and refinement. Utilizing insights from churn modeling to experiment with a variety of marketing campaigns, offering bonuses or fee changes, or simply reminding customers of the benefits and existence of their DFS accounts, can help providers to identify the most effective retention approaches. In particular, A|B testing allows for relatively rapid feedback loops to parse out the most effective approach.

Beyond churn, predictive modeling can be applied to mobile money adoption of MNO customers and identify highly active customers. Once a provider has built up predictive capabilities, models can easily be modified to support business decision making and drive mobile money activity and adoption. IFC's work in Ghana showcased the synergies and benefits in developing a predictive model for customer activation alongside a churn model. Predictive modeling is an important tool for providers who aim to apply cutting edge-methods to inform their actions to reduce inactivity and churn.

Neither surveys nor churn modeling are a panacea for addressing churn and inactivity, but what sets predictive modeling apart is the level of granularity of the analysis. Surveys may look at a limited number of variables to understand churn today, whereas predictive modeling can analyze hundreds of different predictors of future churn. The power of churn modeling does not replace other types of inactivity analysis, but rather it complements and enriches them with a forward-looking perspective that adds detailed insights and opportunities for proactive engagement and retention.

31 McKinsey, 2018. Mobile Money in Emerging Markets: The Business Case for Financial Inclusion.

REFERENCES

Blumenstock, Joshua, Muhammad R. Khan, Joshua Manoj, and Anikate Singh. Jun 2015. "Behavioral Modeling for Churn Prediction: Early Indicators and Accurate Predictors of Custom Defection and Loyalty." Big Data Congress. 2015. 107.

Boston Consulting Group. 2017. "How Banks Can Close the Back Door on Attrition."

Buri, Sinja, Robert Cull, Xavier Giné, Sven Harten, and Soren Heitmann. "Banking with Agents: Experimental Evidence from Senegal." Policy Research Working Paper; No. 8417. World Bank: Washington, DC.

Chamboko, Richard, Fabian Reitzug, Robert Cull, Xavier Giné, Soren Heitmann, and Morne van der Westhuizen. 2019. "The Role of Gender in Agent Banking: Evidence from Democratic Republic of Congo". Forthcoming.

CGAP. 2015. "Understanding Customer Inactivity with Customer Data from Kenya."

Demirguc-Kunt, Asli, Leora Klapper, Dorothe Singer, and Saniya Ansar. 2018. "The Global Findex Database 2017. Measuring Financial Inclusion and the Fintech Revolution." World Bank Group, Washington, D.C.

GSMA Intelligence. 2018. "Global Mobile Trends Database."

GSMA. 2017. "State of the Industry Report on Mobile Money."

Harvard Business Review. "The Value of Keeping the Right Customers." Amy Gallo, October 29, 2014.

IFC. 2016. "Find the Gap: Can Big Data Help to Increase Digital Financial Services Adoption?"

IFC. 2017a. "A Sense of Inclusion: An Ethnographic Study of the Perceptions and Attitudes to Digital Financial Services in Sub-Saharan Africa."

IFC. 2017b. "Data Analytics and Digital Financial Services: Handbook."

IFC. 2018a. "The Mobile Banking Customer that Isn't: Drivers of Digital Financial Services Inactivity in Cote d'Ivoire."

IFC. 2018b. "Banking on the Future: Youth and Digital Financial Services in Sub-Saharan Africa."

IFC. 2018c. "Women and Digital Financial Services in Sub-Saharan Africa: Understanding the Challenges and Harnessing the Opportunities."

IFC. 2019. "Do Agent Networks Boost Savings? Effects on Institutional Deposit Mobilization and Customer Saving Behavior."

IMF. 2017. "Financial Access Survey." International Monetary Fund, Washington, D.C.

Lal, Rajiv and Ishan Sachdev. 2015. "Mobile Money Services: Design and Development for Financial Inclusion." Harvard Business School Working Paper 15-083.

Mattsson, Carolina and Guy Stuart. 2019. "Understanding Key Mobile Money Users."

McKinsey. 2016. "Middle East and Africa: Telecommunications Industry at Cliff's Edge."

McKinsey. 2018. "Mobile Money in Emerging Markets: The Business Case for Financial Inclusion."

Safaricom. 2017. "News Release: Un-Audited Results Half Year Ended 30 September 2017." Safaricom Limited.

AUTHORS

Fabian Reitzug is a research consultant supporting the work of the Applied Research and Learning team. His focus is on data analytics and the impacts of gender and culture on the take-up and usage of DFS in Sub-Saharan Africa.

Soren Heitmann leads the Applied Research and Learning (ARL) team within IFC's Financial Institutions Group in the Middle East and Africa. His background is in data science, development economics, and cultural anthropology.

John Irungu Ngahu is a specialist and project lead within IFC's Financial Institutions Group in the Middle East and Africa. He has extensive DFS experience, with specialties in project design and management, agent network management, customer acquisition, and product development and innovation.

Contact the Publisher:
Devon Maylie
dmaylie@ifc.org
+27(0) 11-731-3000
IFC, Sub-Saharan Africa
14 Fricker Road, Illovo,
Johannesburg

V5 May 2020