12-7-2022

Targeting Aid Beneficiaries: Insights from Bank Transaction Dataset Analytics

Mesfin Fikre Woldmariam
Assistant Professor at Addis Ababa University, mesfin.fikre@aau.edu.et

Tesfahiwot Tefera
Commercial Bank of Ethiopia, tesfahiwot3@gmail.com

Targeting Aid Beneficiaries: Insights from Bank Transaction Dataset Analytics

Full research paper

Woldmariam Mesfin

College of Business and Economics; Institute for Development Policy Research

Addis Ababa University

Addis Ababa, Ethiopia

mesfin.fikre@aau.edu.et

Tefera Tesfa

College of Business and Economics

Addis Ababa University

Addis Ababa, Ethiopia

tesfahiwot3@gmail.com

Abstract

Identifying the poorest individuals for aid and development intervention is challenging. The survey approach is prone to biases by local surveyors and supervisors. Respondents usually give the wrong response if they know the data is collected for aid purposes. In an alternative approach to the survey, researchers are exploring the use of night-time satellite imagery and mobile phone data. But these approaches have limitations too. In developing countries, the rural poor are not connected to the electric grid; hence, there is no night-time satellite image. In this paper, we demonstrate how bank transaction datasets give insightful information for targeting. These datasets are collected in real time, are accurate, and cost-efficient to prioritize counties for targeting. The knowledge gained from this study also provides valuable insights to develop data sharing polices, as input to other targeting models, to develop machine learning models, and to develop a national digital poverty map.

Keywords: targeting aid, data analytics, time series, bank transactions

1 Introduction

In general, targeting can be defined as"... the act of ensuring that aid reaches people who need it most at the right time and place, in the right form and quantity—and at the same time, does not go to people who don't need it" (Schwartz 2019). Scarce resources should reach those that need them the most. To achieve this, different targeting approaches and criteria were in use.

Aid organizations rely on government and international agencies' supplied datasets and information. These datasets are usually census data that governments collected many years in the past. They are out dated and misleading. When government officers collect data that involves financial matters, respondents do not provide accurate information. They suspect the government is collecting such data to levy more tax. In addition to this, if respondents suspect data is gathered for aid purposes, they usually exaggerate their problem and demand aid (humanitarian or developmental). This leads to poor decisions and wrong targets. Besides these challenges, the survey approach is expensive, subject to biases, and based on sample representatives. In poor countries, statistics are used less, and data for statistical analysis is not captured much (data intensive development lab). In much of Africa, for example, national statistics on economic output may be off by up to 50% (Blumenstock 2016). Globally, in the decade between 2004 and 2013, at least 74 countries had no more than one estimate of the national distribution of poverty (Serajudding et al. 2015). This calls for a new, innovative, efficient, and effective approach.

Targeting criteria can be based on the characteristics of individuals, communities, or geographic places. In the process, there are many errors (inclusion and exclusion errors). Inclusion errors occur when people who should not benefit from an aid program get the benefit, while exclusion errors happen when individuals who would have benefited from an aid program are excluded. It is challenging to determine efficient and effective targeting criteria and approaches that minimize targeting errors. Measuring the impacts of these hazards is not an objective and uniform through survey approach. Thus, it could lead to wrong target.

Traditionally, beneficiaries are targeted by assessing geographic factors like population density (urban, peri-urban, and rural), economic zone (low income areas or farming), pastoral, hunting, mining, logging, and ecological zone (littoral dry, humid plateau, humid mountain, dry mountain), ethnic or sociocultural/linguistic groupings, infrastructural and service conditions (availability of services such as schools or infrastructure such as water, roads, and electricity), and or security situation or restrictions (conflict or high crime areas, zones of guerilla activity). Such an approach is also misleading. There are people living in arid or semiarid places who are not as poor as people living in other places. A study in Kenya by (World Food Program Kenya) revealed that 43%–64% of the total number of acutely and chronically malnourished people and food insecure households in the country live in non-arid and semi-arid land areas and are excluded from aid schemes by design.

Once priority areas are identified through one or more of the above criteria, the next step is to actually identify individual beneficiaries. This can be done in a number of ways: self-targeting or community targeting. Targeting datasets can be gathered by establishing dedicated data collection and curating canters or programs like (Kaba et al 2018; Raj & Relton 2013), which is based on community participation and a rating approach. Identifying trusted individuals (men and women) in the community who are the centre of gravity is another approach (Foster, 2021). These individuals can give their genuine and honest assessment of their community that can be considered for targeting. This approach has a 95% success rate and appears to be more successful in selecting households that are, on average, more 'vulnerable' than others as measured by their food (in)security status, proxies of wealth, and levels of education.

Community based targeting also appears to be inefficient. Community leaders can be misleading and crate variations from place to place. Local notions and understandings of poverty and vulnerability by community leaders are not the same as those developed by the government or development partners (Conning & Kevane 2002). There is also a lack of transparency, discriminatory practices, and the exclusion of the poor (McCord 2013). But this approach is considered valuable and appears to be legitimate by the community as it brings community knowledge to the targeting process, which is not the case in other forms of targeting (Devereux et al., 2015). Each of these approaches has its own limitations, and there is no one best way (Jaspers & Shoham 1999; Schwartz, Hernandez, & DeMattee; Sharp 1998). Success metrics are determined by how well the targeting mechanism is designed, implemented, and reduces both inclusion and exclusion errors. A study by Clay, Molla, & Habtewold (1999) indicated that there is a huge error of inclusion and exclusion while targeting for food aid.

Besides the inclusion and exclusion errors, aid organizations should change their reactive approach into a proactive or anticipatory approach (Arendt-Cassetta 2021; Hernandez & Roberts 2020). Technologies like artificial intelligence and predictive technologies can support anticipation by enabling earlier, faster, and potentially more effective humanitarian and development interventions. For example, the targeting algorithms used by Facebook and Google to place their ads in front of their target audience are efficient and effective. This could be a good inspiration to explore the use of machine learning and artificial intelligence techniques in humanitarian targeting (data intensive development lab n.d).

Following the recent cashless transaction policy, Ethiopians are forced to use banks for payments and transfers. As a result of the policy, individuals are discouraged from using cash-based transactions and transfers. This creates an interesting opportunity for everyone to come to banks, which again helps banks to collect detailed and quality datasets. Unlike datasets captured through surveys (which can easily be manipulated), bank transaction datasets are accurate. Such datasets also capture seasonal variability due to different factors. In any case, bank transaction datasets can help to make better and more informed decisions. In this regard, the following section discusses studies on the use of machine learning techniques for targeting purposes.

2 Related Works

As an alternative solution to the above challenge, machine learning and artificial intelligence scholars have started exploring the use of datasets like night time satellite images of places and mobile phone datasets (Blumenstock 2016). Algorithms can play a significant role in identifying targets to provide aid to. Unlike human officers, computer algorithms can continuously track and capture the activities (behaviour) of individuals (Abebe & Goldner 2018). Such objective behaviour tracking in a more continuous manner helps aid organizations.

Specifically, researchers from UC Berkeley's Centre for Effective Global Action, Stanford University, and Facebook's Data for Good have explored alternative targeting approaches that are based on the average wealth of each location. They developed a relative wealth index of places at 2 by 2 kilometres for 135 low-income countries. The wealth index combines data like demographic health survey (DHS) data, satellite imagery, topographic maps, mobile network data, and connectivity data. According to them, counties with high population densities, lower road infrastructure, fewer smartphones, Wi-Fi, shorter phone call durations, and lower mobile internet usage suggest a lower relative wealth of communities than counties with higher values for these variables. By computing and comparing the values of these variables for different counties, governments and aid agencies can target their development and humanitarian aid. But in developing countries like Ethiopia, the majority of counties that seek aid are not connected to the electricity grid, and hence there is no night time satellite imagery.

Besides the use of night time satellite images and mobile phone datasets, there are other insightful proxy measurements (indirect measurements) in the areas of artificial intelligence for social goods. For example, predicting floods from Flickr data tags (Tkachenko, Jarvis, and Procter 2017); analysing business recovery from natural hazards from Facebook datasets (Eyre, De Luca, & Simini, 2020; Lam et al. 2022); analysing disaster damage from Twitter datasets (Kryvasheyeu et al. 2016); understanding disaster resilience from Twitter datasets (Zou et al. 2018); and developing disaster maps of people looking for aid from Facebook usage patterns (Maas 2019). These studies show how the use of proxy measurements (assessments) is becoming common.

In a similar analogy, behavioural economics and finance literature offer great insight and proxies to measure poverty and or wealth levels and hence assist in targeting. According to Karlan, Ratan, and Zinman (2014), the poor can save at banks to avoid the temptation to spend it on unnecessary things, and being stolen by close friends and individuals (Banerjee and Duflo 2006; Hoos 2010; Reyers 2019) and as a buffer against financial shocks. Reyers (2019) found that financial capability is highly associated with saving for emergencies. Those with greater financial capability and those with access to a bank account are more likely to have emergency savings compared with those with lower levels of financial capability.

During emergencies and hardship times, individuals have lower saving amounts (Eskander et al. 2018) and more withdrawals (Brei, Mohan, & Strobl 2019) from their accounts. This means bank transactions like deposits, withdrawals, transfers, etc. are indicative of the wealth or poverty of a country. In light of this knowledge and evidence, by looking at bank transaction datasets and their trends over time, this study documented interesting insights that can help decision makers prioritize counties for targeting. Even better than satellite night time light imagery datasets, bank account financial activities of individuals can reveal the reality on the ground in a real-time and granular manner. These datasets are also captured automatically by bank infrastructure. The following section describes the nature of bank transaction datasets and the analysis or insights derived from them.

3 Research Dataset

As stated above, almost all targeting research uses night time satellite imagery, mobile phone data, and demographic and health (DHS) datasets. With respect to developing countries, these datasets have some limitations. Thus, in this paper, we resorted to using bank transaction datasets of counties for targeting purposes. These kinds of datasets are real-time and cost-efficient (as they are collected by banking applications). Such datasets are also genuine as they are far from direct human manipulation.

Based on United Nations Office for the Coordination of Humanitarian Affairs (OCHA 2001) reports about the number of individuals seeking aid, 25 bank branches operating at the most in-need aid sites are selected. The branches made over 23 million daily transaction records, with an amount of over 18 billion Ethiopian Birr in 2020 and 2021. But, in this paper, we present findings based on datasets of 12 bank branches operating in the most aid seeking counties in the Amhara region of Ethiopia.

The majority of the branches have transactional datasets like AC (account to account transfer), deposit (saving), withdrawal, mobile transfer, internet banking, ATM withdrawal, cheque withdrawal, POS (Point of Sale), local money transfer (LMTS), local money transfer at interest free, and business transfer over the web. Based on literature and our own understanding about which transactions are informative about the poverty of communities around those bank branches, we analysed deposits, withdrawals, account transfers, ATM withdrawals, mobile transfers, and POS transactions. Table 1 gives a brief explanation of why we selected these transactions and their respective meanings in the context of banking practices in Ethiopia. The following section elaborates on the steps followed in data processing.

Trans type	Transaction mode	Brief explanation
AC: Account to account transfer using transfer form	It needs physical presence of account holder at the bank's branch	The account may/may not be opened at that specific location. But it shows money transfers took place at that particular area branch. This can show more volume of AC means communities around that specific place are better than other places.
Deposit: Cash deposit using deposit voucher	It needs physical presence of the depositor at the bank's branch	The account may/may not be opened at that specific place. But it shows physical cash is deposited at that particular area branch. Higher deposit amounts means the communities around such bank branches are economically better off than others.
Cash withdrawal using Withdrawal voucher	It needs physical presence of account holder at bank's branch	The account may/may not be opened at that specific place. But it shows physical cash is withdrawn from that particular area branch. This can show higher withdrawal amount means communities around such bank branches are economically better off than others with lower withdrawal amounts.
Account to account transfer using mobile banking	It doesn't require physical presence of individuals.	It shows the account holder needs to open a bank account at that specific bank branch. But transfers can be made from anywhere in the country. The higher the volume of such transactions, the wealthier the surrounding communities are (in relative terms).
Cash withdrawal from ATM	Needs physical presence of card holder at bank	The account may/may not be opened at that bank branch. But the transaction took place at that particular area branch. This shows the higher the ATM withdrawal amount, the better the community around such bank branches.
POS: Cash withdrawal/money transfer using POS	It needs physical presence of ATM card holder at bank branch	The account may/may not be opened at that specific branch. But cash is withdrawn or transferred from that bank branch. This shows the higher the volume of POS the better the wealth of the communities around such bank branches.
LMTS: Local money transfer	Need physical presence of a person who transfers	It doesn't touch any accounts; just physical cash is transferred from one person to another person through the bank. This means the higher the volume of LMTS the better the wealth of the community around such bank branches.

Table 1: Bank transaction types and descriptions

3.1 Outliers and dataset format

In financial modelling and analysis, a seemingly outlier observation may not really be an outlier, even though we think it to be. Until we understand the nature of the dataset, we cannot consider it as an outlier and remove it from the dataset. The existence of outliers can signify the presence of unnoticed scenarios and may have the most salient information. Sometimes we are actively looking for outliers, even if outliers can even be the goal! Most analysts recommend not removing outliers because of being outliers. One needs to have a scientific reason that justifies the removal. In the context of this paper, we prefer not to remove outliers. The final curated dataset has the structure as shown in table 2. The worda column shows the location of the bank branch; the date column shows specific dates the transactions happen; Trans column shows the kind of transactions (as indicated in table 1); the count column shows the number of times (frequencies) each of those transaction types happens; and the Amount column shows the volume of each transaction on each of those dates.

Worda	Date	Trans	Count	Amount
Dejen	Fri 1 Jan 2021 00:00:00	AC	105.0	21 316 165.
Dejen	Fri 1 Jan 2021 00:00:00	ATM Withdrwal	797.0	806 429.
Dejen	Fri 1 Jan 2021 00:00:00	Cheque Withdrwal	14.0	246 621.
Dejen	Fri 1 Jan 2021 00:00:00	Deposit	372.0	2 922 435.

Table 2: format of curated data set

3.2 Computing Ratios

The number of customers or communities using each bank is different. So, we cannot compare the wealth or poverty of different counties or communities directly. Seen in aggregate, counties with a higher number of customers can have more transactions and a higher volume of transactions. Thus, it is important to find the ratio of volume of each transaction type to the number of each transaction type across a two year period and assess the trend.

3.3 Construct Time Series models

As can be seen, the dataset has a time stamp. This helps us to approach this study as a time series problem. We can create a time series object for selected transaction types (deposit, account transfer, withdrawal, mobile transfer, POS transactions, ATM withdrawals, and local money transfers) for the selected branches as shown in figure 1(a): Wolfram Mathematica's in-built TimeSeries function was used.

This kind of data transformation (process of converting raw data into a format or structure is suitable for model building and data discovery in general). It helps to minimize algorithm bias when the data distribution is skewed (Gong, 2021), which is the case in this study. Such dataset observations need to be continuous in time and allow only one path. But bank transaction datasets are not continuous, as banks do not operate on holidays and Sundays. But some transactions, like ATM withdrawals and mobile transfers, do happen during such days. To minimize the adjusted loss function and prevent over-fitting or under-fitting, we regularized each dataset by using an interpolation technique. In time series analysis, models assume current values will be expressed as driven in some ways from past values. Each dataset is resampled by using the TimeSeriesResample function. This function uses (linear interpolation methods by default). To make the visualization more visual for easy comparison, datasets are resampled into months (figure 1 (b)) except for POS transactions, which are resampled into weeks.

```
Time: 01 Jan 2020 to 31 Dec 2021
Data points: 576
Regular: False    Output dimension: 1
Metadata: None    Minimum increment: {1, Day}
                  Resampling: {Interpolation, 1}
```

Figure 1 a: Sample time series data object (not regularized)

```
Time: 01 Jan 2020 to 31 Dec 2021
Data points: 24
Regular: True     Output dimension: 1
Metadata: None    Minimum increment: {1, Month}
                  Resampling: {Interpolation, 1}
```

b: Sample regularized time series data object

Then, to analyse how poverty or wealth has trended over the two years for communities around selected branches, we have computed percentage changes in the following subsequent months with respect to the values of the first month. Figure 2 shows a sample time series object showing the precent change.

```
Time: 01 Jan 2020 to 31 Dec 2021
Data points: 24
Regular: True      Output dimension: 1
Metadata: None     Minimum increment: {1, Month}
                   Resampling: {Interpolation, 1}
```

Figure 2: Sample time series data object of percentage changes with respect to the first month

Depending on the type of transaction, we can gain insights about how the wealth or poverty of communities around those bank branches is trending. For example, places with better economic situations will reveal an increase in the deposit as compared to other places. Contrary to this, if the economic situation of a place gets worse as compared to the first month, then the ratio will be lower (below zero). Those with zero values for this ratio mean there is no up or down with respect to the value of month one. The following subsection reveals the visualization and hence wealth or poverty trends.

4 Visualize seasonality and trends of selected transactions

As shown in figure 2, once the dataset objects are prepared, we can visualize the patterns and contrast the wealth of the local communities surrounding those bank branches for particular transactions. This is covered in more detail below.

4.1 Trends in Percentage Changes in Deposit Transaction

Close observation of cash deposits by communities using those branches reveals that there are huge ups and downs. As can be seen from figure 3 (a), none of the study sites reveal a uniform positive or negative trend. Counties like Werebabu, Dehana, and Jamma reveal relatively different behaviours. Compared to the first month, there are months where they have more than three time deposit amounts. But, the majority of the counties show a negative, which means their subsequent deposit amount, is lower than the first month. Some counties have the least and decreasing deposit amounts, though the majority of the months. For example, Wereilu appears to have the least of all others, followed by Tehuleder, Habiru, and Wadla counties. One can also easily observe which county gets the lowest and in which month. This is an interesting insight for targeting aid seekers. As discussed in the literature, counties with lower saving (deposit) amounts are relatively poorer than others. Another interesting insight about these transaction types is that each country has ups and downs at almost the same time (month), irrespective of how high or low the volumes are.

4.2 Trends in Percentage Changes in Withdrawal Transaction

Regarding withdrawal, figure 3 (b) shows that majority of the study sites have a below zero (negative) trend compared to their respective values at the beginning of the month. From this figure, we learned that there are counties (East Belesa, Habiru, Raya Kobo, and Metema) whose withdrawal is reduced by over 50% with respect to the first month. Of course, there is an up and down with time. Similar to deposit transactions, withdrawal transactions also do not have uniform ups or downs. But, in general, there is a tendency towards negative secular trends. It can also be seen which county has the least withdrawal amount as compared to other counties and can inform targeting counties for aid.

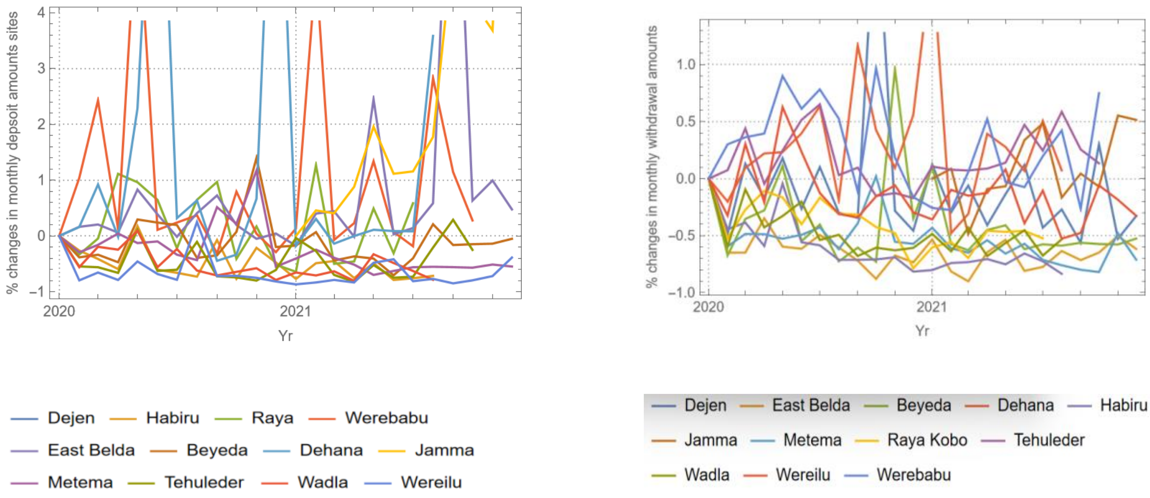

Figure 3 (a) trends in percentage changes in deposit, (b) trends in percentage changes in Withdrawal

4.3 Trends in Percentage changes in Account to account transfer

Account to account money transfer figure 4 (a) shows a relatively positive secular trend for a few of the counties. Similar to deposit and withdrawal transactions, account to account transfer transactions do not show a uniform up or down trend. Again, this transaction has highs and lows for the majority of branches (counties) almost at the same time (months), similar to deposit and withdrawal. It should be understood that the highs and lows are in relative terms. As with poverty or wealth analysis, even though many branches show a decreasing trend (negative-below zero) in 2020, branches like (East Belesa and Dejen) exhibited improvement and a positive trend in 2021. One branch (Beyeda) shows the least across the two years' period. This insight adds value to development and humanitarian aid agencies in their targeting tasks.

4.4 Trends in Percentage Changes in ATM withdrawal Transaction

Unlike other transaction types, the percentage changes in ATM withdrawal amounts at the different bank branches oscillate between an increase and a decrease of 50% of the value in the first month (figure 4 b). At one branch (Wadla), there is a clear outlier. This branch has a significant high percentage increase (50%-100%). From the dataset we cannot tell why this happens. This branch has no data for 2021. Again, the majority of the branches have highs and lows at almost similar time periods. From the figure, one can also tell which branch has encountered the least ATM withdrawal amount and when. For example, Dehana, Werebabu, and Beyeda had the least withdrawal changes over the two years' time (with respect to the first month).

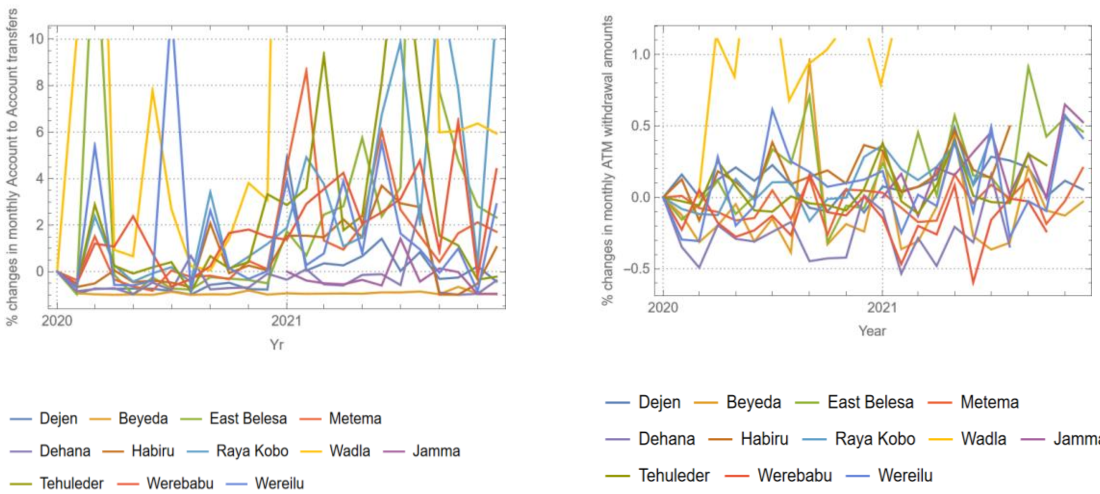

Figure 4: (a) Trends in percentage changes in AC, (b) Trends in percentage chnages in ATM

4.5 Trends in Percentage Changes in Mobile Transfer Transaction

Generally, mobile transfer transactions show a decreasing trend (figure 5 a). There is a clear difference between the two years. In 2021, almost all branches showed a decrease in the amount of mobile transfers, while in 2020 there were both increase and decrease trends. Close observation of individual bank branches reveals Raya Kobo is the least of all in 2021, while Werebabu, East Belesa, and Tehuleder have shown the least in 2020. It is also clear when and which branch has got the least.

4.6 Trends in Percentage Changes in Point of Sale Transaction

Unlike the other datasets, the number of data points for the point of sale transaction type is small (in relative terms). Thus, this dataset is resampled into weeks. Even though this dataset reveals both positive and negative changes (increase and decrease), the amount of increase is significantly high. As can be seen by figure 5 (b), the percentage increase in POS transactions for the majority of the branches is very high, from 200% to 850%. This is quite unusual compared to other transaction types discussed above. This might be due to the war in the region; people might resort to using digital money rather than cash. Still, there are branches with the least percentage changes across the two years. For example, branches like Raya Kobo, Habiru, Dejen, Werebabu, and Tehuleder show decreasing trends for POS transaction types.

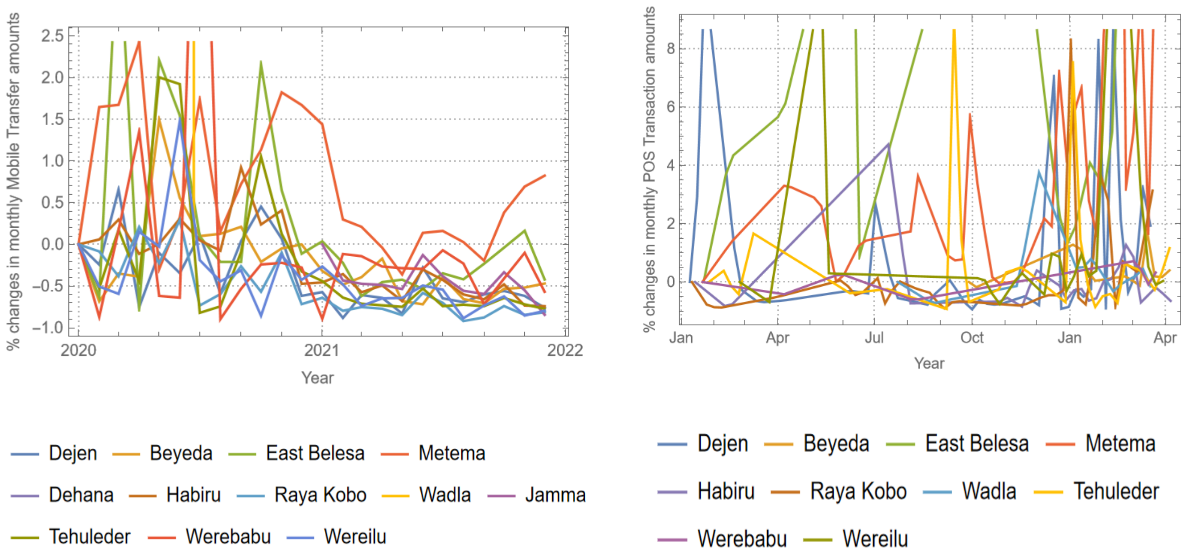

Figure 5 (a) trends in changes in Mobile transfer, (b) trends in percentage in POS transaction

The variability of these transactions can be attributed to different factors like internal conflicts and displacements, natural disasters, COVID-19, and others. Whatever the factors might be what is important for aid agencies is to look at into how financial capabilities of those communities trend over time. Relative comparison matters a lot for such decisions.

4.7 Fitting Transaction Data Points

Time series models are used to forecast the future based on previous observed or collected datasets at regular time intervals (Engineering Statistics Handbook, 2010). By using the TimeSeriesModelFit function, we try to fit the datasets with different process models (AR, MA, ARMA, ARIMA, SARIMA, SARMA, FARIM, and ARCH). Table 3 summarizes the quality of selected process models for each transaction type for sample branches. The quality of the process model is evaluated by looking at the values of AIC (Akaike information criterion), BIC (Bayesian information criterion), and SBC (Schwartz-bayes information criterion). Models with smaller AIC and BIC are considered the better models.

Branch	Criteria	Deposit	Withdraw	Acct to acct transfer	ATM withdrawal	Mobile Transfer	POS trans
Selected Model		FARIM	SARMA	SARMA	FARIM	ARMA	SARMA
	AIC	-121.783	-22.6525	11.869	-140.502	54.0882	112.934
	AICc	-118.583	-20.0811	15.069	-137.302	58.1935	115.134
	BIC	-114.206	-25.2849	9.38105	-133.958	60.245	114.689
Dejen	SBC	-119.427	-21.4745	14.2251	-138.146	57.6224	115.077
		FARIM	ARCH	AR	SARMA	SARMA	SARMA
	AIC	-41.8762	-71.4633	24.4644	56.737	90.1902	-47.3722
	AICc	-38.6762	-67.358	27.6644	59.3084	92.7616	-44.8922
	BIC	-42.6739	-65.6659	21.0774	54.6305	92.4991	-45.9967
Beyeda	SBC	-39.5201	-679291	26.8205	57.915	91.3682	-46.04
		SARMA	SARMA	SARMA	FARIM	SARMA	SARIMA
	AIC	30.347	-84.788	94.1364	-74.1634	6.66051	269.06
	AICc	32.9184	-82.2166	96.7078	-70.9634	9.86051	271.775
	BIC	27.7338	-83.9599	93.5034	-70.5708	5.42077	273.675
East Belesa	SBC	31.5251	-83.61	95.3144	-71.8073	9.01661	275.393
		FARIM	FARIM	SARMA	FARIM	ARMA	SARMA
	AIC	-58.4665	-82.2705	36.1437	-192.275	71.4305	171.741
	AICc	-55.2665	-78.1652	38.7151	-189.075	74.6305	173.944
Metema	BIC	-54.8644	-80.5385	36.8225	-181.408	75.5158	173.425
	SBC	-56.1104	-78.7363	37.3218	-189.918	73.7866	173.868

Table 3: Data fitting process models for selected branches

As can be seen from table 3, every transaction type fits different process models. FARIM (which represents an autoregressive fractionally integrated moving-average process) and SARMA (which represents weakly stationary seasonal autoregressive moving-average) models better fit the majority of datasets as compared to the other process models. The ups and downs (seasonality nature of most of the dataset is also captured by SARMA). Bank transaction types (first row of table 3) are distinctive of one another. Each of these transaction types are better to be represented as independent data models per bank branches.

As bank transaction datasets come in a stream every minutes, day, month, or year, there is seasonality in the trends. This is reflected on how majority of the transaction types fit with SARMA model than other models. These models are not same every month or year. By nature, bank transaction datasets come in a stream, which suggests continuous update on the type of models. The online learning (a modified form of supervised) machine learning paradigm is a good fit to serve this purpose. It gives more importance to recent datasets than older datasets. Figure 6 depicts its schema. According to this figure, as datasets come in a stream, new models should be trained based on the newly incoming datasets. Similarly, trend analysis of aid places should be done continuously, giving more emphasis to recent datasets. Unlike many traditional machine learning problems, poverty targeting problems need to be based on recent datasets and should be updated continuously.

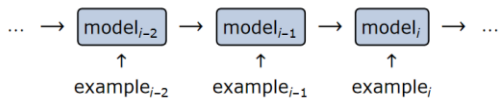

Figure 6: schema of online machine learning paradigm (source: Bernard pp. 18)

5 Conclusion

This study explored the use of machine learning/ artificial intelligence techniques to identify higher and lower level financial capability assessment that can help aid agencies and government agencies make better and informed decisions to prioritize and target beneficiaries for aid. Trend analysis of transactions at each bank branch serving the surrounding community (proxy approach) can reveal how wealth or poverty trends in a more granular manner. These datasets are automatically captured by the banks' IT infrastructure. They are accurate, captured in real time, detailed, and not manipulated directly to the benefit of others and thus they can be used to identify the neediest communities for aid. By developing the necessary data sharing policies with banks, governments and aid agencies can monitor poverty and wealth in real time across the nation.

6 Contribution

As stated in the literature transaction types like bank deposits (saving) and withdrawals can signify a lot about financial/ wealth capabilities of individuals. In this study, we found that trend analysis of individuals' bank account transaction datasets like deposit, withdrawals, and transfers can signify how poverty/ wealth trend at specific location or community. So far, these datasets have not been analysed and used for decision-making by either banks or other institutions.

7 Future Directions

In the future, we intend to scale up this approach. From all the banks and their branches operating in the nation, we will collect transaction datasets and develop a national digital poverty map accessible by governments and aid agencies on the web. The national digital poverty map will be updated every month through an online machine learning model. To make the map more valuable to users, we will develop an online machine learning model that feeds and updates the map every month for better decision making.

8 References

Abebe, R., & Goldner, K. 2018. Mechanism design for social good. *AI Matters*, *4*(3), 27-34.

Arendt-Cassetta, L. 2021. From digital promise to frontline practice: new and emerging technologies in humanitarian action. New York: OCHA.

Banerjee, A. V. 2006. The Economic Lives of the Poor Abhijit V. Banerjee and Esther Duflo October 2006. Africa. pp 164.

Bernard E. 2021. Introduction to Machine Learning. Champaign. Wolfram Media, Inc.

Blumenstock, J. E. 2016. Fighting poverty with data. *Science*, *353*(6301), 753-754.

Brei, M., Mohan, P., & Strobl, E. 2019. The impact of natural disasters on the banking sector: Evidence from hurricane strikes in the Caribbean. The Quarterly Review of Economics and Finance, 72, 232-239.

Clay, D. C., Molla, D., & Habtewold, D. 1999. Food aid targeting in Ethiopia: A study of who needs it and who gets it. Food policy, 24(4), 391-409.

Conning, J., & Kevane, M. 2002. Community-based targeting mechanisms for social safety nets: A critical review. *World development*, *30*(3), 375-394.

Data intensive development lab. Accessed June 2022. (https://didl.berkeley.edu/projects/, accessed on May 10, 2022)

Devereux, S., Masset, E., Sabates-Wheeler, R., Samson, M., te Lintelo, D., & Rivas, A. M. 2015. Evaluating the targeting effectiveness of social transfers: A literature review.

Eskander, S., Fankhauser, S., Jha, S., Batool, S., & Qaisrani, A. 2018. Do natural disasters change savings and employment choices? Evidence from Pakistan.

Eyre, R., De Luca, F., & Simini, F. 2020. Social media usage reveals recovery of small businesses after natural hazard events. Nature communications, 11(1), 1-10.

Foster, S. 2021. Alyans Peyizan Kafé D'Ayiti: A Copowerment Model for the Producers and Consumers of Coffee in Haiti.

Gallant, J. 2016. Financial Decision Making in the Foxhole: The effects of Conflict on saving Behavior in Kenya (Doctoral dissertation, Queen's University).

Gong D. 2021. (3 Common Techniques for Data Transformation | by Destin Gong | Towards Data Science, accessed on January 10, 2022)

Hernandez, K., & Roberts, T. 2020. Predictive analytics in humanitarian action: a preliminary mapping and analysis.

Hoos, K. 2010. Urban People: Too Poor to Save Money? The Case of Micro Deposit in the Philippines. Philippine quarterly of culture and society, 38(1), 75-83.)

Jaspers, S., & Shoham, J. 1999. Targeting the vulnerable: a review of the necessity and feasibility of targeting vulnerable households. Disasters, 23(4), 359-372.

Kaba, M. W., Baesel, K., Poch, B., Bun, S., Cerceau, S., Bury, L., ... & Rasanathan, K. 2018. IDPoor: a poverty identification programme that enables collaboration across sectors for maternal and child health in Cambodia. *bmj*, *363*.

Karlan, D., Ratan, A. L., & Zinman, J. 2014. Savings by and for the Poor: A Research Review and Agenda. Review of Income and Wealth, 60(1), 36-78.

Kryvasheyeu, Y., Chen, H., Obradovich, N., Moro, E., Van Hentenryck, P., Fowler, J., & Cebrian, M. 2016. Rapid assessment of disaster damage using social media activity. Science advances, 2(3), e1500779.

Lam P., Lyer S., Giraudy E., Pompe A., 2022. Facebook Business Activity Trends for Disaster Recovery. 272137999_3000199660246151_8348540010599272876_n.pdf (fbcdn.net)

Maas, P., Iyer, S., Gros, A., Park, W., McGorman, L., Nayak, C., & Dow, P. A. 2019. Facebook Disaster Maps: Aggregate Insights for Crisis Response & Recovery. In KDD (Vol. 19, No. 3173).

McCord, A. 2013. Community-based targeting in the Social Protection sector. Working Paper No. 514.

OCHA. HDX Accessed on May 2022. (Welcome - Humanitarian Data Exchange (humdata.org), accessed on April, 2022)

Raj, J. A. J., & Relton, A. 2013. Community Driven Development Approach as the New Avenue for Social Development-A Study With The Experience Of Vazhdhu Kaatuvom Project In Ramanathapuram District. *Indian Social Science Journal*, *2*(2), 1.

Reyers, M. 2019. Financial capability and emergency savings among South Africans living above and below the poverty line. International Journal of Consumer Studies, 43(4), 335-347.

Schnitzer, P. 2019. How to target households in adaptive social protection systems? Evidence from humanitarian and development approaches in Niger. The Journal of Development Studies, 55(sup1), 75-90.

Schwartz, T. 2019. A model for humanitarian Aid Beneficiary Targeting. Accessed on June 2022. A Model for Humanitarian Aid Beneficiary Targeting - Schwartz Research Group (timothyschwartzhaiti.com) , accessed on March 20, 2022)

Schwartz, T., Maass, H., & DeMattee, A. Coordination National de la Securite Alimentaire

Serajuddin, U., Uematsu, H., Wieser, C., Yoshida, N., & Dabalen, A. 2015. Data deprivation: another deprivation to end. *World Bank policy research working paper*, (7252).

Sharp, K. 1998. Between relief and development: targeting food aid for disaster prevention in Ethiopia. Overseas Development Institute.

Tkachenko N, Jarvis S, Procter R 2017. Predicting floods with Flickr tags. PLoS ONE 12(2): e0172870. (https://doi.org/10.1371/journal.pone.0172870, accessed on Dec 23, 2021)

World Food Program Kenya n.d. Assessment of the Geographical and community-based targeting of WFP's Cash and food for asset program in Kenya. Accessed on July 2022. WFP-UNICEF-Assessment-of-the-Geog-and-Community-based-Targeting-of-Asset-Creation-Programme_0.pdf (alnap.org), accessed on May 20, 2022)

Zou, L., Lam, N. S., Cai, H., & Qiang, Y. 2018. Mining Twitter data for improved understanding of disaster resilience. Annals of the American Association of Geographers, 108(5), 1422-1441.

Acknowledgement

We would like to express our gratitude to scholarship funders - International Development Research Centre (IDRC) and Swedish International Development Cooperation Agency (SIDA), under scholarship programme- Artificial Intelligence for Development (AI4D) Africa, and scholarship fund manager- Africa Centre for Technology Studies (ACTS).

Copyright