Leveraging customer complaints data to monitor consumer protection in mobile services in Uganda

Matthew Bird, Universidad del Pacífico
Kyla Longman, Innovations for Poverty Action, Uganda
Rafe Mazer, Innovations for Poverty Action, Consumer Protection

Innovations for Poverty Action
Presented to the Uganda Communications Commission

September 2021

Contents

Figures

Tables

Key recommendations from MNO customer complaints data analysis

1. **Develop a standardized method for classifying complaints which all MNO providers will use.** The lack of full standardization of complaints categories and other variables creates limitations in comparability across providers in areas such as resolution time and concentration of complaints by products and services.

2. **Include more demographic indicators such as gender and location to segment which types of consumers may face different challenges with the products and services**.

3. **Test merging of customer care and transaction data analysis**. Where customer care logs identify common or recurring problems, transaction data on service use and sales could identify any practices of concern and ensure affected customers were refunded or otherwise compensated.

4. **Integrate complaints statistics into the periodic reporting on market trends and aggregated statistics.** Aggregated complaints volumes and key indicators like resolution time and complaints by product types could be shared periodically to help the industry measure improvements in customer service over time and address emerging issues.

5. **Expand the use of methods such as topic modeling and predictive modeling to better understand issues and target the most at-risk customers.** New machine learning methods such as topic modeling and predictive modeling helped yield insights that were not immediately identifiable through the Exploratory Data Analysis. These tools are particularly useful to properly categorize unstructured data like free text and to identify factors highly correlated with certain types of complaints or consumer issues.

6. **Consider new methods to increase the use of formal complaints channels by consumers underrepresented in the data—such as women and rural populations.** Just as UCC has rolled out a widespread awareness campaign regarding phishing scams during COVID-19, new campaigns could be designed and tested to increase consumer engagement with formal complaints channels. These campaigns should utilize segmentation and predictive analysis to improve the targeting and impact of messages on relevant populations.

1. Complaints data in consumer protection supervision

The right to redress—the ability to raise and resolve complaints—is an essential aspect of any consumer protection regime. A provider's relationship with a client rests on a contractual promise made with the consumer. Enforcing this promise requires a mechanism for making information available and a means for addressing the wrong once it is identified.[i] For consumers to trust in and benefit from products and services, they must be able to raise legitimate concerns and, where justified, receive restitution for any misconduct or errors on the part of the service provider or a third party. With a redress contract enforcement mechanism in place, the markets for the products and services are then more likely to expand.

According to the Alliance for Financial Inclusion (AFI), complaints data can serve multiple functions, including (a) identification of service improvement areas; (b) monitoring consumer protection progress and holding providers accountable; (c) informing leaders at provider organizations about service delivery opportunities; and (d) identification of areas in need of legislative and regulatory interventions.[ii] Complaints data is therefore an important element for expanding consumer protection supervision and policy formation, and regulators are increasingly making use of this data source for these purposes.[iii]

In particular, regulators have explored how to leverage digital tools to collect complaints records, monitor provider conduct, and address issues as they arise in the market.[iv] The Consumer Financial Protection Bureau (CFPB) in the United States developed a series of online complaint submission and data analysis tools which allow the regulator and general public to review complaints by volume, location, issue type and a range of other variables.[v] In the Philippines, the central bank created a chatbot which automated complaints receipt and escalation via mobile, web, and social media channels.[vi]

---

[i] Douglass North. 1990. Institutions, Institutional Change, and Economic Performance. Cambridge: Cambridge University Press. Avner Gref. 2005. "Commitment, Coercion, and Markets: The Nature and Dynamics of Institutions Supporting Exchange." In C. Menard and M. Shirley (eds.), Handbook for New Institutional Economics (Norwell, MA: Kluwer Academic): 727-786, and Avinash Dixit. 2009. "Governance Institutions and Economic Activity," American Economic Review, 99(1): 5-24.

[ii] Alliance for Financial Inclusion. 2020 "Complaint Handling in Central Bank Framework." AFI: Kuala Lumpur.

[iii] S. di Castri, M. Grasser, A. Kulenkampff. 2018. The RegTech for Regulators Accelerator (R2A) Process: Giving Financial Authorities Superpowers. BFA Global

[iv] Denise Dias. 2013. "Implementing Consumer Protection in Emerging Markets and Developing Economies" CGAP: Washington, D.C. and D.W. Arner, J. Barberis, and R.P. Buckley. 2017. FinTech, RegTech, and the Reconceptualization of Financial Regulation, 37 Northwestern Journal of International Law & Business 371

[v] See https://www.consumerfinance.gov/data-research/consumer-complaints/

[vi] See https://www.bsp.gov.ph/Pages/InclusiveFinance/ConsumerAssistanceChannelsChatbot.aspx

But more can still be done to strengthen and leverage redress systems.
Consider the example of Uganda. An Innovations for Poverty Action (IPA) survey of digital financial service (DFS) users in Uganda found that while 79% of respondents reported having experienced at least one of several common DFS challenges, only 39% of them sought to resolve these complaints through formal channels.[vii] Of those who used formal channels to address these complaints, a mere 40% said that they resolved the issue. In other words, only 12% of consumers who reported a problem managed to resolve it via a complaints and redress system. Even still, the Uganda Communications Commission (UCC) receives monthly customer care databases from all mobile network operators (MNOs), reporting on average over 200,000 formal complaints per month during the period of analysis covered in this report. With improved awareness and access to formal redress channels, and increasing use of mobile technology, the number of complaints submitted to MNOs per month only stands to expand further.

The expansion of digital complaints data increases the opportunities for consumer protection supervision and improved consumer protection measures to reduce the risks raised by mobile services and related value-added services (VAS) such as DFS. To take advantage of these opportunities, several research questions need to be addressed:

- How can the UCC – or a similar regulatory body – better leverage complaints data to inform consumer protection supervision and policy development for telecommunications, DFS, and other products operating on MNO platforms?
- What existing service areas can be improved, and what unexplored opportunities exist for providers and regulators to improve complaints handling and redress?
- How can the complaints data increase provider accountability and help better target legislative or regulatory interventions?

This report presents findings from research conducted by IPA and the UCC to develop new tools for collecting, categorizing, and analyzing complaints data from MNOs. The methods and insights achieved by the UCC and IPA offer a set of new solutions which can make better use of the increasing volume of digital consumer feedback that form an important part of consumer protection supervision for the UCC.

[vii] Matthew Bird, William Blackmon, and Rafe Mazer. 2020. “Consumer Protection Survey of Digital Financial Services Users in Uganda.” IPA: Washington, D.C.

2. Background and context

Consumer protection is a key aspect of the mandate of the UCC, as outlined in the Uganda Communications Act (2013). **As part of their regulatory requirements, MNOs in Uganda are required to submit monthly complaints reports to the UCC to monitor consumer experience and consumer protection issues.** The large volume of complaints logs the MNOs collect offers a rich database for market monitoring and the identification of areas for improvement in customer experience, product delivery, and complaints resolution across the digital economy. These new methods of data collection and analysis could similarly be used to analyze technology firms, e-commerce platforms, and DFS providers' complaints data for consumer protection monitoring.

The UCC has prioritized expanding consumer protection activities through more evidence-based interventions which leverage new data sources, such as complaints logs. To this end, the UCC and IPA conducted an analysis of 2019 and 2020 MNO complaints data (Table 1). The purpose of this research was twofold:

1. Analyze existing complaints datasets from MNOs to better understand consumers' mobile service issues and their experiences with customer care channels; and
2. Assess how complaints data can be utilized by UCC to monitor consumer protection issues across the sector and drive evidence-based policy.

Table 1 Complaints data submissions - January 2019 to August 2020

Provider	Months submitted	Customer care logs	Average complaints / month	Complaint channels (2019)	Complaints categories	
					2019	2020
MNO 1	20	688,634	34,432	4	26	143
MNO 2	17	2,159,196	127,012	36	18,576	71
MNO 3	17	733,927	43,172	9	34	25
MNO 4	3	4,107	1,369	5	20	Not received

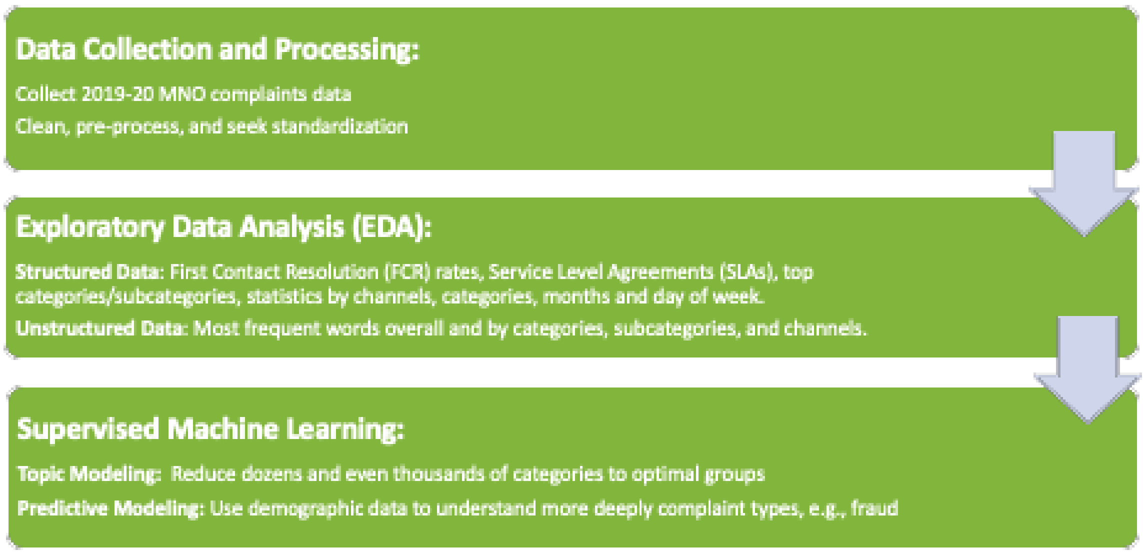
3. Methods of analysis

The objective of the analysis was to extract patterns from the MNO complaints data which the regulator could use to inform decisions on setting consumer protection policies and conducting customer care unit operations. However, the complaints records included nearly four million complaint submissions, each consisting of 10 to 15 variables—including a mix of category variables and free text—submitted in different formats and categorizations from four MNOs. This raised challenges to standardize the data and analyze quantitative and qualitative variables. The complaints also needed to be put in context, since certain population segments use MNO complaints channels more or less than the general population, leaving their population segment and complaints types overrepresented in the sample. To better contextualize the complaints logs, a nationally representative phone-based market survey was conducted with DFS users to identify biases in the MNO complaints logs.[viii]

To analyze the MNO complaints data, the research team modified a standard process for working with large datasets known as the Knowledge Discovery Process (KDD) (see Figure 1).[ix] KDD maps steps for understanding the data, selecting it to create a target data set, cleaning and

[viii] Ibid.
[ix] Usama Fayyad, Gregory Piatesky-Shapiro, and Padhraic Smyth. 1996. "From Data Mining to Knowledge Discovery in Databases." AI Magazine, Fall, 37-54

Figure 1 Modified Stages of Knowledge Discovery Process



preprocessing it, transforming and/or reducing it to identify useful features, and mining it to extract descriptive or predictive patterns.[x] **For the Uganda complaints data analysis, the modified KDD approach consisted of three main steps: (i) data collection and processing, (ii) Exploratory Data Analysis (EDA), and (iii) supervised machine learning for deeper understanding**. These steps should be seen as an iterative as opposed to linear process and are described in detail below.

3.1. Data collection and processing

The first task is to establish a basic understanding of the data. What are the data characteristics, how were the complaint registers generated, and for what goals were the data collected? The complaints data submitted by MNOs did not follow standard formats and thus required cleaning to standardize the databases as much as possible across MNOs.

For example, a critical difference between MNOs' databases was the multiple types of categories and sub-categories of complaints used by each MNO. These ranged from 20 categories in one MNOs database to more than 18,000 for another MNO, with the latter due to lack of standardization, leading to many identical categories with minor spelling or other differences, which hindered accurate measurement of complaints across categories.

Before measuring the types of complaints consumers report to MNOs, these categories had to be standardized as much as possible across MNOs. The research team aggregated these categories into smaller sets beginning with addressing such

[x] A similar but alternative approach is the Cross-Industry Standard Process for Data Mining (CRISP-DM), which iterates between different phases of analysis: (a) business understanding, (b) data understanding, (c) data preparation, (d) modeling, (e) evaluation, and (f) deployment.

issues as case sensitive categories and punctuation. **For future complaints reporting and analysis, harmonizing categories across providers would help to address this challenge and facilitate the analysis going forward.** Once the data were collected and processed, the team began to identify useful patterns of complaints handling through the EDA process.

descriptive and summary statistics from the complaints data provided performance metrics and indicators which could be used for ongoing supervision and measurement of quality of customer care by UCC. **From the EDA in this project, the research team identified the following key performance metrics, summarized here and detailed in Section 4.**

a. **Volumes of complaints across providers.** This compares volumes of complaints between providers over the same periods of time, which allows us to see if providers are receiving similar levels of complaints. If they are not, further inquiry would be merited to explain the variation. In this case, it is clear that at least part of the variation is due to providers' holding different understandings of complaints definitions and reporting requirements, such as some providers not including all customer care inquiries in their category of what is considered a complaint.

b. **Complaints submission by channels**. Some MNOs track the channels customers use to engage customer care, which allows for analysis of how many consumers use different channels such

3.2. Exploratory Data Analysis (EDA)

The EDA analysis served two purposes. First, it was an intermediate diagnostic step for testing assumptions and formulating hypotheses about the complaints data, which could then be examined via data mining, while iteratively identifying data components to continue to clean and code for subsequent analysis. Second, the

as a customer hotline or physical service centers. Understanding the channels used by consumers to register complaints can signal what channels are most relevant for different segments and help to set strategies for outreach and engagement that use consumers' most preferred channels. Unsurprisingly, our analysis found that virtually all complaints came through MNO call centers, which would be interesting to track over time to see how this shifts as social media increases in use beyond younger, urban populations.

c. **Most common complaint categories**. As noted previously, sorting data by complaints categories was an important step to harmonizing complaints databases across each MNO. Once categories are relatively similar, they can then be used to identify what types of issues were raised most often, and how these vary across providers. Variations in most frequent complaint categories across providers can signal areas where individual providers or the industry could improve service delivery. For instance, being 'unable to pay bill' was the most common mobile money complaint for one provider, while this did not appear often in the other provider's mobile money complaints. As these providers offer the same types of services for paying bills with mobile money, this suggests that one provider is facing greater issues than the other in this area and warrants further investigation and possible improvements in bill payment services.

d. **Time series analysis**. By looking at complaints volumes across nearly 20 months, the research team identified shifts in volumes of complaints by different topics or issues. For example, we were able to document how complaints volumes decreased at the beginning of the COVID-19 pandemic, likely due to both a fall in new mobile subscriptions and reduced staffing at customer care centers.[xi]

e. **Complaints resolution rates.** Complaints resolution focuses on three metrics: i. First Contact Resolution (FCR) rates (what portion of complaints were resolved on the first contact); ii. overall rates of resolution; and iii. resolution times (how long it took to resolve a complaint). These rates are analyzed by provider, complaint type, subscription-type, and complaint channel, and can be used as benchmarks for quality of complaints handling, to ensure systems improve over time in their ability to address consumer concerns.

f. **Consumer segmentation**. The addition of gender, age, and location data in the MNOs' 2020 complaints data yielded important insights on who does—and does not—access formal complaints channels. For example, complaints by female customers were only 35% of the overall volume, despite female subscribers accounting for 45% of total subscribers for MNOs. This raises questions as to whether female subscribers are less likely to experience challenges or are less likely to raise these challenges to customer care when they occur. By adding in demographics to the complaints reporting requirements, UCC and other regulators can better target different consumer segments to increase their use of complaints channels or support them if they are disproportionately affected by certain problems.

Additional EDA was conducted with unstructured text data, including analysis of most frequent keywords overall and by categories, subcategories, and channels. These initial patterns helped generate hypotheses about different forms of classification and suggested that the application of supervised machine learning techniques may reveal deeper insights.

[xi] New mobile subscriptions fell by ~3 million from Q1 to Q2 in 2020, coinciding with the start of the pandemic. By the end of 2020, this had returned to pre-COVID-19 levels. Uganda Communications Commissions, Market Performance Report 4Q2020. https://www.ucc.co.ug/wp-content/uploads/2021/04/UCC-Q4-2020-Market-Perfomance-Report-compressed.pdf

3.3. Machine learning

The EDA provided useful metrics and indicators that form the foundation for tracking complaints handling by MNOs, improving customer care practices, and targeting different consumer segments. These insights also helped to generate hypotheses that can be tested via supervised machine learning algorithms. Machine learning is a field of study where algorithms and other techniques are developed, which train machines to learn how to perform tasks of analysis over time.

In this research, machine learning tools were tested to develop hypotheses regarding how complaints relate to each other across the sample, and what traits may be predictive of consumers experiencing certain types of complaints. For example, patterns seen in word clouds developed during the EDA process led to hypotheses about new types of customer complaint categories. Descriptive data on call volume and complaint types over time also helped formulate hypotheses about client characteristics and complaint types at different times of the day, week, or month. **By applying supervised machine learning algorithms, the research team was able to deepen the insights from the EDA phase of analysis using two primary approaches: Topic modeling and predictive analytics.**

a. **Topic modeling.** Topic modeling is a Natural Language Processing (NLP) technique used to extract latent themes from a text, based on keywords and their combinations. The 2019 MNO complaints data contained unstructured data or text—such as notes from call center staff describing the complaint interaction and measure taken—as well as the categories of complaints and outcomes. Unstructured text provides the opportunity to identify empirically the most common themes or topics using text descriptions from the customer care personnel, instead of providers categorization. In this case, the complaints data from 2019 were analyzed using a topic modeling algorithm to identify the most common topics based on their description.

   Topic modeling analysis led to substantial category reductions for each MNO, as the text analysis was able to discover similar complaints that may have been placed in different categories by the call center staff. Further refinement of this work could result in additional patterns in complaint types and improvement in category reporting, benefitting both the UCC and MNOs.

b. **Predictive modeling.** The research team used supervised machine learning algorithms to identify characteristics of the complainant most predictive of the complaint type presented. The purpose of this analysis is to identify whether complaints about certain problems were associated with characteristics of the complainant or channel in which the complaint was presented.

   Based on the 2020 MNO complaints data (which included customer demographic data including gender, age, and district of residence) the research team was able to see which individual consumer characteristics were most likely to predict the kinds of problems presented to customer care. For example, predictive modeling was used to identify the client characteristics that make them most likely to present a complaint related to fraud, which are: The gender of the person calling, where the call is coming from, the time of day, day of the week, and month of the year. **Predictive modeling analysis combined with demographic data could help supervisors to identify and provide preventive measures to support demographic segments which suffer disproportionately from a particular issue or challenge, and then, for example, develop specific anti-fraud messaging targeting these populations.**

3.4. Sequence of complaints data analysis methods

Researchers, regulators, and financial service providers seeking to leverage complaints data can consider using both EDA and machine learning methods when determining their plan for analysis. The methods have been ordered above from least to increasingly complex, with an incremental approach as an appropriate starting point (See Table 2).

For example, a regulator could begin with a review of the current categories used in customer care logs and work with providers to simplify and harmonize these categories. From there, the harmonized data could be used to conduct an EDA, which will yield the key performance indicators such as resolution rates, turnaround times, and complaints volume by product and service type. These indicators alone would be sufficient to begin complaints-based supervision. Better still, if the data were collected on a real-time or next-day basis, they could flag problematic issues as they arise and allow for immediate corrective actions to be taken, instead of waiting for monthly or quarterly reporting. Complaint data analysis insights may then be used to inform real-time preventive consumer protection actions.

Moving beyond this descriptive analysis, use of supervised machine learning analysis—including topic modeling and predictive modeling—could be a longer-term aspiration for complaints data monitoring. Compared to EDA, these methods require more investment in data generation, database management, and data science analysis. However, this investment may be worthwhile where machine learning can improve complaints categorization or predictive modeling of vulnerable consumer segments, as seen with the Ugandan MNO complaints data.

Table 2 Stages of complaints data analysis employed by IPA & UCC

Method of analysis	Primary objective of analysis	Data types	Examples	Level of difficulty
Data collection and processing	Ensure quality and consistency of complaints data across providers and channels.	Structured data are typically numerical or categorical. Unstructured data may include text, images, audio.	Categories of complaint channels (e.g., call center, store, social media). Numerical measure of time to resolution. Free text notes or description of the complaint and action taken.	Low. Basic experience in data cleaning.
Exploratory Data Analysis (EDA)	Build key performance indicators to determine most common types of complaints and assess quality of complaints handling.	Structured Data: Organized categorical and numerical variables Unstructured Data: Text	Visualization techniques for structured data may include line graphs, time-series graphs, histograms, box plots, and scatter plots. Visualization for unstructured text may include word clouds and n-gram lists.	Moderate. Requires familiarity with Stata or similar such software to properly execute.
Machine Learning	Identify insights for improving complaint categorization or anticipate complaint types to inform prevention efforts.	*Topic Modeling:* Unstructured data *Predictive Modeling:* Structured and unstructured data	Topic Modeling Predictive Modeling	High. Requires familiarity with programs, such as R or Python, and use of NLP libraries using these scripted programming languages.

4. Creating Consumer Protection Indicators via EDA of Complaints Data

The exploratory data analysis of Ugandan MNO complaints logs demonstrates how supervisors and service providers could better leverage this data to monitor and improve customer care, and in turn consumer experience. Even with non-standardized data from the 2019 and 2020 MNO customer care logs, IPA and UCC were still able to measure key performance indicators which could be used as benchmarks for consumer protection supervision and monitoring going forward.

From this analysis, **we developed six types of indicators, which could be prioritized as the foundation of complaints-based supervision for the UCC and similar regulators overseeing MNOs:**

1) Complaints volumes by providers
2) Complaints volumes by channels
3) Distribution of complaints across product and issue categories
4) Complaints variations over time
5) Complaints resolution times
6) Demographic segmentation of complainants

4.1. Complaints volumes by providers highlights differences in complaints categorization

Holding other factors equal, one would expect that a providers' volume of complaints would be proportionate to their market share. For example, if a provider has around 70% of the market, we would expect their complaints submitted to UCC for a given month would total around 70% of the total complaints the UCC received that month. However, in the 20-month sample considered in this study, this was not the case, as shown in Table 3.

Despite relatively similar market shares between MNO 1 (51%) and MNO 2 (44%), MNO 2 submitted more than three times the number of complaints per month to UCC. MNO 3—despite having only 5% of the market—submitted 21% of the average monthly complaints received by UCC. MNO 1 is extremely underrepresented in the data, while MNO 2 and MNO 3 are overrepresented.

Table 3 Complaints volumes by MNO compared to market share of subscriptions (2020)

Provider	Average # of complaints per month	Provider's market share	Provider's share of complaints data
MNO 1	34,432	51%	16.8%
MNO 2	127,012	44%	62.1%
MNO 3	43,172	5%	21.1%

There are several possible explanations for this variation:

Hypothesis 1: MNO 1 is simply a better service provider than the other two MNOs, resulting in fewer customer issues and thus fewer customer complaints.

Hypothesis 2: Customers across the MNOs face similar levels of issues, but MNO 2 and MNO 3 have more accessible customer care channels and thus customers can better express their complaints.

Hypothesis 3: Customers across the MNOs face similar levels of issues, but MNO 2 and MNO 3 are more reactive to customer care complaints than MNO 1. Perhaps customers at MNO 1 have stopped expressing their complaints because they believe that when they raise complaints, their issues are not addressed.

Hypothesis 4: MNOs have different understandings of what constitutes a 'complaint' and therefore what is supposed to be included in their complaints submission to UCC. For example, it could be that MNO 1 filters out customer care inquiries that do not result in a complaints investigation or any follow-up action.

By triangulating our complaints data with other data sources, we sought to confirm or refute these possible explanations. IPA's 2020 survey of DFS users in Uganda found that:

- DFS users across the providers were just as likely to face common DFS customer care issues.[xii] This does not support hypothesis 1 (that MNO 1 is simply a better service provider).

- DFS users across the providers who faced issues were just as likely to contact providers to resolve these issues.[xiii] This does not support hypothesis 2 (that customers of MNO 1 lack accessible complaints channels compared to MNO 2) and hypothesis 3 (that customers of MNO 1 face similar levels of issues but are less likely to report them compared to customers of MNO 2).

- However, DFS Users from MNO 2 were more likely to report that when they reached out to their provider to solve their problem their issue was successfully resolved compared to MNO 1.[xiv] If in the future once templates and guidelines are standardized, we continue to see low levels of MNO 1's

[xii] For context, these were 79% of MNO1 and MNO2 users. Matthew Bird, and Rafe Mazer. 2020. "Consumer Protection Survey of Digital Financial Services Users in Uganda." IPA: Washington, D.C.

[xiii] For context, these were 17.6% of MNO1 and 16.3% of MNO2 users. This difference is not statistically significant. Matthew Bird, and Rafe Mazer. 2020. "Consumer Protection Survey of Digital Financial Services Users in Uganda." IPA: Washington, D.C.

[xiv] For MNO2, 47% of respondents who reached out to the provider to solve their issue reported that the issue was resolved, compared to 32% of respondents for MNO1. This difference was not statistically significant at conventional levels (p=.13). Matthew Bird, and Rafe Mazer. 2020. "Consumer Protection Survey of Digital Financial Services Users in Uganda." IPA: Washington, D.C.

customer complaints volumes compared to MNO 2, then it would be worth considering hypothesis 3 (that customers at MNO 1 have stopped expressing their complaints due to perceived inaction in resolving complaints) as a potential explanation for the difference in volumes of complaints by providers. If MNO 1's complaints volumes decrease as time goes on, this would also support this hypothesis, as more consumers may be learning that it is not worth raising complaints because they believe the MNO is not addressing the issues.

The 2020 IPA survey findings suggest that hypotheses 1-2 are not likely the cause for the discrepancy we found in providers complaints data, and hypothesis 3 would require further data to determine. **Hypothesis 4 is the most likely candidate: Perhaps providers have different understanding of what classifies as a complaint and what should be reported to UCC.** Indeed, MNO 3's 2019 data include a variable indicating that 39% of their submitted customer care logs are 'inquires', 35% are 'complaints', and 26% are 'others.' The other MNOs' datasets do not include this variable, so it is unclear if they are including inquiries in their submissions or excluding them, and how exactly they make this distinction.

Providers' understandings of complaints and reporting requirements play a role in why some MNOs report different numbers of complaints. The provision of clear guidelines for the type of MNO customer care logs submitted each month and a clear explanation of what should be classified a complaint versus an inquiry should help solve this issue. **UCC should consider not only standardizing the complaints reporting templates, but also work with MNOs to develop consistent definitions of what the threshold is for complaints to be logged and submitted to UCC, as well as any exclusion criteria of general inquiries from the reporting requirements.**

4.2. Call center accounts for nearly all complaints by channel used

For each of the providers, between 97 to 99% of all complaints came from MNO call centers. Complaints expressed by social media or in-person represented less than 3% of complaints. One explanation is that these alternative channels are underutilized across all the providers. Different consumer segments might be more likely to express complaints through particular channels, e.g. WhatsApp or Twitter, since users of these channels are more likely to be urban, younger, and have higher incomes on average.[xv] Thus, if these channels are not made easily accessible and complaints reported through those channels are not responded to in a timely manner, the

[xv] For an analysis of social media usage in the context of consumer protection in Uganda see: https://www.poverty-action.org/publication/social-media-usage-digital-finance-consumers-analysis-consumer-complaints

complaints experiences of particular consumer segments may be missing. **Improvements to the accessibility of alternative channels and the timely resolution of complaints raised through them may provide consumers more preferred options for complaints handling and increase overall capture and resolution of complaints.**

4.3. Complaints across product and issue categories are highly varied

In the four MNOs reviewed, there were considerable differences regarding which issues were most commonly raised to customer care. Table 4 provides an example, listing top customer issues raised by providers in the 2020 data.

For two of the providers, mobile money was the main area for complaints. MNO 3, despite having a mobile money product, did not receive many complaints in this area, with the bulk of the complaints focused on traditional telecommunication services (data and voice calls). This suggests that customers of this MNO are not utilizing its mobile money product as much as their counterparts at MNO 1 and MNO 2 are. This finding coincides with IPA's 2020 survey, which found that while 17% of respondents used this provider for voice calls, only 1% had ever used them for mobile money.

With both MNO 1 and MNO 2, we find that mobile money-related complaints tend to concentrate on one or two primary issues, but the issues vary between the MNOs. For MNO 2, the inability to pay bills was the most common mobile money issue, accounting for more than half of mobile money complaints. Whereas, for MNO 1, only 2% of mobile money complaints related to being unable to pay bills.[xvi] The types of bills one can pay using mobile money are identical across providers (e.g. school fees, electricity, solar power, water, television). Thus, if users from one MNO are having more issues using mobile money to pay their bills than users of the other MNO, it could suggest problems with that MNO's bill payment services that may warrant further investigation into this matter to determine the causes.

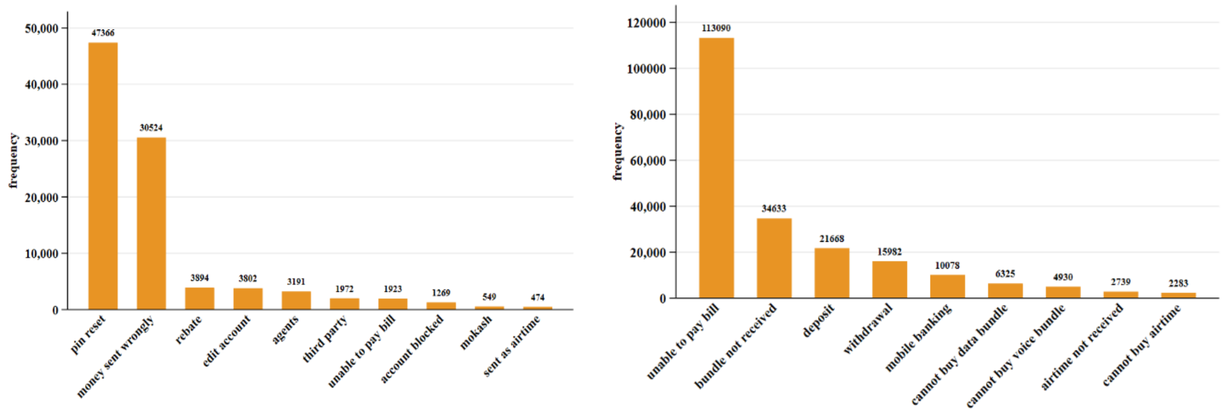
Using the sub-categories from the provider data, we can dive deeper into mobile money complaints for the two dominant mobile money providers (Figure 2).

Table 4 Most common customer issues in 2020 complaints data

MNO 1	1) Mobile money 2) Lost airtime 3) Blocked sim	MNO 2	1) Mobile money 2) Platform 3) Data	MNO 3	1) Mobile internet 2) Failed calls 3) Account info and modification

[xvi] Note, being 'unable to pay bills' refers to there being an issue with the mobile money mechanisms which prevented the money transfer from going through, not that customers did not have sufficient funds to pay their bills.

Figure 2 Mobile money complaints by issue type 2020 (MNOs 1 and 2)



4.4. Monthly variations in complaints volumes signal emerging customer concerns

Tracking complaints over time enables identification of shifts in consumer protection issues and provider responsiveness across MNO products and services. **For providers, regulators, and consumer rights organizations, setting up a periodic review of complaints data can help to flag concerns with the effectiveness of complaints channels as well as new issues increasing in incidence that may be addressed with policy reforms.** One striking finding from a time series analysis was the shift in complaints volumes during the COVID-19 pandemic. Shifts in complaints data revealed both how providers managed their customer care at the start of the pandemic, as well as how the crisis impacted issue prevalence, such as consumer-affecting fraud.

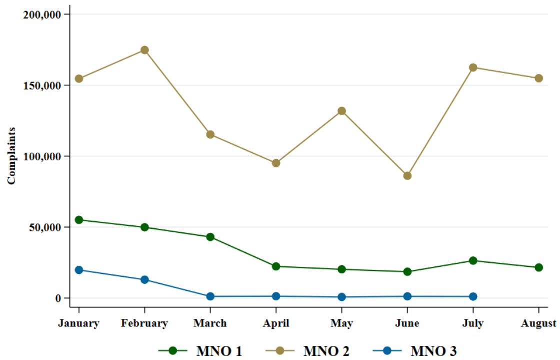
Tracking complaints volumes

Complaints volumes declined between March and June 2020 (Figure 3), which coincides with Uganda's most restrictive COVID-19 lockdown period. This may seem counterintuitive, since the use of mobile and mobile financial services increased across the country during this time. IPA's 2020 survey found that since COVID-19 begin, 48% of respondents reported using mobile money more than before.[xvii] Furthermore 55% of respondents said they had transitioned to mobile money use for transactions previously done with cash.[xviii]

[xvii] 16% of respondents use mobile money as the same level as before, and 36% use it less often. Matthew Bird, and Rafe Mazer. 2020. "Consumer Protection Survey of Digital Financial Services Users in Uganda." IPA: Washington, D.C.

[xviii] Matthew Bird, and Rafe Mazer. 2020. "Consumer Protection Survey of Digital Financial Services Users in Uganda." IPA: Washington, D.C.

Figure 3 Complaints volumes - January - August 2020



Further challenging the decline in customer care volumes at the outset of COVID-19, a parallel IPA study of social media data in Ugandan financial services from July 2019 through June 2020 found an increase in outreach to customer care on Twitter and Facebook after the World Health Organization declared COVID-19 a pandemic on March 11.[xix] **This implies that complaints likely did not decrease during the pandemic, rather providers were not responding to as many customer care inquiries via their customer care channels, perhaps due to staffing issues during the pandemic, resulting in more unresolved issues for consumers.** Considering the slower resolution time observed for complaints raised via social media, there could be a backlog of unresolved issues from the COVID-19 pandemic that are worth examining through a combination of resolution rates data from call center records and response rates to customer complaints raised via social media.

Monitoring fraud issues during COVID-19

While all the MNOs saw a large reduction in complaints during the most severe Covid lockdown period, only one of the three MNOs had since returned to its pre-pandemic level of monthly complaints volumes by the end of the data sample. By July 2020, MNO 2 was back to its pre-Covid volume, signaling that this provider had resolved whatever issue was driving the lack of customer care responsiveness and reestablished its regular level of customer care service.

By contrast, as of August 2020, MNO 1 and MNO 3 had still not returned to their pre-COVID 19 levels of complaints volumes. From a monthly average of 37,990 before COVID-19, MNO 1 averaged 21,793 complaints a month from March – August 2020, a 43% reduction. MNO 3 experienced an even greater drop off, from a monthly average of 56,124 before COVID-19 to 1,080 in August, a 98% reduction. This may signal a gap in customer complaints handling during the pandemic and calls for further inquiry as to the causes.

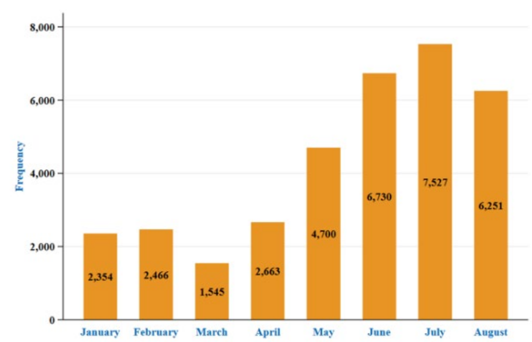
[xix] Melissa Tully and Dani Madrid-Morales. 2020. "Social media as a tool for consumer protection monitoring." Innovations for Poverty Action: Washington, D.C. https://www.poverty-action.org/sites/default/files/Social%20Media%20as%20a%20Tool%20for%20Consumer%20Protection%20Monitoring_circulation.pdf

During the pandemic, fraudsters took advantage of mobile subscribers through phishing calls impersonating the government or customer care staff. IPA's consumer protection survey in Uganda found that 46% of respondents had received a scam call since COVID-19 began, with 49% of those cases involving the fraudster impersonating MNO customer care staff.[xx] At the same time, the UCC sought to raise consumer awareness of these fraud risks and how consumers can protect themselves through messaging on traditional and social media.[xxi]

While the overall number of complaints received per month fell with the onset of COVID-19, the number of fraud complaints rose dramatically for one MNO in our sample. From April to May, the volume of fraud complaints increased by 76% for this provider (Figure 4). From May to June, it this closely and consider whether fraud prevention methods are needed to combat this trend of COVID-19 fraud cases. jumped again by 43%, and another 11% from June to July. UCC will want to monitor

Complaints data from the other two MNOs did not reflect this same fraud trend; however, the research team believes this may be a result of an incompatibility in how these two MNOs are categorizing and reporting fraud complaints rather than a lack of fraud complaints per se. For instance, for MNO 3 in 2020, only 76 complaints out of 38,245 (0.2%) mentioned the word 'fraud' anywhere in the complaint entry. While it is possible that this reflects the reality – given that MNO 3 relies less on mobile money services and thus may be less likely to receive any fraud-related complaints – such a low rate is still unlikely. **Adoption of a standardized method for classifying complaints across MNOs will address this issue and allow UCC to understand whether instances of fraud trends are common across providers or unique to a particular provider.**

Figure 4 Volume of fraud complaints MNOs



[xx] Matthew Bird, and Rafe Mazer. 2020. "Consumer Protection Survey of Digital Financial Services Users in Uganda." IPA: Washington, D.C.

[xxi] https://twitter.com/UCC_Official/status/1271707277737832456

4.5. Complaints resolution times depend on channel and issue

Resolution time is a key benchmark for monitoring the quality of complaints handling mechanisms. Consumers should be able to expect swift and effective resolution of complaints they present to service providers. Without timely resolution, consumers may suffer financial loss, or decrease product usage. The IPA survey of DFS users found that those who had an unresolved problem with their DFS products were more likely to report having stopped using that DFS product, reduced usage of the DFS product, or switched providers compared to DFS users whose complaints were resolved.

In Uganda complaints data from MNOs included complaint status—e.g. resolved or unresolved—and the time and date of complaints submission, which enabled construction of complaints resolution indicators for rate and time of resolution.

Table 5 depicts how many complaints were resolved in first contact in 2020 for a provider, and how long it took on average to resolve complaints that were not handled in a single contact. These are the types of benchmarks on resolution which UCC, and other supervisors could track over time to monitor industry performance. **By monitoring resolution times and resolution rates over time, it will be possible to track whether customer care standards are improving or not. Where regulators and providers have put in place maximum turnaround time policies for complaints handling, these indicators could be used to assess how well the industry meets turnaround time requirements and ensure continuous improvements in resolution times over months and years.**

Table 5 Complaints resolution times (2020)

	% complaints resolved in first contact	Average resolution time not resolved in first contact
MNO 1	84%	47 hours
MNO 2	65%	16 hours
MNO 3	9%	44 hours
MNO 4	Not provided	Not provided

Resolution times across providers

There was a large variation in the resolution times across providers, suggesting that either some MNOs are quicker at resolving complaints, or that the MNOs do not share the same understanding and approach to marking complaints as resolved. With a standardized reporting template and clear guidelines for logging complaints, UCC will be able to monitor resolution times to identify issues and areas for improvement in complaints handling.

Resolution time by complaint category

The research team wanted to understand how resolution rates may be affected by the characteristics of each complaint. **Focusing only on the overall MNO rates of resolution and turnaround time can mask significant variation in resolution rates across different products and complaints channels. The research team thus analyzed the resolution rates against several variables captured in each complaint log, including the type of issue raised and the channel the complaint was raised on**.

Figure 5 shows how resolution rates vary considerably by type of issue or complaint channel. In this case, the resolution time analysis separates between those issues resolved within 15 minutes, and those which take more time. A higher percent of complaints resolved within 15 minutes is not always associated with lower median resolution time for complaints not resolved within 15 minutes. This may mean that some types of issues are generally easy to address, but those that take longer vary greatly in resolution time.

For example, 97% of fraud complaints are resolved within 15 minutes, but those that take longer are resolved in an average of 24 hours. Unfortunately, the complaints dataset does not provide any detail on the nature of these fraud complaints, so it is not clear why some are resolved so quickly, while others take multiple days to resolve—although severity or complexity of the issue is one probable explanation. In the new template UCC plans to issue, fraud complaints will be categorized according to the relevant product. With this added information, UCC will be able to cross-reference resolution times with categories and subcategories to monitor which types of issues are taking the longest to resolve.

Resolution time by complaint subcategory

Crossing the mobile money complaints subcategories with resolution times shows how resolution times can vary greatly even within a particular complaint category. Figure 6 demonstrates this using the mobile money subcategories for one MNO's data.

'Unable to pay bill' was the most common mobile money complaint for this MNO in 2019, representing 56% of the total complaints. Yet, these complaints are resolved relatively quickly: 83% are resolved within 15 minutes, and only 3% take more than 24 hours to resolve. By contrast, the second most common type of mobile money complaint, 'Data not received but charged', which represents 22% of mobile money complaints, takes much larger on average to resolve. Only 42% of these complaints are solved within 15 minutes, and 43% take more than 24 hours to resolve.

Figure 5 Speed of resolution by issue experienced, MNO 2 (2019)

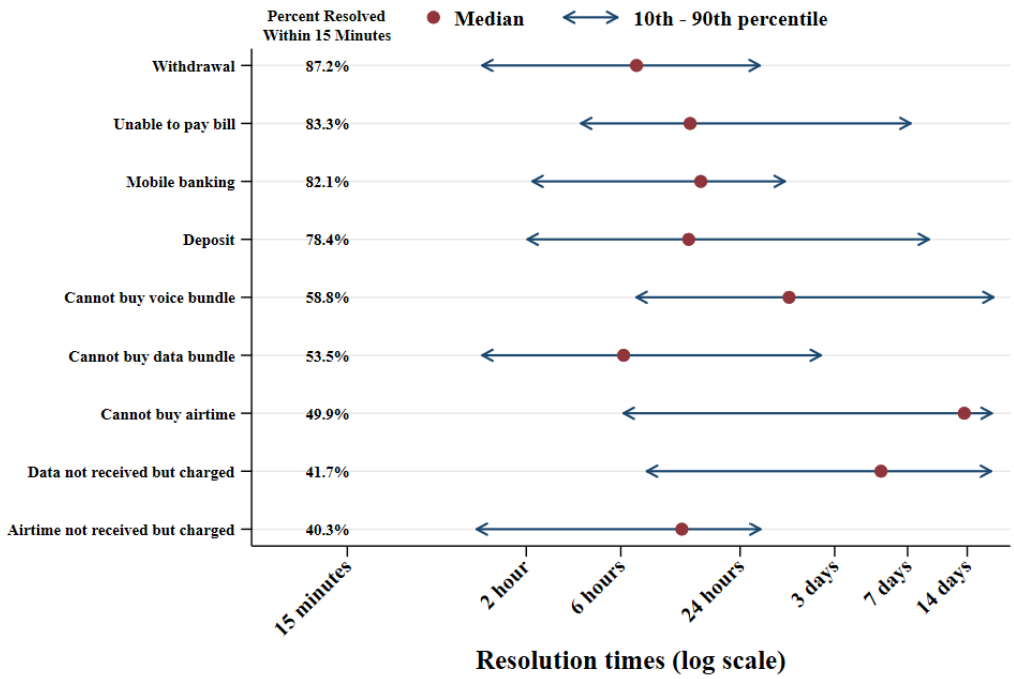
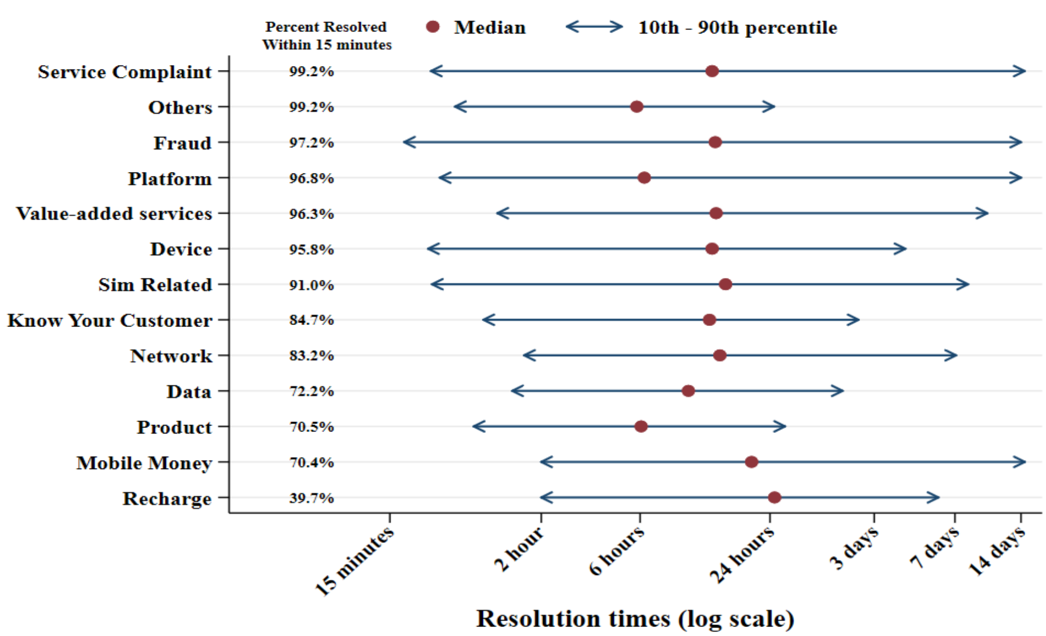


Figure 6 Resolution times by complaints category



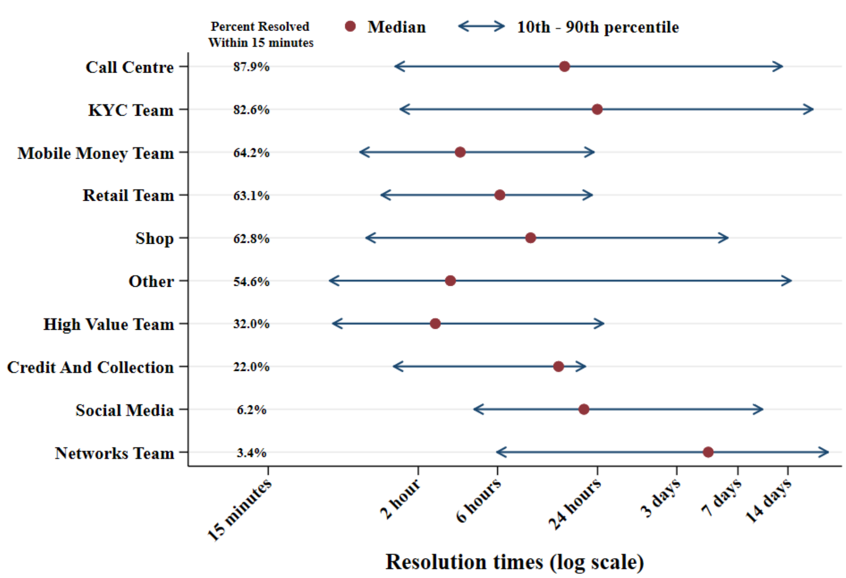
Resolution time by complaint channel

Figure 7 uses the same provider's 2019 dataset to evaluate how fast complaints are resolved depending on the channel through which the customer presents the complaint.

As seen in Figure 7, the speed of resolution varies by channel. For example, when complaints are handled by the Know-your-customer (KYC) team, 83% of complaints are resolved within 15 minutes. By contrast, only 6% of complaints presented via social media are resolved in 15 minutes. This may indicate a lack of sufficient investment in social media responsiveness by the customer care team, which has become an increasingly important complaints resolution channel. The Mobile Money Team has a relatively high 15-minute resolution rate, as well as a low median resolution. This appears to contradict the findings from Figure 5 that mobile money issues were less likely to be resolved within 15 minutes overall, which may mean that some questions best directed to mobile money operations are instead directed to other channels which are less equipped to handle these types of issues.

As this MNO example shows, **reviewing resolution time by category and by complaints channel can allow regulators like the UCC to drill down further on which types of issues do—and do not—get resolved in a timely manner. Over time, these data can be used to set and track benchmarks for resolution time by issues, channels and outcomes—e.g. customer refunded/transaction cancelled, request denied, referred to further investigation—to monitor effectiveness of customer care on a monthly basis.**

Figure 7 Speed of resolution by complaint channel used, MNO 2 (2019)



4.6. Demographic segmentation by gender, location, and age

To help identify potential biases in the representation of consumer segments in the complaints data, UCC amended the 2020 complaints data template to include the demographic variables of gender, age and location of SIM registration for each complaint. This analysis revealed several biases towards certain consumer segments. For example, in 2020, women represented 45% of all subscribers, yet only between 30% and 34% of customer care complaints in the datasets belong to accounts registered to women.

Location

Complaints are coming mainly from users in urban areas in the Central region of Uganda.[xxii] For example, for one of the MNOs nearly half the complainants are from Kampala (42%), followed by Wakiso (7%), Mukono (1.75%), and Mbarara (1.5%).[xxiii]

Subscriber figures by district were not available for this study, so it is not clear how these complaints proportions by district compare to the market share of mobile phone users by district. While exact market shares are not known, IPA's 2020 DFS consumer protection survey supports the idea that the large representation of Kampala in the complaints data is proportionate to its market share. In this survey, mobile users in Kampala were just as likely to face mobile challenges as users outside of Kampala.[xxiv] In addition, users in Kampala who faced these types of challenges were just as likely to report them to the provider as users outside of Kampala.[xxv] **Still, it would be beneficial for the UCC to utilize the subscriber figures by district from the MNOs for future complaints data monitoring and analysis. Combining subscribers by district with complaints data by district would allow the UCC to identify the existence of hotspots or determine if users in different districts are facing distinct types of issues by weighting the portion of complaints against the overall subscribers in that district.**

[xxii] Only one of the MNOs reported location of SIM registered data in a way that could be used for analysis.
[xxiii] Kampala, Wakiso, and Mukono are urban centers located next to each other in the Central region. Mbarara is a district in Western region and contains Uganda's fourth largest city.
[xxiv] 80% of Kampala respondents reported that they have faced a DFS challenge (114 out of 143 respondents). 76% of non-Kampala respondents reported that they have faced a DFS challenge (500 out of 659 respondents). The difference between these two figures was not statistically significant (p=.277). Matthew Bird, and Rafe Mazer. 2020. "Consumer Protection Survey of Digital Financial Services Users in Uganda." IPA: Washington, D.C.
[xxv] 16% of Kampala respondents who have faced a DFS challenge reported it to the provider (18 out of 111). 14% of non-Kampala respondents who have faced a DFS challenge reported it to the provider (75 out of 523). The difference between these two figures was not statistically significant (p=.612). Matthew Bird, and Rafe Mazer. 2020. "Consumer Protection Survey of Digital Financial Services Users in Uganda." IPA: Washington, D.C.

Age

In comparison to the broader country population,[xxvi] the complainants in the 2019-2020 MNO data skew younger. However, this may also reflect that in general mobile subscribers tend to skew younger. We did not have access to the Uganda mobile subscriber age figures for this study, so it unclear if this bias is genuine.

Going forward, through comparison of subscriber ages within the complaints data and the overall subscriber data from MNOs, UCC could monitor on a regular basis several important aspects of the complaints and redress process, including answers to questions such as, “Are consumer of different age groups facing more or different types of issues?” and “Are consumers of different age groups using different channels to report their complaints?” If the answer to both questions is yes, then providers could consider targeted methods to address complaints depending on who is complaining where about what types of issues. **With improved collection of subscription figures by gender, location, and age, UCC would be able to benchmark these demographic splits in the complaints data against the overall subscribers in telecommunications, mobile money, and other services as needed.**

[xxvi] Age in graph is from Uganda Bureau of Statistics projection for 2020. Made in 2018. https://www.ubos.org/wp-content/uploads/statistics/Population_Projections_2018.xlsx

5. Expanding Complaints Data Insights via Machine Learning

To expand the descriptive data analysis, the research team further explored the application of topic modeling and predictive modeling techniques within the 2019-2020 MNO complaints databases. These machine learning techniques were used for two test cases which sought to explore their relevance for complaints data analysis: (a) Topic modeling of complaint categories to improve standardization of classification in complaints databases; (b) predictive modeling of fraud complaints to identify more at-risk populations or contexts.

5.1 Topic modeling

A major challenge in the EDA analysis was the lack of standardization in complaint categories. This is critical because it hinders the regulator's ability to track complaint types for individual MNOs and compare across providers. A complaint category at one provider may not have the same definition as another provider, or providers may use different category labels to represent the same types of complaints. Topic modeling reduces these categories by identifying common words or expressions across complaints which indicate empirically that they may be related.

As a Natural Language Processing (NLP) technique, **Topic Modeling is useful for identifying latent themes from text using keywords and their combinations. The method may be useful for validating existing MNO complaint classification or discovering new patterns that could later be incorporated to improve MNO complaint categorization and reporting.** Ideally, a regulator could create systems that sort unstructured data, such as notes from customer care records, and use the results to better consolidate and monitor market-level issues.

As seen in Table 6, the number of complaints categories used by each MNO ranged from 20 to 18,576 in 2019 and 25 to 143 in 2020. The number of categories also varied from year-to-year for each MNO. This suggests an opportunity to improve the standardization of complaint classification to the benefit of MNOs and the UCC. The 2019 MNO complaints data contained notes from call center staff describing the complaint interaction and measure taken. This unstructured text provided the opportunity to identify the most common themes or topics using Topic Modeling. Further refinement of this work could result in additional patterns in complaint types and improvement in category reporting, benefitting both MNOs and UCC.

Table 6 Complaints data submissions January 2019 – December 2019

MNO Provider	Months available	Customer care logs	Average complaints per month	Complaint channels	Complaints categories
MNO 1	12	421,855	35,155	4	26
MNO 2	9	1,084,348	120,483	36	18,576
MNO 3	10	695,681	69,568	9	34
MNO 4	3	4,107	1,369	5	20

Preprocessing text data

Topic Modeling analysis begins with preprocessing text data to prepare it for algorithm application. Preprocessing steps include: (a) import original data, (b) convert to lowercase, (c) remove special characters, (d) remove stopwords, (e) stemming, and (f) create term-document matrix.

Preprocessing seeks to reduce noise in the data and distill the text to its most meaningful parts for extracting topics relevant to complaints. Steps b and c "tokenize" the text by breaking the original sentences into words and eliminating cases and punctuation that do not contribute meaningfully to the documents. For example, capital letters and common words like "an" or "the" (otherwise known as "stopwords") add much meaning, yet they may affect algorithm performance. Downloadable dictionaries exist to quickly identify common stopwords before specifying additional stopwords for the data. Step e or "stemming" takes variations of words with the same meaning such as "complain", "complained" or "complaining" and consolidates them as a stem so they are counted as one term for analysis.

The final output of the preprocessing task is construction of a term-document matrix whereby words that appear in each document are counted and compared across a set of documents or a corpus—in our case the corpus is the entire sample of customer complaints to MNOs which also included notes from call center staff. The textual components of this sample, in this case the individual words, are turned into "vectors" meaning they are assigned a unique numerical value. The algorithm then searches for patterns in these vectors. See Table 7 for an example of one complaint (d4) transformed into a term-document matrix. In practice, preprocessing is performed using programs such as Python or R, which allow for management and manipulation of a large amount of data.

Topic modeling analysis of text data

Once the data are ready, we can apply Topic Modeling algorithms. One of the most common is Latent Dirichlet Allocation (LDA)[xxvii], which has been used with Consumer Financial Protection Bureau (CFPB) complaints,[xxviii] but other algorithms optimized for different textual data structures exist. For example, biterm topic modeling (BTM) is useful for short texts, such as tweets or social media.[xxix] Algorithm selection depends partly on the data structure. While LDA and BTM algorithms were applied to the UCC complaints data, the BTM results were used because of the short length of the call-center descriptions.

[xxvii] Blei, D. M. , Ng, A. Y. , & Jordan, M. I. (2003). Latent Dirichlet allocation. Journal of Machine Learning Research, 3 (Jan), 993–1022 .

[xxviii] Kaveh Basatani, Hamed Namavari, and Jeffrey Shaffer. (2019). Latent Dirichlet Allocation (LDA) for topic modeling of the CFPB consumer complaints. Expert Systems with Applications 127, pp. 256-271.

[xxix] Xiaohui Yan, Jiafeng Guo, Yanyan Lan, Xueqi Cheng. (2013). A Biterm Topic Model for Short Texts. Proceedings of the 22nd international conference on World Wide Web, ACM 978-1-4503-2035-1/13/05

Table 7 Document Preprocessing Example

a. Original Document (d4)
Sub complained about losing their airtime after loading it and not using it. On checking sub had opted in to auto renewal of unlimited <name> bundle. Also was advised to the airtime was used to renewal bundle and to avoid such that can buy using option 2 for buy one time and to stop auto renewal. They can choose option 3 for cancel auto renewal off their number. i as well stopped the auto renewal.

b. Convert to lowercase
sub complained about losing their airtime after loading it and not using it. on checking sub had opted in to auto renewal of unlimited <name> <name> bundle. also was advised to the airtime was used to renewal bundle and to avoid such that can buy using option 2 for buy one time and to stop auto renewal. they can choose option 3 for cancel auto renewal off their number. i as well stopped the auto renewal.

c. Remove special characters
"sub", "complained", "about", "losing" "their", "airtime", "after", "loading", "it", "and", "not", "using", "it", "on", "checking", "sub", "had", "opted", "in", "to", "auto", "renewal", "of", "unlimited", "name", "bundle", "also", "was", "advised", "to", "the", "airtime", "was", "used", "to", "renewal", "bundle", "and", "to", "avoid", "such", "that", "can", "buy", "using", "option", "2", "for", "buy", "one", "time", "and", "to", "stop", "auto", "renewal", "they", "can", "choose", "option", "3", "for", "cancel", "auto", "renewal", "off", "their", "number", "I", "as", "well", "stopped", "the", "auto", "renewal"

d. Remove stopwords
"complained", "losing" "airtime", "loading", "using", "checking", "opted", "auto", "renewal", "unlimited", "bundle", "advised", "airtime", "used", "renewal", "bundle", " "avoid", "buy", "using", "option", "2", "buy", "one", "time", "stop", "auto", "renewal", "choose", "option", "3", "cancel", "auto", "renewal", "off", "number", "stopped", "auto", "renewal"

e. Stemming
"complain", "los" "airtime", "load", "using", "checking", "opted", "auto", "renewal", "unlimited", "bundle", "advis", "airtime", "used", "renewal", "bundle", " "avoid", "buy", "using", "option", "2", "buy", "one", "time", "stop", "auto", "renewal", "choose", "option", "3", "cancel", "auto", "renewal", "off", "number", "stop", "auto", "renewal"

f. Construct term-document matrix					
Terms	d1		d4		d89
…	…	…	…	…	…
complain	1	…	1	…	1
airtime	0	…	2	…	2
renewal	0	…	5	…	2
bundle	0	…	1	…	1
advis	1		1		0
stop	1	…	1	…	0
…		…		…	

g. Construct term-document matrix								
	…	complain	airtime	renewal	bundle	advis	stop	…
Document d1	…	1	0	0	0	1	1	…
	…	…	…	…	…	…	…	…
Document d4	…	**1**	**2**	**5**	**1**	**1**	**1**	…
	…	…	…	…	…	…	…	…
Document d89	…	1	2	2	1	0	0	…

We introduce the logic of topic modeling and demonstrate how it works with LDA, while discussion of topic modeling results is based on BTM output.

The LDA model models the probability of the existence of latent topics in a given document, such as an individual call-center complaint text. For example, the only observable variables in the complaints database are words within documents. But we do not know the topics, the distribution of topics within each document, and the assigned words in each topic. The topic modeling algorithm calculates how words are distributed in the corpus (i.e., the full set of complaint documents) and determines the relative weights of topics in each document.

Consider the processed customer care notes example from Table 7 for a consumer who had airtime concerns:

```
Complain lost airtime load using checking
opted auto renewal unlimited bundle advis
airtime used renewal bundle avoid buy using
option 2 buy one time stop auto renewal
choose option 3 cancel auto renewal off
number stop auto renewal
```

The only known variables are these words – and the words in all the other complaint documents. The words "airtime", "unlimited," and "renewal" in blue may be assigned with more weights to one topic category, while "bundle", "lost", and "cancel" in orange may weigh more in another topic category. Based on the distribution of these words and their topic weights an overall topic proportion can be assigned for each document. For example, the topic weight for the category with "airtime", "unlimited", and "renewal" may be 0.55, while the weight for the topic which includes "bundle", "lost", and "cancel" may be 0.25.

While the LDA examines word-document co-occurrence, the BTM model follows a similar procedure except that it focuses on modeling word-word or biterm co-occurrence and thus examines word relationships in short text. Conceptually, the BTM analysis does something similar to the LDA example above except this more sophisticated algorithm is recommended for short texts such as tweets, captions, or headlines ranging between 140 and 200 characters, which fit the UCC complaint text lengths. For example, the sample text in Table 7 was typical and, in fact, longer than the average complaint. After preprocessing the data, this sample complaint consisted of 38 relevant words, compared to the original text which had 75 words before cleaning and preprocessing.

To demonstrate results from the full topic model analysis using BTM, consider first the original topic categories assigned by one MNO in 2019 to its logged complaints. As seen in Table 8, for the MNO complaints for which there were notes, the service provider originally classified the complaints into the following categories: (a) failed calls, (b) account information and modification, (c) dropped calls, (d) mobile Internet, (e) airtime product, (f) mobile money product, and (g) prank caller.

BTM discovered seven new categories. As seen in Table 9, New Category 1 is made up of 55% of complaints originally categorized as 'account info and modification', 36% of

Table 8 MNO's Original Classifications

MNO's Original Category	Number of Complaints	Percentage
Failed calls	33,333	18.27%
Account information and modification	32,899	18.04%
Dropped calls (occurred during complaint)	30,931	16.96%
Mobile Internet	20,938	11.48%
<<Airtime Product>>	16,980	9.31%
<<Mobile Money Product>>	11,213	6.15%
Prank Caller	8,339	4.57%
Others (as categorized by the call center)	4,878	2.67%
Top up	4,299	2.36%
All other categories combined	18,597	10.20%
TOTAL	182,407	100.00%

Table 9 Distribution of Original MNO Categories into 7 New Topics

New Unnamed Category	Original Categories	% of new category consisting of complaints from original category
New Category 1	Account Modification	55%
	Mobile Internet	36%
	Airtime Product	5%
New Category 2	Airtime Product	91%
	Account Modification	2%
	Failed Calls	2%
New Category 3	Failed Calls	39%
	Airtime Product	19%
	Dropped Calls	16%
New Category 4	Airtime Product	90%
	Account Modification	4%
	Failed Calls	3%
New Category 5	Account Modification	72%
	Failed Calls	21%
	Airtime Product	4%
New Category 6	Airtime Product	37%
	Failed Calls	35%
	Account Modification	12%
New Category 7	Mobile Internet	91%
	Account Modification	4%
	Dropped Calls	2%

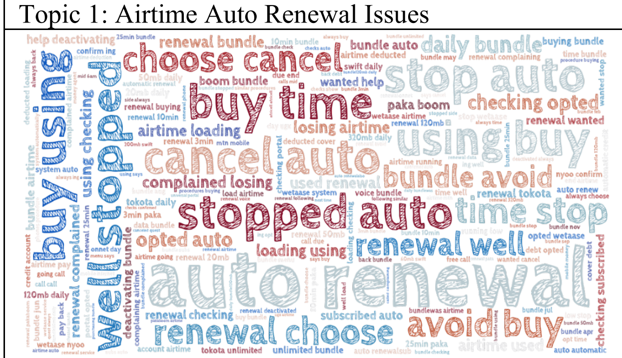
complaints originally categorized as 'mobile internet', and 5% of complaints originally categorized as 'airtime product.' New Category 2 is made up 91% of complaints originally categorized as 'airtime product', 2% of complaints originally categorized as 'account modification', and 2% of complaints originally categorized as 'failed calls.'

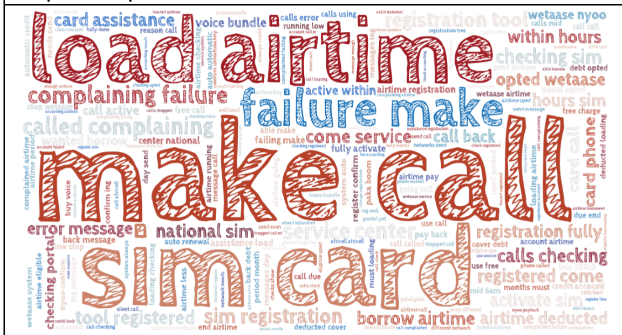
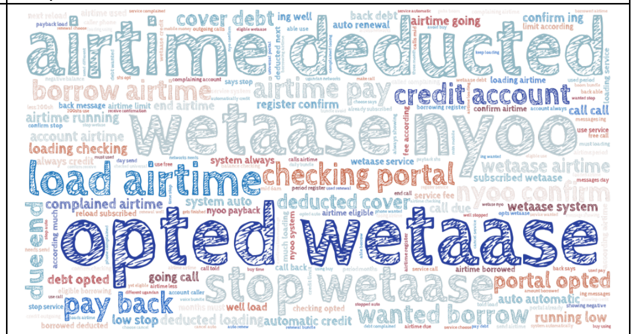
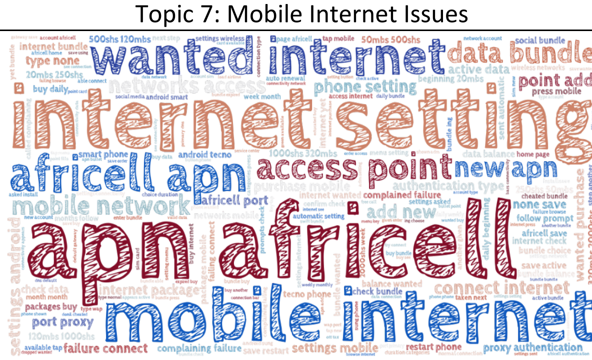
Several insights emerge from these outputs. First, if one looks at the original first tier categories used by the MNO (Table 8), the complaints appear to be related to technical issues (e.g., dropped calls, failed calls) and account maintenance inquiries (e.g., account modification and information). Only around a third of complaints (e.g., mobile internet and airtime product) seem connected to specific products and services. Yet as we will see, **the new latent topic categories reveal that the majority of complaints do relate to specific products/services (e.g. airtime, mobile internet).**

Second, three of the original first tier categories used by the MNO – dropped calls, failed calls, and prank caller – accounted for roughly half of the complaints. These categories are problematic because visual inspection of the notes indicates that some calls classified as pranks were cases in which the caller was silent, indicating a possible technical issue and not the intention to prank call the center. Furthermore, **many dropped call notes indicate enough information to better classify the complaint because the call dropped midway through the customer care handling. However, it is not evident that the customer was re-contacted or the complaint pursued further by the MNO.**

The topic modeling better extracts pertinent information and regroups these original MNO categories across a new topic classification. This can be seen by examining the top bi-terms associated with each new category (Figure 8).

Figure 8 New MNO Categories by Bi-terms

Topic 1: Airtime Auto Renewal Issues	Topic 2: Airtime Borrowing and Deductions I
	
auto_renewal, buy_time, stop_auto, cancel_auto, cancel_auto_renewal, stop_auto_renewal, stopped_auto, well_stopped, well_stopped_auto, using_buy	borrow_airtime, wanted_borrow, wanted_borrow_airtime, 161_register, register_confirm, 161_register_confirm, loading_airtime, call_back, months_must

<table>
<tr><td></td><td></td></tr>
<tr><td>Topic 3: Operational and Service Issues</td><td>Topic 4: Airtime Purchase Issues</td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>make_calls, sim_card, load_airtime, failure_make, complaining_failure, failure_make_calls, called_complaining, complaining_failure_make, service_center, sim_registration</td><td>opted_wetaase, airtime_deducted, wetaase_nyoo, stop_wetaase, load_airtime, load_airtime, checking_portal, pay_back, credit_account, borrow_airtime, airtime_pay</td></tr>
<tr><td>Topic 5: Voice Bundle Issues</td><td>Topic 6: Airtime Borrowing and Deductions II</td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>paka_boom, voice_bundle, auto_renewal, buy_voice, buy_voice_bundle, 134_paka, 134_paka_boom, boom_bundle, paka_boom_bundle, bundle_134</td><td>sim_card, make_calls, load_airtime, auto_renewal, borrow_airtime, failure_make, failure_make_calls, complaining_failure, called_complaining, service_center, wanted_borrow, wanted_borrow_airtime</td></tr>
<tr><td colspan="2">Topic 7: Mobile Internet Issues</td></tr>
<tr><td colspan="2">
</td></tr>
<tr><td colspan="2">apn_africell, internet_settings, mobile_internet, wanted_internet, africell_apn, africell_apn_africell, access_point, wanted_internet_settings, new_apn, mobile_networks</td></tr>
</table>

For example, the original MNO category "prank caller" disappeared completely (compare with Table 8).[xxx] In fact, only one category (Topic 3: Operational and Service Issues) clearly corresponded to failed and dropped calls. Instead, the other failed and dropped call items appear as secondary issues to complaints relating to specific products and services.

Third, **whereas the original MNO first tier classification identified three product and service groups relating to airtime, mobile money, and mobile internet, the new topic categories reveal better the contours of the product and service-related issues, while parsing more Airtime issues**. Topics clearly group around airtime (Topic 1: Airtime Auto Renewal; Topic 2: Airtime Borrowing and Deductions I; Topic 4: Airtime Purchase Issues; and Topic 6: Airtime Borrowing and Deductions II), voice bundles (Topic 5: Voice Bundle Issues), and mobile internet (Topic 7: Mobile Internet Issues). Notably, the mobile money category is not salient in the main topic categories. Upon further analysis, this may be attributed to two related reasons. Further analysis indicated that over half of the mobile money complaints were related to pin reset issues which appear to fall most in Topic 3. It should also be noted that mobile money was not a central service for this MNO, as also seen in the low rates of fraud, and which further supports the hypothesis that the complaints relating to the MNO's original mobile money category may have been reclassified to other technical issues subsumed by the new categories. As noted above, the remaining category (Topic 3: Operational and Service Issues) appear to have related to most SIM card issues and failed calls.

Fourth, **the newly identified latent topics appear to specify better the desired consumer actions and issues relating to specific products.** Consider the example of airtime – auto renewal complaints (Topic 1) vs. airtime borrowing/deduction (Topic 2 and 6) vs. airtime purchases (Topic 4). While all of these pertain to airtime—a higher-level category of product type—they are different issues which reflect different types of service challenges and customer complaints. Although further analysis is needed, the distinction between Topic 6 vs. Topic 1 and 2 appears to be differentiated by "wanted_borrow" issues. Interestingly, **the centrality of these airtime issues as a portion of the complaints received was not apparent from the MNO's original tier one classification, with the category of airtime only accounting explicitly for 9% of the MNO's original complaint classification.**

Fifth and finally, **the account modification category disguised issues of interest to a regulatory body, such as the request to stop auto renewal of airtime (Topic 1) and voice bundles (Topic 5).** In many cases the consumer request was to stop the auto renewal, which signals that there may be issues with opt-in vs. opt-out practices for

[xxx] Inspection of the prank caller complaints found that many of these calls were either people asking for the time or nobody answered the call center personnel, the latter of which could also be attributed to signal issues.

value added services. These services generate additional fees and debt for consumers, and if improperly administered, they are an important consumer issue a regulator would want greater visibility to conduct further inquiry.

While there are other insights, the five findings described above highlight the potential for topic modeling to improve learning and monitoring of consumer issues by both the MNO and a regulatory body. For the MNO in question, common complaints were made in each of the seven newly identified topic areas while call center personnel responded in standard ways explaining how to resolve these issues. **One could envision an MNO using these insights to design a customer care chatbot which could help customers solve the most common straightforward issues, thus helping to free the time of critical call center personnel for more complex issues.** A regulatory body could extend this idea and provide a chatbot for users in the larger ecosystem, customized but standardized according to the most salient and common issues in the marketplace. This chatbot could even filter for individual MNOs and highlight categories of most concern to the regulatory body for each MNO.

Another application of topic modeling is to better understand provider issues beyond their classification approaches, or to suggest new reporting categories to the MNOs. Issues evolve over time and topic modeling could help identify and track emerging problems or new complaints types, which may not be reflected in existing categories. In our analysis this was the case for complaints related to airtime autorenewal, and could similarly be used to track new issues such as challenges with digital loans or other third-party services that emerge on MNO platforms. **These insights may also flag emerging and growing consumer protection risks which can lead to new legislative or regulatory interventions. Such uses of topic modeling tools to analyze the complaint log data would transition the analysis from reactive to more proactive responses, in which emerging issues are identified early on and policy interventions are implemented to address them in a proactive manner.** For example, customer care data could be leveraged to monitor fraud market-wide fraud, a test case explored in the following discussion of predictive modeling.

5.2 Predictive modeling

Fraud, especially by third parties, is a substantial threat to the DFS ecosystem and to MNOs themselves. As discussed in section 4.3, a nationally representative DFS survey found that between May to September 2020, 46% of respondents received a fraudulent call, with 49% of cases involving the fraudster impersonating MNO staff and 2% impersonating UCC staff.[xxxi] In

[xxxi] Matthew Bird and Rafe Mazer. 2020 "Consumer Protection Survey of Digital Financial Services Users in Uganda." IPA: Washington, D.C.

other words, over 50% of fraudsters preyed on the trust in recognized private and public bodies to deceive consumers for financial gain—with 70% of fraudsters requesting the targeted consumer send money. Although 77% of those contacted ignored these fraud attempts, relying mostly on warnings from peers or preexisting knowledge to identify the fraud attempt, 16% of recipients responded, with a quarter of those admitting to falling victim.

Considering that the survey was of a nationally representative sample, the successful third-party fraud rate in Uganda during the first six months of the pandemic was an estimated 2% among DFS users. **Since Uganda's mobile subscription market is around 25 million, roughly 500,000 consumers fell victim to third-party fraud between March and September 2020, making this issue a serious consumer protection risk.**

Yet, just as the increasing scale and complexity of DFS created more opportunities for fraud, the parallel development of AI and machine learning approaches also enabled proactive oversight. **One such opportunity is using past complaints data to predict which consumer and product segments may be the most likely to fall victim to future fraud attempts—and how regulators and providers can more effectively target these consumers with preventive measures.**

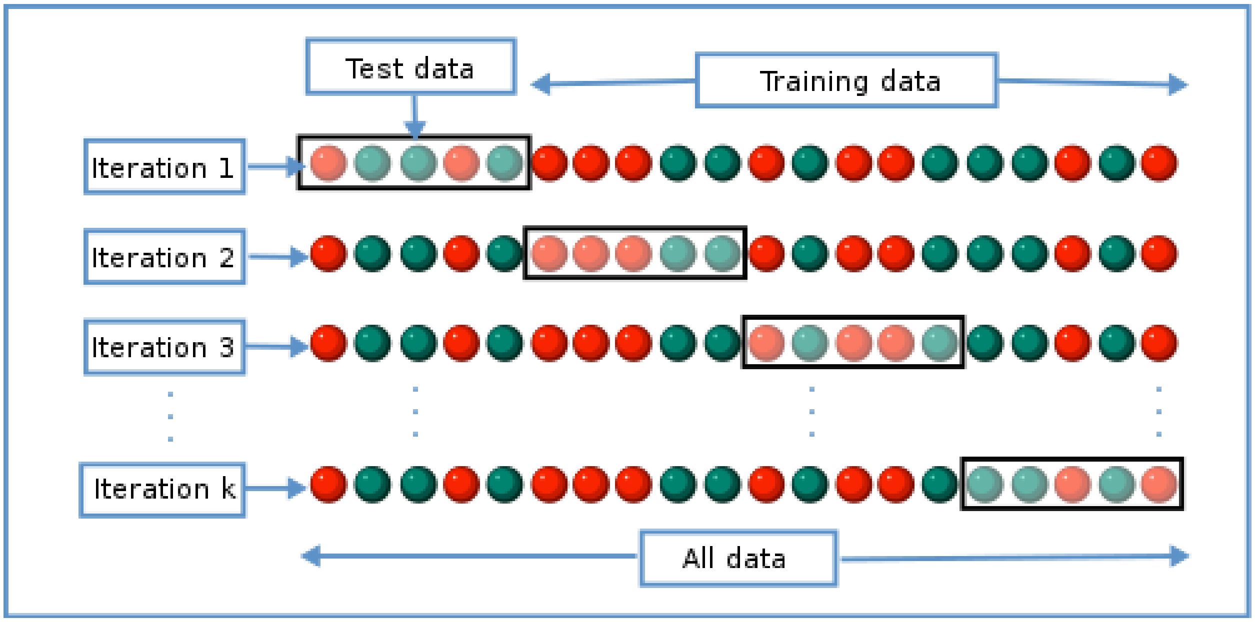
Complaints data analysis is primarily an ex-post monitoring activity—where problems from prior time periods are flagged and corrective actions are proposed to reduce the level of these incidences going forward or improve service delivery. However, past data can still inform future consumer protection practices designed to reduce the likelihood of future incidences, such as fraud. **In this research, algorithm-driven predictive modeling was used with the MNO complaints data to identify the characteristics of a complainant that make them most likely to be presenting a complaint related to fraud.**[xxxii]

A test case was developed with one MNO for fraud-identified complaints between January and June 2020. Variables used in this test case included age, sex, time as client, whether the client was a mobile money user, location of SIM registration, and hour, day, and month of the fraud complaint. A method known as k-cross-validation was then used, whereby the dataset was randomly split into multiple training and testing groups.[xxxiii] This process was repeated multiple times (see Figure 9 for a visual demonstration). The goal of

[xxxii] There are multiple algotrithms and strategies for developing predictive models, depending on the nature of the data and modeling objectives. For this exercise, we used LightGBM, a variation of Gradient Boosted Decision Tree (GBDT), developed by Microsoft. Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu, "LightGBM: A Highlly Efficient Gradient Boosted Decision Tree," 31st Conference on Neural Information Processing Sytems, 2017. Code available at: https://github.com/Microsoft/LightGBM.

[xxxiii] D. Anguita, L. Ghelardoni, A. Ghio, L. Oneto, & S. Ridella. (2012). The 'K' in K-fold Cross Validation. In ESANN (pp. 441-446).

Figure 9 K-cross validation



cross-validation is to prevent overfitting, i.e., making sure that you do not overtrain the model to your data and ensuring that when you apply the model to new data it will retain its predictive ability.[xxxiv]

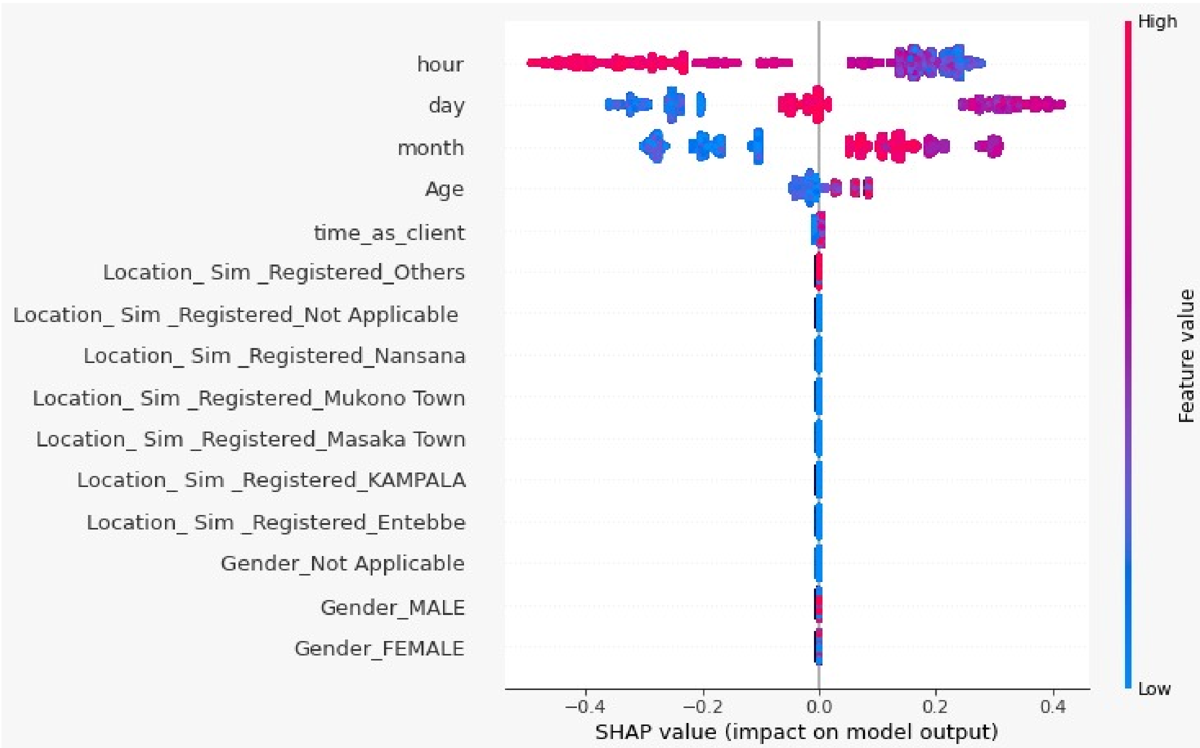
Once concluding the predictive model training process, you can evaluate the predictive accuracy and, perhaps just as important, seek to understand what elements contribute to the model's explanatory power or precision. A challenge with predictive models is trouble understanding the results or inscrutability. The algorithms perform sophisticated computations and it can be difficult to understand how they make predictions based on the variables. As with topic modeling, visualization tools aid analysis.[xxxv]

Figure 10 is a scatter plot representing the predictive elements of the fraud model. The vertical bar on the right, running from blue (low) to red (high) indicates how much the model's individual variables contribute to the prediction of whether a complaint is about fraud. The horizontal bars spanning each variable row show the variable's prediction across a linear range (e.g., low to high values). For example, for "hour" in Figure 10, the blue bar at the lower range indicates less probability of a fraud complaint, while the red bar later in the day indicates increased complaint likelihood.

[xxxiv] In this work, k value was fixed to 10. Other techniques were also used to prevent overfitting the model. For example, instead of using all available locations of SIM registrations, only the top 10 locations were used. Otherwise, the model would risk becoming only applicable to the dataset and would perform poorly when applied to new data in the future.

[xxxv] This exercise used the SHAP (SHapley Additive exPlanations) method, based on game theoretically optimal Shapley Values, to explain individual predictions.

Figure 10 Predictive Modeling Results



made at those hours, since calls to customer care about fraud complaints are more likely to occur at that time of day. While there could be other factors which lead to more calls later in the day, this signals a possible relationships between hour of day and fraud attempts or at least fraud reporting.

Looking across the variables from the complaints data analyzed in Figure 10, the most important variables to predict a fraud call, in ranked order, were:

1. Hour: Calls made EARLIER in the day, more likely to be fraud
2. Day: Days in the middle of the month, with next likelihood at end of month*
3. Month (January – August): MIDDLE followed by later months in timeframe more likely fraud
4. Age: Older people more likely to report a fraud complaint
5. Time-as-client: MORE time as MNO client, more likely to report a fraud complaints
6. Location: Locations outside of major urban areas more likely of fraud
7. Males: Slight tendency for men to make more fraud complaints

In other words, **if you are an older, established, male MNO client from a rural area calling earlier in the day in the middle of June, you are more likely to be calling about a fraud-related matter** .

5.3 Potential next steps for predictive modeling for fraud complaints

The model predicted fraud complaints with 84.6% accuracy. Yet, this example was built to provide proof of concept. Despite use of cross-validation methods, care should be taken to test for overfitting the model, while more data may enable more precise prediction. Regardless, this exercise demonstrates the untapped potential for MNOs, FSPs, and regulatory bodies to better leverage customer care logs to address consumer protection issues such as fraud. **MNOs and FSPs could use supervised machine learning to better anticipate, customize, and prevent fraud, while regulatory bodies could better target fraud intervention and prevention efforts to the most at risk consumer segments.**

Fraud detection at root is pattern identification, be it by a human investigator, rules-detection software, or algorithms. Just as the increasing scale and complexity of DFS enabled fraud opportunities, artificial intelligence and machine learning have enabled better fraud detection.[xxxvi] Typically, financial companies have used proprietary data to identify fraud harmful to their businesses and customers, though these algorithms are not free of biases.[xxxvii] More recently "Regtech" interventions which use new technologies for supervision have emerged and begun testing how bodies can leverage digital tools to monitor markets and activities.[xxxviii] For example, the Consumer Financial Protection Bureau (CFPB) in the United States developed a publicly available complaints database which allows for identification of emerging issues including fraud schemes.[xxxix] Big data analytics of social media data also hold promise.[xl] Relatedly, governments have used machine learning to improve poverty targeting in Afghanistan and identify beneficiaries during the COVID-19

[xxxvi] N. Ryman-Tubb, P. Krause, and W. Garn. (2018). "How Artificial Intelligence and machine learning research impacts credit card fraud detection: A survey and industry benchmark." Engineering Applications of Artificial Intelligence 76: 130-157

[xxxvii] J. West, and M. Bhattacharya. (2016). "Intelligent Financial Fraud Detection: A Comprehensive Review." Computers & Security 57 (March): 47–66; C. O'Neill, (2017). Weapons of Math Destruction: How Big Data Increase Inequality and Threatens Democracy. New York: Crown; C. Criado Perez. (2019) Invisible Women: Data Bias in a World Designed for Men. New York: Abrams Press.

[xxxviii] D.W. Arner, J. Barberis, R.P. and Buckley, (2017). FinTech, RegTech, and the Reconceptualization of Financial Regulation, 37 Northwestern Journal of International Law & Business 371; S. di Castri, M. Grasser, and A. Kulenkampff. (2018). The RegTech for Regulators Accelerator (R2A) Process: Giving Financial Authorities Superpowers. BFA Global. Alexandra Rizzi, Alexandra Kessler, and Jacobo Menajovsky, (2021). The Stories Algorithms Tell: Bias and Financial Inclusion at the Data Margins. Center for Financial Inclusion, Accion.

[xxxix] K. Bastani., H. Namavari, and J. Shaffer. (2019). Latent Dirichlet allocation (LDA) for topic modeling of the CFPB consumer complaints. Expert Systems with Applications 127: 256-271.

[xl] R. Mazer, R., and D. Onchieku. 2019. "Did You See My Tweet? Monitoring Financial Consumer Protection via Social Media." Nairobi, Kenya: FSD Kenya. https://fsdkenya.org/publication/did-you-see-my-tweet-monitoring-financial-consumer-protection-via-social-media/; Tully, M., and Madrid-Morales, D. (2020). Social Media as a Tool for Consumer Protection Monitoring. Presented at the Innovations for Poverty Action Consumer Protection Practitioner's Forum, September 29. https://www.poverty-action.org/event/ipas-consumer-protection-research-initiative-holds-first-practitioner%E2%80%99s-forum-meeting.

pandemic in Togo.[xli] Yet evidence of effective Regtech interventions is limited, especially for solutions that use detection to target and forewarn consumers rather than merely respond to complaints.

The Uganda Communications Commission (UCC) already recognizes the threat of third-party fraud to consumers, especially vulnerable segments. The regulator's current fraud prevention communications involve mass media campaigns, mainly through social media, and guidelines for mobile network operators (MNOs) who are encouraged to send individual fraud prevention messages to its users, mainly via SMS. Figure 11 shows sample messages.

Figure 11 Sample of Existing UCC Fraud Messages

[xli] Blumenstock J.E. (2016). "Fighting Poverty with Data." Science, 353(6301): 753-754; J. Blumenstock. (2020). "Machine learning can help get COVID-19 aid to those who need it most." Nature, 581: 7807.

From a practitioner perspective, responses to several important questions could better inform the UCC's efforts:

1. Do these communications work?
2. How can the communications be improved?
3. Do these communications target the right populations at the right moments?

Predictive models of fraud could better inform these prevention efforts by generating near real-time models that will inform to whom the prevention messaging can be sent and when. Preventing third-party fraud in places such as Uganda may thus lie in combining machine learning fraud detection with insights from digital-based financial education interventions to deliver timely, relevant content on the communications channels most used by targeted financial consumers—in this case DFS customers.

6. Conclusion and Recommendations

6.1 Key conclusions from initial analysis of MNO complaints data

The new methods for complaints data analysis piloted in this research have advanced understanding of the scale, nature, and outcomes of consumer redress for customers of MNOs in Uganda. The research intended to answer several questions of importance to UCC's mandate and to supporting effective consumer protection:

1. **How can the UCC – or a similar regulatory body – better leverage complaints data to inform consumer protection supervision and policy development for telecommunications, digital financial services, and other products operating on MNO platforms?**

The EDA phase of the analysis points to several ways in which complaints data analysis can inform consumer protection supervision and analysis:

- First, standardized performance benchmarks such as resolution time, resolution rate, and volumes of complaints will allow for the UCC and MNOs to monitor their performance over time to ensure continued improvements in complaints handling and redress;
- Second, by monitoring types and volumes of complaints over time, UCC will be able to identify emerging consumer protection issues, or any sudden increases in complaints of a particular nature, and engage providers to ensure these systemic issues are addressed;
- Finally, by integrating demographic data into complaints records, it is possible to identify where use of—or outcomes within—complaints channels vary across population segments, and to better ensure equal access and outcomes for all complainants. This demographic analysis becomes even more useful when it can be triangulated with other demand-side data such as the IPA-UCC consumer protection survey, which also captured data on consumer protection challenges across age, gender, geography and socio-economic status.

2. **What existing service areas can be improved and what unexplored opportunities are there for providers and regulators to improve complaints handling and redress?**

Analyzing complaint type and resolution time proved particularly useful for insights on improving complaints handling and redress. For example, mobile money complaints represent a substantial portion of customer care issues, and further analysis revealed some of the most common issues to include failed bill payment and sending money to the wrong recipient. With such insights on the common challenges for mobile money users, UCC and providers can focus efforts on improving consumer use and outcomes in the specific areas of mobile money where they may still face challenges. Resolution time analysis proved particularly insightful when combined with complementary data such as the type of channel the complaint came in on or the type of issue experienced. For example, resolution times were particularly slow for complaints raised via social media, despite separate IPA analysis of social media data finding this was a growing channel for raising consumer protection issues during the COVID-19 pandemic.

3. How can the complaints data increase provider accountability and where can actors better target legislative or regulatory interventions?

As noted in the first two research questions, metrics such as resolution time can increase accountability of providers for quality complaints handling, while demographic segmentation and time series analysis can identify particular populations experiencing challenges or sudden increases in a type of consumer challenge faced by an MNO's customers. Beyond descriptive statistics, predictive modeling techniques expand the ability to both understand drivers of different consumer complaints and develop more targeted consumer protection interventions. Leveraging complaints related to fraud issues, the research team was able to identify several likely drivers of fraud attempts, including day and time of customer complaint, and location of the customer. Going forward this predictive analysis should be expanded to identify with more certainty what are the most likely drivers of different consumer protection challenges, and which populations are most affected. Robust predictive analysis will enable more precise targeting of consumer awareness campaigns, anti-fraud monitoring, and other preventive measures, potentially increasing the benefit of such efforts by MNOs and the UCC while reducing the costs of delivery.

6.2 Recommendations for next steps

The analysis has yielded a rich set of insights, described in detail in this report. Based on this analysis, the following policy priorities have been identified for continued use of complaints data for consumer protection monitoring:

1. **Develop a standardized method for classifying complaints which all providers will use**. The analysis has demonstrated the utility of the complaints data to gain insights into the conduct of individual MNOs. However, the lack of complete standardization of complaints categories and other variables creates limitations in comparability across providers in areas such as resolution time, or concentration of complaints by products and services. UCC has sought to address this with updates to the complaints template for 2020 data, and has issues a new standardized complaints data reporting template to be used from October,2021 going forward.

2. **Include new demographic indicators in future reporting activities**. MNOs were able to integrate age, gender and

location data into the 2020 complaints data, demonstrating the benefits of adding in data captured during SIM registration to the complaints logs. These variables should be added on a permanent basis, and other demographic indicators should be considered to segment out which types of consumers may face different challenges with the products and services.

3. **Test merging of customer care and transaction data analysis.** Analyzing complaints types and volumes over time can identify any sudden shifts in customer care issues across days, months, or years, allowing for focused interventions to address emerging customer care issues. For example, a complaints spike related to Value Added Services for one provider over a few days in September 2019 is easily identified when the customer care data is separated out by product categories. This issue happened to relate to customers paying for airtime bundles and not receiving them. Where customer care logs identify common or recurring problems like these, the service provider or relevant regulator could investigate further by reviewing transaction data on the sales of these services to identify any practices that may be of concern and ensure affected customers were refunded or otherwise compensated.

4. **Integrate complaints statistics into the periodic reporting on market trends and aggregated statistics.** Aggregated complaints volumes and key indicators like resolution time and complaints by product types could be shared periodically to help the industry measure improvements in customer service over time. Many regulators globally report out market-level complaints data on a periodic basis to measure progress and hold providers accountable for high standards. Summary statistics generated in this report could be integrated into the quarterly subscriber reports from the UCC to similar effect.

5. **Expand the use of methods such as predictive analysis and topic modeling to improve ability to target the most at-risk customers**. New methods of analysis such as topic modeling and predictive modeling have been shown in this project to yield insights that were not immediately identifiable through the Exploratory Data Analysis phase. With a larger sample these tools could be further refined to better determine which populations are most at risk for different issues, and lead to more targeted interventions such as fraud awareness campaigns or publicity regarding issues with services such as automated subscriptions, which were commonly raised in customer care complaints.

6. **Consider new methods to increase the use of formal complaints channels by consumers underrepresented in the data—such as women and rural populations**. The demographic data has identified consumer segments that may not be using formal complaints channels

to the extent of other segments. These consumers may not be aware of, or need more support, to utilize formal complaints channels. Just as UCC has rolled out a widespread awareness campaign regarding phishing scams during COVID-19, new campaigns could be designed and tested to increase consumer engagement with formal complaints channels.

IPA is grateful for the support and partnership of the UCC in conducting this exploratory analysis of complaints data, and looks forward to continued collaboration in the future to implement these and other recommendations to help expand data-driven consumer protection in Uganda's growing mobile and digital economy.

7. Annex 1: Working with Administrative Complaints Data: Case Study and Lessons

The following is an overview of the administrative complaints data used and the steps taken to identify, gather, and clean the data to prepare it for use in this project, as well as our recommendations for researchers considering engaging in similar work.

1. Nature of the Complaints Datasets

The Uganda Communications Commission (UCC) instructed all communication service providers operating in Uganda to submit each month a dataset containing the details of all complaints received that month as well as the redress process that took place to resolve that complaint. The variables captured in the datasets include: respondent ID (an anonymized ID number that the provider could use to re-connect the anonymized complaint to the complainant's personal details within the provider's system), complaint creation date/time, complaint closing date/time, complaint category/subcategory, channel through which the complaint was raised (e.g. call center, social media), and whether the complaint has been resolved or is still pending. In most of the provider's datasets, there is also a free text variable for additional comments, though it is often left blank.

Beginning in 2020, UCC also began instructing providers to submit the gender, data of birth, and location of sim registration for the person who the phone number is registered for all complaints. This was added in order to assess the demographic variation of complaints raised and the redress process. E.g. Are women/people above a certain age/people living in a certain areas expressing different types of complaints? Are complaints raised by different demographic groups being resolved slower or quicker than others? Are different demographic groups using different channels to raise their complaints?

The gender, date of birth, and location captured reflects the person who registered the SIM card that is linked to the complaint. It does not necessarily reflect the person who raised the complaint, as someone might be sharing a phone line with someone else. Ideally, the data would capture the demographics of the complainant him/herself, but this would require the providers to ask this additional information of the complainant during the conversation in which the complaint is raised. This would create further work for the providers and an additional burden for consumers to provide this information. Because of this, the UCC has chosen to use this proxy of the demographics of the person who registered the SIM card, as this information can be linked to the complaints automatically through the provider's internal systems.

2. Project Steps

1. Data Consolidation

While all providers are required to submit their complaints datasets every month to the UCC, some providers do not submit every month, and some have not begun submitting this data at all. Providers are supposed to upload this data on the UCC web portal, but due to technical issues, some of these datasets were lost from the portal in the months after they were submitted. In some cases, the data submissions had been previously shared internally around UCC via hard, so the data was retrievable.

The goal with this project was to consolidate, clean, and analyze all complaints data submissions for communication service providers in Uganda for the period of January 2019 to August 2020. To collect and organize the complaints datasets for sharing with IPA, a process of consolidation was conducted internally with a UCC staff member gathering datasets from the web portal, as well as following up directly with various UCC staff who may have had copies of the datasets that had been inadvertently lost from the UCC portal in order to retrieve as many of the datasets as possible.

In the end, UCC was able to share 57 monthly datasets from 4 different mobile network providers for this exercise.

Lessons Learnt

IPA recommends that researchers who are engaged in a similar process to sit down early in the project with a focal point from the partner organization who will be responsible for the data consolidation and submission. Fully understand the context around the partner's data regulations and guidelines for data reporting. Key questions include:

- How many and which actors are required to submit data, and does the data required vary by actor either in official regulation or in practice? E.g. is the UCC prioritizing getting the bigger providers to comply with data submissions first over the smaller providers?
- Are different types of providers (e.g. MNOs and ISPs) supposed to use the same template?
- How do data submissions occur?
- Do some organizations submit their data differently than others?
- Did the process of data submissions change at some point? This may affect the process/ability to locate data received prior to/after that point, so consider if this will have implications for the project's study design.
- Was the mandate to submit data rolled out equally at the same time to all organizations, or was there a targeted/phased in implementation? E.g. Were MNOs instructed to begin submissions at the same time as ISPs?
- Does the organization provide regulations/guidelines for providers as to how to address complaints and classify/organization their complaints data?
- What documents exist that would lay out these regulations/guidelines?
- Do the providers have internal guidelines/protocols for this themselves, and what documents exist that would lay out these guidelines/protocols? E.g. does the provider have an internal complaints

manual that is providing to call center staff as to how to log and respond to complaints?

This type of information is often not clear just from the data itself and can provide crucial context to understanding the data. E.g. Has the regulator told organizations that they expect complaints of a certain type to be resolved within 24 hours? If so, keep this in mind when analyzing resolution times.

Designate a focal point at the partner organization who will be responsible for the data collection and submission process and direct data requests/updates to this focal point.

Collecting data from different actors within the same organization can lead to duplication of efforts and confusion. Allow for more than ample time for the partner's data consolidation process to occur. While organizations often have clear protocols for how data is meant to be received and stored, in practice these protocols can founder when dealing with large data submissions that come in on a regular basis from many different actors.

If you anticipate this process taking a long time, we recommend coordinating reception of a partial, smaller data submission to do some exploratory analysis on and gain a better understanding of the data source. When following up on data requests to the partner organization, sharing insights from this smaller dataset can demonstrate the utility of the project and generate enthusiasm and momentum within the partner organization.

At the same time as the datasets are being consolidated, track down the relevant materials for contextualizing the data. The most important materials for this project were any instructions/guidelines given to the providers from UCC as to how to address and log complaints, as well as the materials that providers use internally to address and log complaints. In this case, these materials either did not exist or were not provided by the MNOs. This prompted many large lingering questions when analyzing the complaints data. For instance, "How are providers deciding when a complaint is marked a resolved, and do all providers have the same protocol for this?"

2. Data Transfers

The complaints submitted to UCC each month are intended to be anonymized with no personally identifiable information (PII) contained anywhere in the submission. In practice, however, there is often PII in these complaints data submissions. The UCC cannot share complaints datasets in this form with external actors like IPA as this would violate privacy policies. However, the UCC did not have the capacity to anonymize these complaints internally before sharing with IPA. To solve this, an IPA staff member with UCC staff in-person at the UCC offices and wrote a Stata code to anonymize the data. The UCC staff then ran this anonymization code on the data with support from IPA and shared the anonymized complaints dataset with IPA on a hard drive.

Upon reception of the complaints datasets, the first task for IPA was to review the data for completeness. Questions we investigated included:

- Are all the required variables submitted?
    - For several providers, we noticed that they did not include the required, gender, date of birth, and location variables as instructed.

- Did the 'creation date' variable match the month indicated in the title of the data submission
    - E.g. The complaints excel file was titled "MNO 1_October_2019", but creation_date for complaints were all from August and September 2019.
- All answers all within logical windows?
    - E.g. Does the date/time of the complaint closed variable occur after the data/time of the complaint logged variable (e.g. complaint logged 11:00am September 21, 2019, closed 11:59am September 20, 2019)
    - Is the date of birth given reasonable? Several complaints had year of birth listed as 1910 or earlier

- Do any of issues identified prevent the analysis from being carried out as planned?

After these completeness checks, UCC followed up with providers asking them to clarify and address the issues and submit corrected data promptly. The back-and-forth with the providers took a significant amount of time, and clarifications were not provided for all the questions. We would advise accounting for this back-and-forth period in the project timeline and budgeting accordingly.

We would also recommend direct engagement with the providers early on in the provider to build a relationship that will allow for back-and-forth clarifications as data is shared and reviewed. It was our experience that some providers were skeptical of and/or uninterested in the project and did not respond fully to the questions asked. We believe that direct engagement with the providers early on to explain the project and to demonstrate the value to the provider could have help prevented these later issues.

3. Data Cleaning

As the complaints datasets included over 3 million customer care logs, preparing the data to be used for analysis was a lengthy process. Most of the variables used by providers did not include restrictions on the data entered so were in effect free-text variables. Basic data cleaning was done first to all string variables including removing leading/trailing/superfluous spaces, unnecessary punctuation, abnormal characters, as well as converting text to lowercase. However, after this process significant issues remained, which required extensive string and fuzzy matching to convert the data into a usable form.

For example, after basic cleaning MNO 1 still had 509 different values for the variable ‘district of SIM registration’, even though Uganda only has 135 districts. Many of the values contained for this variable referred instead to neighborhoods, streets, or landmarks contained within a district. String-matching cross-referenced with google maps was done to link these locations to their appropriate districts. Analysis of the numerical variables also revealed many errors, e.g. complaint closing date pre-dating the complaint raised date, date of births being before 1900. In these cases, values were dropped.

Another example that demonstrates the challenges faced in cleaning such unstructured data was the issue of creating clear category and sub-categories for each provider. MNO 2 had 18,576 complaints categories in its 2020 dataset. A large portion of this can be explained by spelling mistakes/slight variations between phrasing of complaints categories. However, even after addressing these issues, it was difficult to identify the main overarching categories represented, as they were so numerous and seemed to overlap with each other frequently. For example, often ‘service complaint’ was listed as a sub-category under ‘mobile money’ wherein other cases, ‘mobile money’ was considered main category with ‘service complaint’ listed as one of the sub-categories under it. ‘Voice combos’ would often come up as its own category but also appeared also as a subcategory under ‘product’, ‘service complaint’, ‘beerako,’ ‘promotion’, and ‘tariff’.

Clarifications were sought from the providers as to how to interpret the often-conflicting data points presented in the complaints datasets, but unfortunately these were not provided. Again, we believe that early and congenial engagement with the providers could have gone a long way to clarifying these issues.

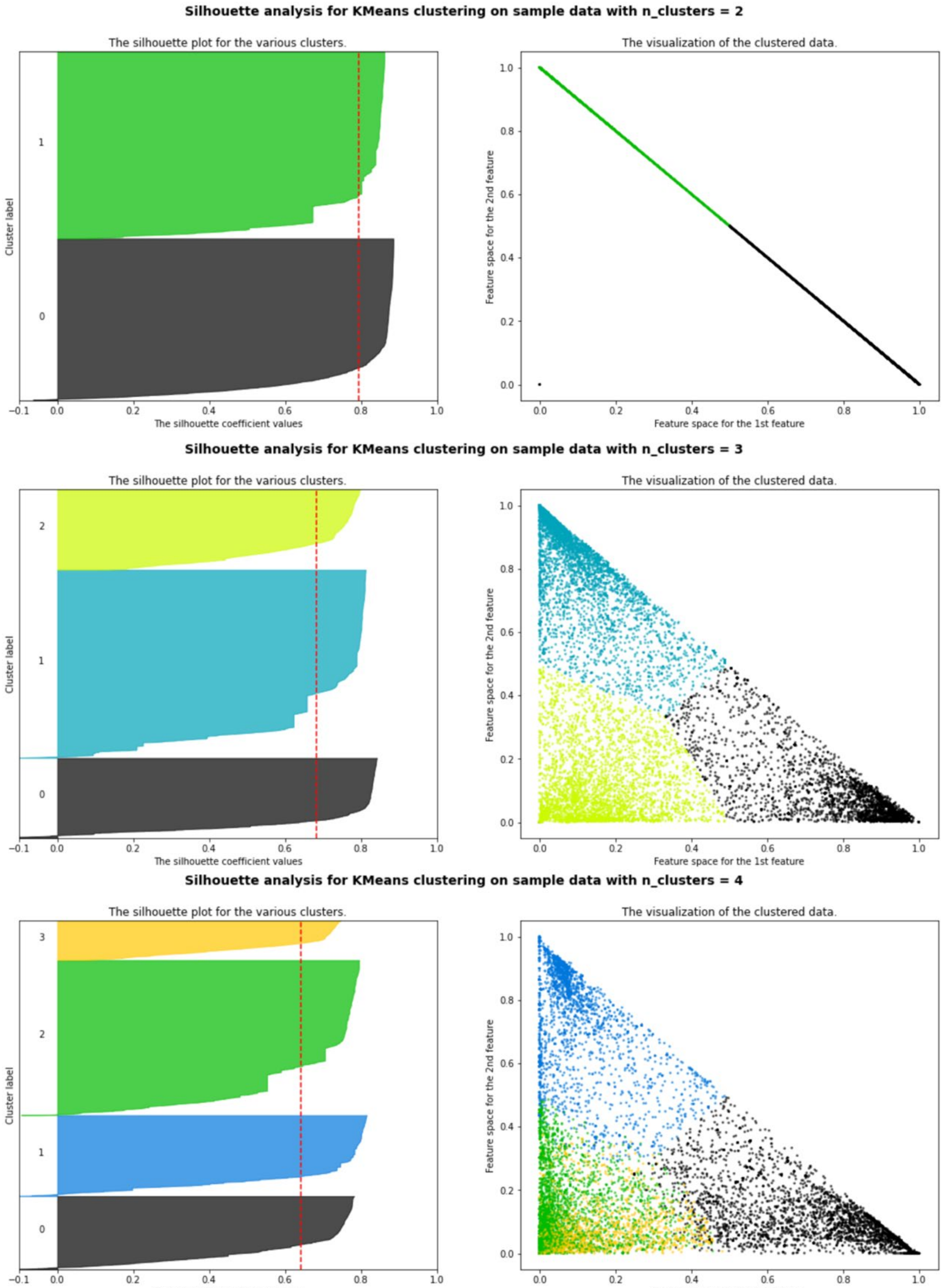
8. Annex 2: Topic Modeling Evaluation Metrics

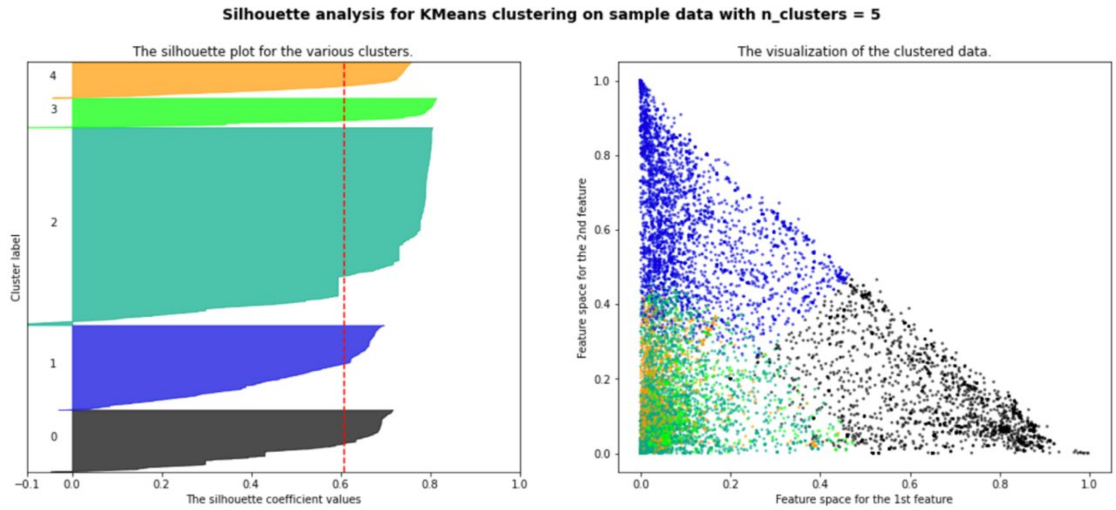
Various metrics can be used to assess the optimum number of topics. One of the most common is the Silhouette coefficient, which consists of two scores. The first score type measures how similar (or close) the points are in the same cluster. The second measures how different (or distant) any point is compared to other clusters. A coefficient is then computed, ranging from -1 (poor clustering) to 1 (dense clustering), with 0 indicating clusters that overlap. The higher the score the more clearly defined the clusters. Figure 12 visualizes the results. The first panel plots the cluster size and coefficient (red line) for each topic, while the second panel plots points in a geometric frame. As we increase the imposed number of clusters, we see how the topic modeling algorithm regrouped the categories (e.g., left panel bars, right panel points).

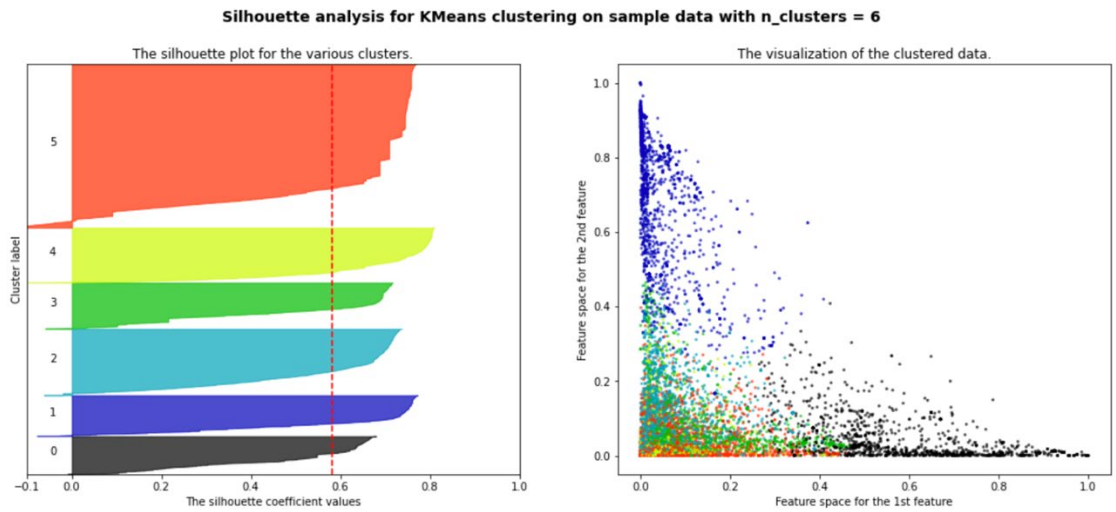
If one were to select the optimum number of topics based solely on the Silhouette score, two clusters would appear as the best option. However, for practical purposes this division may not be useful. Additional criteria may be used, including maximization of the number of clusters with minimum loss in the Silhouette score, a rule that every individual cluster score reaches the overall Silhouette score, and the relative balance in cluster sizes. A final validation is qualitative inspection of the generated clusters to confirm meaningfulness of the results and interpretation. This process is not iterative. For example, after generating results and evaluating them quantitatively and qualitatively, adjustments may be made which may involve adding new stopwords or creating new stems.

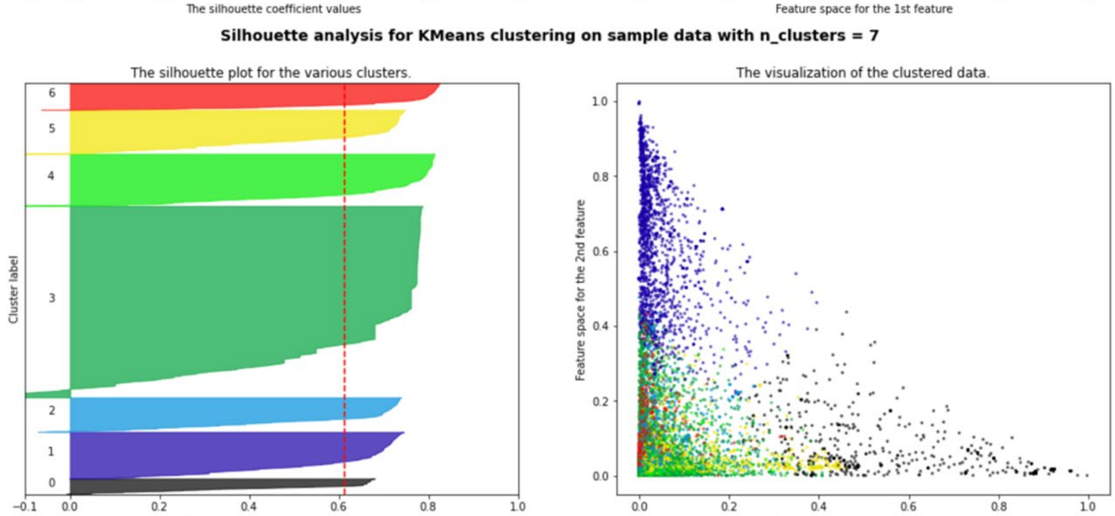
According to the Silhouette score and cluster balance for the UCC topic modeling exercise, one could choose having two, three, or four clusters. In fact, these numbers may be the optimum for specifying a first tier category for complaint classification. However, we were interested in exploring how many categories could be identified. We thus systematically increased the cluster sizes. For cluster sizes five, six, and seven, the Silhouette score hovered around 0.6 until it dropped. Six clusters had a score below that of five and seven clusters. The remaining two clusters had virtually identical scores at 0.6088 (five clusters) and 0.6090 (seven clusters). In the interest of selecting the maximum number of clusters, we chose seven. But how meaningful were they?

Figure 12 Topic Modeling Visualization




Silhouette analysis for KMeans clustering on sample data with n_clusters = 5
The silhouette plot for the various clusters.
The visualization of the clustered data.
Cluster label
The silhouette coefficient values
Feature space for the 2nd feature
Feature space for the 1st feature


Silhouette analysis for KMeans clustering on sample data with n_clusters = 6
The silhouette plot for the various clusters.
The visualization of the clustered data.
Cluster label
The silhouette coefficient values
Feature space for the 2nd feature
Feature space for the 1st feature


Silhouette analysis for KMeans clustering on sample data with n_clusters = 7
The silhouette plot for the various clusters.
The visualization of the clustered data.
Cluster label
The silhouette coefficient values
Feature space for the 2nd feature
Feature space for the 1st feature