SCALING UP SOCIAL ASSISTANCE WHERE DATA IS SCARCE

Opportunities and limits of novel data and AI

WORLD BANK GROUP
Social Protection & Jobs

G2Px
PAYMENTS THAT EMPOWER

1818 H Street NW
Washington DC 20433
Telephone: 202-473-1000
Internet: www.worldbank.org

SCALING UP SOCIAL ASSISTANCE WHERE DATA IS SCARCE

Opportunities and limits of novel data and AI

This paper was written by Yuko Okamura, Tim Ohlenburg and Emil Tesliuc. The authors gratefully acknowledge comments and inputs from Minita Varghese, Anit Mukherjee, Laura Bermeo, Tina George, Ana Lucia Cardenas, Fanen Ade, Maheshwor Shrestha, Ubah Thomas Ubah, Muderis Abdulahi Mohammed, Arthur Alik Lagrange, Mitja Del Bono, Steisianasari Mileiva, Julian Koschorke, Sebastian Geschwind, Luis Iñaki Alberro Encinas, Oscar Barriga-Cabanillas, Andrés Parrado, and Tom Mtenje. This paper is being cross-published in both Social Protection & Jobs Discussion Paper Series and G2Px Paper Series, given that this paper was produced in partnership between Social Protection and Jobs Global Practice and the G2Px initiative to produce a collection of papers on the use of novel data for social assistance delivery. The authors are also grateful for guidance from Loli Arribas-Banos and Vyjayanti Desai. The paper benefited from the peer review from the following colleagues: Phillippe Leite, Sveta Milusheva, Jonathan Marskell, and Robert Palacios. Lara Srivastava provided editorial assistance, and Jihane El Khoury Roederer provided design and layout assistance. For comments and queries, the authors can be reached at yokamura@worldbank.org, mail@timohlenburg.com, and etesliuc@worldbank.org.

During the recent Covid-19 shock (2020/21), most countries used cash transfers to protect the livelihoods of those affected by the pandemic or by restrictions on mobility or economic activities, including the poor and vulnerable. While a large majority of countries mobilized existing programs and/or administrative databases to expand support to new beneficiaries, countries without such programs or databases were severely limited in their capacity to respond. Leveraging the Covid-19 shock as an opportunity to leapfrog and innovate, various low-income countries used new sources of data and computational methods to rapidly develop -level welfare-targeted programs. This paper reviews both crisis-time programs and regular social protection operations to distill lessons that could be applicable for both contexts. It examines three programs from the Democratic Republic of Congo, Togo, and Nigeria that used geospatial and mobile phone usage data and/or artificial intelligence (AI), particularly machine learning methods to estimate the welfare of applicants for individual-level welfare targeting and deliver emergency cash transfers in response to the pandemic. Additionally, it reviews two post-pandemic programs, in Lome, Togo and in rural Lilongwe, Malawi, that incorporated those innovations into the more traditional delivery infrastructure and expanded their monitoring and evaluation framework. The rationale, key achievements, and main challenges of the various approaches are considered, and cases from other countries, as well as innovations beyond targeting, are taken into account. The paper concludes with policy recommendations and promising research topics to inform the discourse on leveraging novel data sources and estimation methods for improved social assistance in and beyond emergency settings.

Contents

1. Introduction

Social protection is at a threshold of transformative change propelled by rapid digitalization, new technologies, and the pressing need for adaptive and universal systems. While the use of digital technologies and administrative data has become a norm in social protection, the application of new types of data (i.e., novel data sources,[1] such as satellite and mobile data) is more recent. The rise in these novel approaches over the past few years[2] is due to a confluence of several factors: increased availability of data over time, the evolution of computation methods and data analytics (e.g., machine learning, artificial intelligence), and the need for urgent social protection scale-up to respond to the COVID-19 pandemic.

In response to the COVID-19 crisis, social assistance was expanded at an unprecedented scale, reaching many of those who had previously been left out of any government support. An estimated 1.7 billion people in low- and middle-income countries (roughly one in five people in the world, across most regions) received COVID-response social assistance payments, over half of them for the first time (Marin and Palacios 2022).[3]

During COVID-19, registering and selecting new beneficiaries proved to be a significant challenge in delivering emergency support to a larger segment of the population not covered by social protection programs. This was primarily due to the lack of data: only half of the global population was being covered by social protection at the time (ILO 2021), leaving half of the population, especially those working in the informal sector, "invisible". Informal workers were not covered by any social protection schemes, were absent from any administrative databases, and consequently more vulnerable to the economic shock of COVID-19 without automatic access to any protection mechanisms.

The challenge of registering applicants and selecting beneficiaries is most pronounced in "data desert" countries that lack the traditional infrastructure, data, and digital delivery systems. Designing and implementing cash transfers requires data to identify, locate, screen, and pay beneficiaries (e.g., administrative records and/or data collected from applicants through socio-economic surveys), as well as digital public infrastructure such as an ID system, social registry, management information

1 We use the terms 'non-traditional data' and 'novel data' interchangeably to refer to the various data sources. See Aiken and Ohlenburg (2023) for an overview.

2 See Lowe, 2022 and Ohlenburg, 2022 for reviews.

3 Estimate based on payments made as of May 2021. Recipients are measured as estimated individuals living in households that received a cash transfer, regardless of whether the program targeted individuals or households.

systems, and a payment modality.[4] For example, a social registry plays a critical and enabling role in assessing eligibility for social assistance, identifying individuals or households in need, based on the eligibility criteria such as welfare estimates (poverty) or demographic characteristics. In responding to COVID-19, countries with social registries with up-to-date information and high coverage were able to deliver emergency support to their target population quickly and effectively (e.g., in Turkey, Chile, Brazil, Jordan, Malaysia).

In this context, a small number of countries without pre-existing social protection systems and tools to estimate the welfare of target populations turned to novel data sources, estimation techniques and digital delivery. While most countries lacking the tools to estimate the welfare of their target group resorted to categorical targeting (e.g., the list of employees of affected industries),[5] the countries surveyed in this paper rapidly developed and adopted welfare-based criteria using novel data and estimation techniques. These innovative approaches are the focus of this paper.

Although the use of non-traditional approaches for individual and household targeting is still in its infancy, examples from around the world are emerging and their feasibility is increasingly being recognized. There are a few notable examples on the African continent. Togo's Novissi Model 2[6] and the Democratic Republic of Congo (DRC)'s STEP-KIN are perhaps the best publicized of these applications. There are three additional use cases which followed the pandemic: Malawi pilot which replicated the same approach as Togo's Novissi Model 2; the Togo Social Safety Net and Basic Service Project which is aiming to combine both novel and traditional approaches; and the Nigeria National Social Safety Nets Program – Scale Up (NASSP SU)[7] project which started using a new registry collected through mobile phones. In addition, similar programs are already underway in some African countries or are at an initial stage of policy discussion in other African countries (e.g., Cote d'Ivoire, Mali, Cameroon).

While these examples have yielded some preliminary results, solid evidence of the effectiveness and impact of such approaches is still missing. For example, rigorous evaluations of the comparative performance of novel versus traditional targeting approaches have yet to be carried out, though some are underway at the time of writing. Alternative approaches come with many unknowns but can be suited to emergency situations when speed is more vital than accuracy. In contrast, for regular poverty-targeted social assistance programs under normal circumstances, inclusion and exclusion errors remain critical performance metrics.

The objective of this paper is to distill lessons from recent country case studies and develop guidance on leveraging new forms of data for delivering social assistance in and beyond emergency settings. Novel sources of data, when used to complement conventional approaches and data, have

---

4 These supporting systems are increasingly referred to as Digital Public Infrastructure (DPI). DPI can play a critical role for effective public administration, service delivery, and innovation across multiple sectors. For government-to-person (G2P) payments programs, systems for digital identity verification or authentication, digital payment systems, as well as structures that facilitate data exchange between these and other sectoral databases and applications, are key elements of a modern G2P architecture. Combined, these platforms can work together as a "stack" to support service delivery and a dynamic digital economy. For more information see Metz et al. 2022. "A Digital Stack for Transforming Service Delivery ID-Payments and Data Sharing" World Bank ID4D.

5 Despite the lack of existing registry, some countries succeeded in implementing cash transfers anew, while relying on more traditional approaches and categorical targeting without welfare assessment (ex. Bangladesh and Sierra Leone using the employees' list of the affected industry.

6 This paper focuses on Novissi Model 2 which included ML/AI targeting innovations to expand coverage to rural areas, instead of Model 1 which was purely based on categorical targeting using admin data sources and was implemented in urban areas.

7 For Nigeria NASSP SU, this paper focuses on the use of Rapid Social Registry (RRR), instead of National Social Registry (NSR)

the potential to make social protection systems more adaptive, inclusive, resilient, and efficient. Exploring innovations that support faster responses to crises, such as global pandemics, will help countries tackle and prepare for the myriad of challenges linked to crisis and pandemic preparedness, climate change, economic instability, fragility, food insecurity, war, and displacement.

The paper aims to answer practical operational questions regarding the rationale behind the use of novel data sources for identifying social assistance beneficiaries, as well as the implications of such practices for the future. Three specific country cases are examined in depth, all of which used satellite and mobile data for individual or household level targeting: Togo Novissi Model 2, DRC STEP-KIN, and Nigeria NASSP SU. The next chapter discusses each of the three case studies, their innovation context, key challenges and drawbacks. It also highlights use cases from other countries where innovation has taken place in other parts of the social assistance delivery chain, beyond the targeting process. Chapter 3 distills some preliminary results. Chapter 4 presents two additional cases which adopted non-traditional targeting approaches to deliver assistance post pandemic. Finally, Chapter 5 examines the applicability of such approaches for both regular and emergency situations and highlights existing limitations and unknowns. The paper concludes with a set of recommendations, highlighting areas for future research.

2. Overview of novel approaches to social assistance

This chapter provides an overview of the use of novel data sources for social assistance in three specific cases as follows:

- Togo's *Novissi Model 2*, which started paying out benefits to 139,000 beneficiaries in November 2020, covering more than half of the cantons in the country;
- DRC's *STEP-KIN*, which was able to deliver a first payment in March 2021 to a total 456,000 beneficiaries residing in selected areas of the country's capital; and
- Nigeria's *NASSP SU* selected 636,000 beneficiaries using the registry collected through mobile phones, and paid them in November 2023, three years after the onset of COVID-19.

Table 1 below sets out the timing of the first payment in the three case study programs.

TABLE 1. Timeline of Three Case Study Programs: First Payment by Phase (Number of beneficiaries)

	2020		2021												2022				2023			
	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec				Dec				Nov
Togo Novissi Model 2	100 poorest cantons					Next 100 poorest cantons (total of about 139k)																
DRC STEP KIN					100k								Additional 200k					Additional 250k				
Nigeria NASSP SU (using RRR)																						636k

Source: Author's compilation
Note: Phase 1—Blue; Phase 2—purple; Phase 3— Green

The application of novel data sources in the three countries varied from simple to sophisticated. DRC's STEP-KIN only used simple criteria based on the use of mobile data for program eligibility: this meant that a program operating as a first-come, first-served basis (without any further prioritization) was sufficient. In contrast, Togo's Novissi Model 2 used a more sophisticated targeting mechanism based on machine learning to prioritize the poorest areas and individuals. Nigeria employed a similar process as Togo to develop a poverty map. However, this map was never used for geographic targeting; rather a mobile phone-based outreach and registration strategy helped create "rapid response registry (RRR)" to bridge coverage gaps in the existing social registry. The RRR was used to select beneficiaries only in late 2023. Table 2 below summarizes the salient design features and data points of all three programs.

TABLE 2. Overview of DRC STEP-KIN, Togo Novissi model 2 and Nigeria NASSP SU Cases

	DRC STEP-KIN	Togo NOVISSI Model 2	Nigeria NASSP SU
Implementation agency	Social Fund	Ministry of Digital Economy and Transformation	Ministry of Humanitarian Affairs and Poverty Alleviations (FMHAPA)
Target beneficiaries	300,000 individuals		8.2 million households[8] (selected through both traditional and new registries/approaches)
First payment	March 2021	November 2020	November 2023
Objective (target population)	To provide cash transfers to the vulnerable in the capital affected by the pandemic	To provide cash transfers the poorest rural population affected by the pandemic	To provide cash transfers to urban and rural poor in selected wards
Geographic coverage	**Capital** only (71 most vulnerable neighborhoods out of 369)	**Rural** (200 poorest cantons out of 389 cantons)	All local government areas but with disproportionate focus on urban areas (61% of beneficiaries)
Geographic assessment	**Hotspot (overcrowding) and vulnerability map** (based on satellite imagery, flood data, and commuter survey)	**Poverty map** (based on satellite imagery, population census data and Machine Learning (ML))	**Poverty map** (based on satellite imagery, Demographic and Health Survey (DHS) data and ML)
Individual assessment	Five **exclusion criteria** using anonymized Call Detail Record or CDR (no machine learning, no score)	**Welfare assessment** (score) based on machine learning and anonymized CDR data	**Welfare assessment** (score, adjusted Proxy Means Test or PMT) using the National Social Registry and Rapid Response Registry data[9]
Key partner(s) for individual targeting	Mobile Network Operator (MNO) to discuss/propose criteria based on their data, World Bank to provide Technical Assistance (TA)	MNO to share anonymized data, academia to conduct the analysis	World Bank to provide TA to produce proxy welfare indicator
Verification of ID	None	Voter ID	None[10]
Registration	On-demand self-registration via USSD	On-demand self-registration via USSD	On-demand self-registration via USSD and household visit

8 The planned project seeks to reach over 8 million households (5 million households in urban and 3.2 million in rural areas) from 2023 to 2024, prioritizing poorer wards based on the granular poverty maps.

9 Nigeria Rapid Social Registry (RRR) consists of information submitted by applicants through self-registration platform as well as through in-person household visit which collects about 42 variables (which corresponds to about a quarter of the National Social Registry's size).

10 The bank KYC requirements varied based on the tier of the account.

	DRC STEP-KIN	Togo NOVISSI Model 2	Nigeria NASSP SU
Mobile numbers contacted	370,000 (of 2.3 million eligible numbers)	-	68 million
Number of registrants	315,000 individuals	519,972 individuals	6.9 million
Number of beneficiaries	277,000 individuals	138,531 individuals	2 million households (635,649 selected through RRR; the rest through NSR)
(% female)	(36%)	(52%)	(50%)
(% with existing mobile accounts)	(42%)	–	–[11]
Benefit amount	25 dollars per installment (150 dollars per beneficiary)	Women: 8,170 CFA (13 dollars) per installment Men: 7,000 CFA per installment	N 5,000 (12 dollars) per installment (plan) (while the first beneficiaries received one-off payment of N25,000)
Payment frequency	Monthly	Monthly (100 poorest cantons); bi-weekly (next 100 poorest cantons)	Monthly (plan) (One-off payment for the first beneficiaries)
Duration of benefit	6 months from February 2021 – June 2022	5 months from November 2020 – August 2021 (100 poorest cantons) 2.5 months (next 100 poorest cantons)	6 months (plan) (while the first beneficiaries received one-off payment in November 2023)
Payment mode	Mobile money account	Mobile money account	Mobile money account
Budget	$45 million	$10 million	$800 million
(% Share of operational cost)	(5.7%)		(6.6%)
Source of Funds	World Bank	GiveDirectly	World Bank

Source: Authors' compilation

2.1. EXPLORING THE USE OF NON-TRADITIONAL OR NOVEL DATA SOURCES

This chapter takes a closer look at the targeting approaches used in each case study. It begins with a discussion of the relevance of novel data sources for social assistance and the advantages these have over traditional approaches. It then examines the country conditions and requirements necessary for using such sources to deliver emergency support.

2.1.1 Relevance and benefits

Some countries have turned to novel data sources for delivering social assistance due to a lack of access to traditional approaches, limited time, operational challenges, and data constraints. The design and implementation of cash transfer programs typically require various data points to determine program beneficiaries, and key infrastructure such as an ID system, social registry, and payment system.

11 All RRR registrants already had accounts upon the registrations with their phone number. NSR registrants with existing accounts were selected for the first payments.

Official statistics such as census and household income/expenditure surveys are critical for identifying poorer areas and households. In turn, social registries are needed to determine which households are eligible for programs, through the application of welfare estimates and categorical variables. Given the lack of household survey data or registries with adequately high coverage in an emergency context like COVID-19, some countries had to shift their approach to social assistance, particularly with respect to geographic and individual targeting.

As can be expected, countries with higher quality digital public infrastructure were in a better position to deliver emergency responses during the pandemic. The coverage rate of emergency cash transfers was closely associated with the income level of the country,[12] which in turn was correlated with the level of existing infrastructure. Countries that were able to use existing databases reached an average of 51% of their population, whereas those without reached only 16% (Marin and Palacios 2022).[13] On average, it was much faster to make emergency payments to additional beneficiaries where there was already a social registry in place (Gentilini et al, 2020, 2021, and 2022).[14] In the absence of a social registry, reaching beneficiaries proved to be a much longer process.

Therefore, while continued investment in conventional enablers is advisable in the first instance, innovative and complementary strategies can help overcome some existing limitations and enhance program responsiveness and inclusiveness. Despite progress made over the past few decades, the reality is that many countries lack the necessary enablers for social assistance delivery. It takes continued effort and time to develop dynamic and inclusive social registries which can respond quickly and fully to shocks and crises, and many countries still have long way to go. This is particularly the case with low-income countries, some of which have not yet established any social registry.

The alternative approaches discussed here can be powerful tools not only in emergencies but also in fragile contexts in which investing in traditional infrastructure is less feasible. Nearly one in every five countries is characterized by fragility, conflict, and violence (FCV), without any improvements over the past decade (World Bank, 2009, 2023). Out of 37 FCV countries, 17 are affected by conflict and the remaining 20 countries exhibit institutional and social fragility due to very limited social infrastructure or systems (e.g., small island states), underscoring the relevance of using alternative approaches for social assistance.

In summary, novel data sources for social protection fall into three main categories. The first is as an original information source that enables new approaches to program provision. A social registry with household geo-location combined with satellite data is an example of the emerging frontier of disaster response in social protection systems (see Sundqvist (2023) for the case of Malawi). The second category is as a substitute to existing sources where those are unavailable or of insufficient quality. The case studies examined in this paper fall into this category. The third is as a complement to traditional data to achieve better outcomes. For instance, inclusion of mobile phone usage via call

12 Based on the Global Tracker on Social Protection Responses to COVID (Gentilini et al, 2020, 2021, and 2022), high-income countries provided cash transfers to half of the population (44 percent), while only 8 percent benefited in low-income countries where existing systems were much more limited

13 Research on the social protection response to the pandemic in 85 countries found that those that were able to leverage digital databases or ID records and data-sharing platforms reached over three time more beneficiaries than countries that had to collect anew (Marin, Palacios 2022).

14 For example, the programs which had to call for applications to collect data from interested new beneficiaries took more than one month on average to transfer payments.

detail records in Proxy Means Tests (PMT) may raise targeting accuracy in a setting (Pinxten, 2020). Depending on the context, novel data sources can make a variety of important contributions.

2.1.2 Requirements and bottlenecks

Requirements necessary for successfully adopting non-conventional approaches for social assistance include political leadership and commitment, sufficient mobile penetration, and strategic partnerships with mobile operators. Political commitment is critical to adoption and tends to reduce the influence and discretion of local implementers. Strong political leadership was vital to the launch and success of Novissi 2 in Togo and STEP-KIN in the DRC. Togo's President galvanized different stakeholders, set up an inter-ministerial Steering Committee, and secured resources needed for the program. In Nigeria, the government chose not to target specific geographic areas, for fear that this may create a perception of preferential treatment, but rather opted to send mass mobile messages to a much wider set of locations, well outside those identified as priority areas by poverty maps (based on satellite imagery). This example illustrates the political economy challenge related to geographic targeting.

Second, for phone-based approaches, the mobile penetration rate needs to be sufficiently high, nationwide, or at least in those geographic areas selected for the program. This is one of the first check points at the design stage, as mobile phones are the basis for program inclusion and communication. The three country cases reviewed here (DRC, Togo, and Nigeria) benefited from the relatively high mobile phone penetration rates to deliver social assistance. DRC's STEP-KIN specifically focused on the capital where more than 90 percent of the population owned mobile phones (Mukherjee et al 2023). In Nigeria, the mobile penetration rate was 85 percent[15] and in Togo, 85 percent of the population was estimated to belong to a household with at least one mobile phone.[16]

Third, it is essential that governments involve MNOs as key implementation partners, without whom the use of mobile data is not feasible. Specifically, MNOs need to allow their data to be used to identify and reach the mobile phones of potential beneficiaries. Access to specific personal data is not required, and there are of course privacy and data protection safeguards to be considered (discussed later in this section). In the Togo and DRC country cases reviewed here, building the relationship with MNOs was a learning process, as the public and private sectors had limited experience of working together on this kind of project. Ideally, the partner MNOs should have a high collective market share, as this will maximize reach. In the case of multiple MNOs partners, the programs would require a mechanism to detect individuals with multiple SIM cards, to avoid double coverage or dipping.[17]

Other conditions, such as the availability of specific types of data to facilitate targeting, can be easily met in most countries.[18] For the purpose of developing a geo-spatial map using satellite imagery, low-resolution data is available globally. What is often limited is ground truth data (i.e.,

15 As discussed in the final section, such less-than-complete coverage raises policy issues.

16 Source: *Enquête Harmonisée des Conditions de Vie des Ménages* (EHCVM) 2018-2019, Togo

17 Ideally, a unique identifier of high coverage could have detected such duplications, which was not, however, the option for Togo and DRC when they conceived Novissi and STEP-KIN, respectively. As a result, the verification was limited to check applicants' names on an ad-hoc basis. Togo had recently conducted a voter registration that covered 95% of adults and was biometrically deduplicated was a key to success.

18 In the same vein, the existence of payment service providers (PSP) is another condition which is practically met in many settings. In most countries, certain PSP(s) should exist and can be made available for social transfers, regardless of conventional or non-conventional approaches. While digital payments (including mobile wallets) can expand payment choice and improve efficiency, they are not necessarily a prerequisite as the programs can use manual payments, or more traditional services from the financial sector.

welfare information provided by direct observation, such as household data from existing surveys or from primary collection methods) for training more detailed models. When data is available, it can often be of lower quality or outdated, particularly in low-income and FCV countries. However, Chi et al (2022) demonstrate that even moderately precise estimates can be made in all low- and middle-income countries using Demographic and Health Surveys (DHS) as ground truth data. These provide a relative wealth index that is different from consumption (the welfare metric measured by the World Bank's Living Standards Measurement Surveys, LSMS[19]).

In addition to the above, legal and regulatory safeguards need to be put in place as part of program preparation.[20] Mobile data should be collected and processed in accordance with national law including applicable data protection and privacy law. When partnering with international carriers, depending on their governance structure and the jurisdiction of their headquarters, additional legal requirements (e.g., based on international standards) can be an additional hurdle for low- and middle-income countries.

Program implementers and MNOs need to specify and agree on the terms and conditions for using mobile data in social assistance programs. While cash transfer programs can be designed and implemented without personally identifiable mobile data being shared with external parties, in some cases non-disclosure agreements (NDAs) need to be signed between MNOs and external parties involved in program delivery, to ensure that data access and processing is carried out in accordance with data protection and privacy law. In the cases of Togo, a first NDA was signed with the academics who developed the prioritization algorithm: this was done with the consent of subscribers who were interviewed for the survey. A second NDA was signed with the government implementation agency, which needed to cross-check whether the mobile numbers of program applicants appeared in the MNO's list of eligible numbers. Once again, this was done based on the applicant's consent during the application process. Any agreements with MNOs and external parties must cover requisite data protection and privacy aspects, such as, *inter alia*, the objective of the data use (e.g., the legitimacy of the objectives, the scope of analysis), where the data is stored, what kind of security measures are in place to safeguard the data, who has access to the data (with different authorization levels), and how the results of data processing will be used.

2.2. TARGETING PROCESSES USING NON-TRADITIONAL SOURCES OF DATA

When the COVID-19 shock hit DRC, Nigeria, and Togo, policymakers in these countries sought solutions to provide a temporary minimum income for vulnerable segments of their populations. This support was important for those who were already poor before the shock, as well as those who became impoverished due to the COVID-19 pandemic and associated mobility and employment restrictions. However, the scale of the shock was too large to be mitigated through universal income supportgiven that the fiscal resources of the three countries were constrained. As a result, governments faced a

19 The Living Standards Measurement Study (LSMS) is the World Bank's flagship household survey program focused on strengthening household survey systems in client countries and on improving the quality of microdata to better inform development policies. See https://www.worldbank.org/en/programs/lsms.

20 There are also general issues related to the policy framework which are not limited to the programs using innovative approaches. For instance, Know Your Customer (KYC) requirements are a typical bottleneck in financial services for inclusion of poor and vulnerable populations. The study countries revisited KYC policy to relax/remove the requirements to submit the ID to open an account for emergency transfers. These changes may need to be reviewed to re-balance anti-money laundering considerations against financial inclusion.

choice between a) providing a smaller benefit to the entire population or b) offering a larger benefit but only to the poor and vulnerable. The universal approach of the first option would have resulted in much smaller benefits, with little impact on the livelihoods of the poor and vulnerable. The second option had the advantage of delivering a substantial benefit that could significantly mitigate income losses due to the shock.

However, choosing the second option comes with its own set of challenges and is only feasible when there is an effective method to identify the intended (poor and vulnerable) target group. During normal times, when governments are not facing crises, the selection of a targeting method involves the assessment of various options, weighing advantages and disadvantages, and possibly combining targeting methods. Common options include group-based methods (such as geographical and categorical targeting), economic welfare-based approaches (such as community-based targeting, proxy means testing, or means testing), and even self-targeting or random selection (lottery). Each targeting method can only be successful when certain minimum conditions are met (Grosh et al., 2022).[21] For example, means testing is not effective when there is a high level of informality and a lack of administrative records on income, and community-based targeting is not viable in the presence of mobility restrictions (such as those imposed during the pandemic).

In the cases of DRC, Nigeria, and Togo, the chosen solution involved a combination of geographical targeting and welfare-based targeting. All three countries adopted a two-step approach: first, geographical targeting was used to focus interventions on specific areas with concentrated poverty, and second, individual level targeting was used to allocate resources to those in greatest need. In Togo and the DRC, a form of proxy-means or affluence testing based on available data was employed.[22] While the solutions adopted by these three countries were not unique compared to other middle- or low-income countries, design and implementation constraints were greater than in most other countries. As such, it proved difficult to implement a targeted program using established approaches rapidly and cost-effectively. In each of these countries, the targeting methods had to be adapted to suit local data constraints, deviating from standard, first-best methodologies. This process resulted in several important innovations in the field. The following sections describe how these countries addressed the constraints of their contexts and the specific targeting strategies they employed. It is important to note that satellite data was used in all three programs to determine and select poor areas. In Togo and the DRC, a phone-based selection process further refined the selection of poor individuals.

2.2.1 Geographic targeting

Implicitly or explicitly, geographic targeting relies on maps that estimate poverty rates based on income or consumption at a relatively small geographical scale. To create these maps, researchers combine census data with surveys that capture the income or consumption levels of households in the sample, as described by Elbers, Lanjouw, and Lanjouw (2003). To gather the most accurate data, these surveys often employ detailed lists of consumption items or income sources. Econometric models are then applied to calculate poverty rates for smaller areas, such as districts, parishes, or municipalities, which offer a much higher level of geographic detail. Poverty maps complement other targeting methods and have been extensively applied to allocate services to impoverished areas. Ex

21 See Annex for more details on appropriateness (pros and cons), minimum conditions, shock responsiveness of different targeting methods.

22 Many countries did collect new data in person, including some like Philippines that did have a high coverage social registry before the pandemic. Such a registry was unavailable in DRC or Togo.

post evaluations of these poverty maps have demonstrated their reliability as predictors of regional poverty at the level of small geographic units.

Although all three countries used geographic targeting to focus their interventions on areas of concentrated poverty, due to data constraints none of them were able to develop a consumption-based poverty map using the small area estimation methodology. The most recent census was conducted in 1984 in DRC, in 2006 in Nigeria, and in 2010/11 in Togo. At the same time, since the last census, population was estimated to have increased by more than three times in the DRC, by 48 percent in Nigeria and by 30 percent in Togo.[23] Household surveys with a good consumption module are more recent: 2018 in DRC, but limited to the capital Kinshasa, 2018/19 in Nigeria[24] and 2018 in Togo. In addition, like in other parts of the world, it proved very difficult to track or estimate the welfare of households due to the COVID-19 shock. To identify households impoverished due to COVID-19, countries would have required information on who was affected by the shock and to what extent, which was not available. The countries reviewed here turned to innovative proxies to circumvent these data problems.

In the DRC, where data constraints were the highest and the impact of COVID-19 deemed most severe, the government chose to focus on the capital Kinshasa. To identify potential areas for intervention, a hotspot and vulnerability map was developed combining the identification of potential COVID-19 hotspots with a vulnerability score as a proxy for poverty (Mukherjee et al 2023). Potential hotspots were estimated based on two measures of crowding—population density and estimated number of people per unit of floor space—using satellite imagery with a 100x100 meter grid. To prioritize assistance within potential hotspots, a vulnerability score was estimated based on the following four indicators: precarious construction, density of buildings, risk of flooding and a measure of access to jobs. The precariousness of buildings and their density was estimated based on high-resolution satellite imagery (Batana et al, 2021), and the risk of flooding on pluvial hazard and flash flood data (Fathom 2020). Access to jobs was proxied by travel time derived from a commuter travel survey combined with high-resolution satellite imagery (He et al, 2021).

The selection of areas for the STEP-KIN cash transfer program included the 20 percent most vulnerable areas among the COVID-19 hotspot areas. These areas were overlaid with administrative neighborhoods to identify the geographical target of the STEP-KIN programs: 71 neighborhoods out of a total of 369, covering one-sixth of the area of Kinshasa and containing 43 percent of its estimated population were selected. The 71 neighborhoods were further divided into four zones, with an intention to allocate one zone to each of the four potential MNOs. Allocating one zone per partner was intended to minimize double dipping for those with multiple SIM cards from multiple MNOs. A first-come, first-served approach was then employed to deliver benefits.

For Togo's Novissi Model 2, the government partnered with a group of academics to estimate the poverty rates of the 397 cantons using a range of survey data and non-traditional data (see Figure 1). As in the DRC, neither a recent poverty map nor any poverty estimates, at this level of disaggregation, were available. The non-traditional data used to generate the Togo poverty map included an array of sources including satellite data (high-resolution imagery and nightlights), connectivity data

23 The population in DRC increased from 30 million in 1984 to 96 million in 2021; from 144 million in 2006 to 213 million in 2021 in Nigeria; from 66/67 in 2010/11 to 86 million in 2021 in Togo (World Bank Open Data https://data.worldbank.org/).

24 Nigeria Living Standard Survey (NLLS)

(characteristics on telecommunications infrastructure, like cell towers and Wi-Fi access points, and mobile device features), demographic features (high-resolution population density maps, urban-rural domain markers) and geographic properties (road density, elevation, and precipitation). A total of 112 features derived from such non-traditional data were then matched to ground truth data to find patterns of poverty in geospatial data and calibrate a model to predict consumption in those grid cells or tiles without ground truth data. This methodology (as described in Chi et al. 2021, Aiken et al. 2021, and Lawson et al. forthcoming) ranked Togo's 397 districts from the poorest to the richest (see Figure 1). The 200 poorest cantons were selected to receive benefits as part of the expansion of the program to rural areas. The prioritization of the poorest cantons followed a four-step process, as set out below:

1. The national household survey of 2018-2019 was used to estimate the average consumption of households within 2.4 km2 grid cells. The average was based on the consumption reported in the survey and the approximate geographic coordinates of the 6,172 HHs in the sample. This indicator was then used as the target variable for a machine learning (ML) prediction model.
2. The average consumption of households in each 2.4km2 grid cell was then linked to a range of contemporaneous non-traditional data records to calibrate the ML prediction model for all the tiles covered by the survey.
3. The ML prediction model was then used to estimate the average consumption for all grid tiles in Togo, based on the most recent available records. The result was a map with the estimated average daily consumption per capita at the grid cell level as of 2021.
4. To identify the poorest cantons, the estimates of consumption per grid cell were combined with population density estimates. Using high-resolution population density estimates from the Humanitarian Data Exchange, the team of academics calculated a weighted average of consumption for all 2.4 km2 grid cells within each canton. This allowed cantons to be ranked by their predicted degree of consumption as well as their population density, giving more prominence to cantons with a large population estimated to be poor.

FIGURE 1. Prioritizing the poorest cantons in Togo (Novissi Model 2)

Source: Lawson et al. 2023

In Nigeria, a similar process was supported by the same group of researchers.[25] Machine learning algorithms were developed to predict wealth indicators based on high-resolution satellite imagery data. The algorithms were trained on geocoded Demographics and Household Survey (DHS) data from 2018 which covered 14 percent of the country's wards (i.e., Admin 3 units) and used the relative wealth index of the DHS as the key wealth metric. The algorithms were then applied to the entirety of the country to produce 159,000 high-resolution poverty maps which were then aggregated up to ward (8,808), Local Government Area (LGA) (774), and state (37) levels using approximate population weights. The ML-generated wealth estimates at these different levels were then validated using independently collected ground truth data from the 2018/2019 Nigeria Living Standards Survey (NLSS).[26] The validation confirmed that the satellite image-based estimates were sufficiently accurate to function as a standalone targeting method (see Chapter 3 for more details on the comparative results).

In summary, to mitigate the lack of recent census and household surveys, all three countries employed recent high-resolution geospatial data to estimate poverty levels in small areas. Geospatial data can capture features that are correlated with average welfare levels, while mobile data correlate with individual well-being (inequality in welfare levels). As mentioned above, DRC used five variables derived from a theoretical model of sources of poverty and vulnerability in Kinshasa: overcrowding, precarious constructions, density of buildings, risk of flooding and access to jobs. In contrast, Togo and Nigeria used a large matrix of 112 such variables. Not unlike small area estimation techniques for poverty mapping, Togo modelled the pre-crisis household consumption based on CDR data (pre-crisis, contemporaneous) and the geospatial features of the small areas. To generate these estimates, both Nigeria and Togo used machine learning algorithms trained on areas that overlapped with the primary sampling units of the household survey. This was then used to predict small-area poverty for those grid cells or tiles without household information. In Nigeria, an independently collected survey (not for the purpose of training the machine-learning model) together with the 2018/2019 NLSS were used to validate the out-of-sample accuracy of the model. This helped build confidence across the country in the use of geographic targeting based on satellite imagery. In the DRC, given the absence of household consumption information, an index for geographic vulnerability was developed. The performance of these maps in predicting small-area poverty rates was unknown at the time of writing (see Chapter 5).

2.2.2. Individual-level targeting

During normal times, the best approach to reduce consumption and income-based poverty is through a cash transfer program that targets poor households. Depending on their individual circumstances, countries typically choose one or more of the following targeting methods to identify the poor: community-based targeting, proxy means test (PMT), or a hybrid means test (income and assets). In DRC, Nigeria, and Togo, given that income and asset data was either unavailable or outdated, only community-based and proxy-means testing were feasible. To register households predicted as poor, the program would have either needed to organize a census-sweep operation or devise a system for on-demand applications, where at least one household member would have come in physical contact

25 The development of the machine learning algorithm for Nigeria was done as a part of a broader exercise to produce micro-estimates of wealth for all low- and middle- income countries. For more details, see Chi et al (2022). Full citation: Chi, Guanghua, Han Fang, Sourav Chatterjee, and Joshua E. Blumenstock. "Microestimates of wealth for all low-and middle-income countries." Proceedings of the National Academy of Sciences 119, no. 3 (2022): e2113658119.

26 See Smythe and Blumenstock (2022) for more details. Full citation: Smythe, Isabella S., and Joshua E. Blumenstock. "Geographic microtargeting of social assistance with high-resolution poverty maps." Proceedings of the National Academy of Sciences 119, no. 32 (2022): e2120025119.

with program staff. Such approaches would have required substantial resources and time and would have also risked spreading the virus.

With no existing social registry systems covering their target areas (i.e. Kinshasa and rural Togo),[27] **DRC and Togo chose a contactless, paperless, and cashless delivery system built around mobile phones, designing a program with individuals as the assistance unit, rather than households.** The intention was to use mobile phones for both targeting and program delivery functions, including outreach, registration, eligibility determination, enrolment, and payment. This option, however, had a notable disadvantage: by design, it could not reach those without mobile phones, who are likely to be among the poorest. It also disproportionately excludes women, who have a lower mobile phone ownership rate than men. However, given the unprecedented situation of COVID-19, this was seen as the only option at the time.

Nigeria implemented a similar system, developing a rapid response registry based on a mobile phone enrollment process in urban areas. As a first step, a granular poverty map was used to identify the poorest urban wards. Second, SMS blasts were sent through cell towers located in those areas, inviting people to a secure USSD platform to provide details to be used for enrollment and payments. In contrast to the DRC and Togo, subsequent in-person visits were also organized in Nigeria to validate the selections and collect additional information about individuals, households, and asset positions. However, personalized mobile phone data was not used to target beneficiaries beyond the registration process. Under NASSP SU, the government is considering using asset filters to exclude households at the top of the income distribution through a proxy household wealth indicator. In rural areas, where Nigeria already had a National Social Registry, satellite-based poverty maps are being considered to prioritize wards for support under NASSP SU.

During normal times, many countries operate multiple social assistance programs with different, albeit related, objectives from poverty reduction programs. These programs are based on individual assistance units and use categorial targeting (i.e., child allowances, disability allowances, social pensions). Disability allowances and social pensions aim to provide a minimum income contingent on social risks of disability and loss of employment capacity during old age. Child allowances go towards the cost of raising a child, which is typically a fraction of total household income. Programs implemented during the COVID-19 crisis were comparable in the following aspects: a large proportion of the population was directly or indirectly affected by confinement policies and economic slowdown, and the emergency cash transfer programs supported the maintenance of a minimum income.

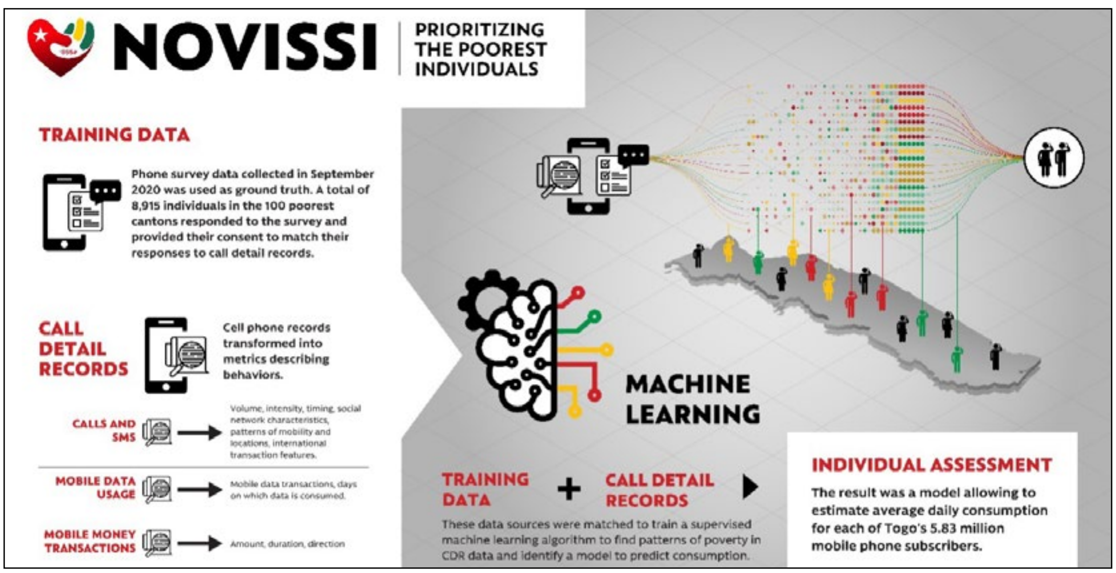
In the DRC, affluence-testing combined with a quota system was used to select individual beneficiaries. The budget for the first phase of the DRC's STEP-KIN program could only reach 250,000 beneficiaries, or about 10 percent of the estimated number of eligible individuals in the areas identified through geographical targeting. Rationing was unintentionally facilitated when only one of the four MNOs operating in Kinshasa (with 16% market share) complied with all bidding requirements and was selected as a partner. The selected MNO produced a list of all its client numbers in the target areas, from which a few customer categories were excluded through a simple affluence test. The program used the following proxies to filter out subscribers with higher incomes: those owning a smart phone, on a prepaid data plan, having monthly expenditures of over $5 per month (which is nearly 70 percent more than average), or making international calls. Only those customers who passed the affluence

27 Prior to Model 2, the lack of a social registry also influenced the delivery model 1 in urban areas.

test were invited to apply. Payments were made on a first come, first-served basis without further prioritization among admissible applicants.

To prioritize among applicants, Togo's Novissi Model 2 used a proxy-means test based on Call Detail Records (CDR), household survey data, a phone survey, non-traditional data, and machine learning. The phone survey was administered to 8,915 individual subscribers[28] over a two-week period in September 2020, with a view to predicting the consumption of active mobile phone subscribers from the poorest 100 cantons.[29] Shorter than traditional household surveys, the 40-minute[30] phone survey covered household demographics, asset ownership and well-being, but not consumption. Consumption was predicted using a proxy means testing model derived from the 2018-2019 household survey administered by the national statistical institute.[31] For this sample of 8,915 subscribers, a second model was developed to estimate consumption using CDR records and a machine learning algorithm.[32] By linking CDR records with self-reported information on asset ownership and other household characteristics, and then estimating consumption, it became possible to gauge the economic wellbeing of households. The ML model was then used to estimate the average daily consumption for all mobile phone subscribers in the CDR database (see Figure 2).

FIGURE 2. Prioritizing the poorest individuals within the poorest cantons (Novissi Model 2)



Source: Lawson et al. 2023

28 This sample represent the final set of interviews being retained, using an innovative sample selection process. To ensure that enough observations were selected from the 100 poorest cantons, the researchers oversampled 40,000 individual subscribers with high probability to reside in these cantons and with a low level of predicted consumption. From this total sample, 10,701 interviews were carried out. After removing low-quality responses, the phone survey sample trimmed to 8,915 individuals whose records were used to predict consumption, merge their records with the CDR data, and then train the ML model using the CDR data. Consent of the subscribers was obtained for the survey, thereby enabling the merging of the CDR data into the overall survey records.

29 Initially, the 100 poorest cantons were targeted. Subsequently, the following 100 next poorest cantons were included (As a result, a total 200 cantons were selected.).

30 The survey aimed to interview only one adult member from the same household. The survey questionnaire is available at https// jblumenstock.com/files/papers/TogoInstrument2020.pdf

31 Institut National de la Statistique et des Études Économiques et Démographiques, INSEED

32 Gualivisi and Newhouse (2022) suggest a similar approach creating training data for a 'fusion' approach of geo-spatial and household survey data.

In summary, DRC's STEP-KIN program and Togo's Novissi Model 2 are two innovative approaches to calibrating and implementing individual-level targeting. Confronted with a "data desert", the DRC opted for a simple approach to identifying eligible beneficiaries, based on an affluence test, and a first-come, first-served system. A master list of subscribers from eligible areas ("whitelist") and simple data from mobile records was used to filter out subscribers with potentially higher incomes. In contrast, Togo combined available traditional data sources (e.g., recent household survey, voter registry[33]) with non-traditional CDR and machine learning tools to estimate the consumption of mobile phone subscribers.

There are similar examples of predictive modelling (based on machine learning and computational statistics) and the use of non-traditional data for household-level targeting in other countries. Bangladesh's COVID-19 response combined mobile phone, administrative, and financial data to reach up to five million beneficiary households. Linked via national ID, their Digital Cash Transfer program compiled public data to exclude civil servants, pension recipients, and recipients of existing social protection programs. In addition, financial data was used to exclude owners of substantial public savings certificates. Any reduced mobile phone activity was used to understand vulnerability to the pandemic's economic impact. In South America, the social enterprise Prosperia.ai has been supporting several countries in the use of more effective predictive models, such as tree-based models, and the curation of data through engineering of highly predictive explanatory variables (e.g., in Colombia, Costa Rica, Ecuador). These approaches seem to have generated better results than traditional proxy means testing based on ordinary least-square (OLS).[34] The use of variable selection approaches, using the LASSO[35] model to select variables in proxy means tests, is another machine learning tool that is increasingly being used in this context.[36]

2.2.3 Key issues and challenges

This subsection raises the need for caution in interpreting the relative success of the novel approaches discussed here as definitive evidence of the broader utility of their application elsewhere. The first part examines specific issues and challenges that pertain to satellite-based methods and the second to mobile phone-based approaches.

Use of satellite-based methods

First, it is important to note that the evidence base for calibrating COVID-19 social protection responses, including eligibility assessments, was collected in normal times and therefore does not take into account any changes resulting from the COVID-19 shock itself. For example, in the absence of real-time information about living conditions, social registries were often used for the vertical expansion of support. However, in the main, these registries contained household information that preceded the outbreak of the coronavirus. The same is true for surveys used to calibrate the assessment tools discussed above (e.g., Nigeria's ground truth consumption survey data and Togo's

33 The voter's registry was used to verify the identify and location of applicants (i.e., whether they live in an eligible canton), not to estimate the consumption of mobile subscribers.

34 Feature engineering refers to transformations of explanatory variables that increase their predictive power, such as squared terms. Tree-based models are a statistical modeling paradigm that splits the data into groups according to explanatory variables value threshold, rather than using fixed proportions as with ordinary least-square or OLS method.

35 The LASSO (Tibshirani, 1996) is a linear statistical model based on OLS that can select variables and reduce their coefficient estimates to enhance out-of-sample predictive power.

36 For more information on different terminologies and methods, a USAID report can be found at https://www.usaid.gov/digital-development/machine-learning/AI-ML-in-development.

consumption data used for estimating household assets). It is likely that those that were already poor at the beginning of the pandemic remained so throughout due to the absence of economic opportunities. However, a lack of evidence on crisis impacts made it difficult to identify those who were in a comfortable economic position prior to the crisis, but subsequently suffered hardship, and required emergency support as a result.

Similarly, neither satellite imagery nor mobile phone-based methods were designed to estimate the impacts of the crisis. Geographic targeting was used to identify those areas of Togo and Nigeria that had been poor at the time of data collection. The static nature of satellite images means that they had little prospect of spotting temporary changes. For instance, the Togolese cantons selected for receiving Novissi funds may not have remained the poorest if they were able to maintain consumption more consistently than other more vulnerable areas. The Novissi Model 2 also identified as poor those mobile phone subscribers whose phone usage patterns were similar to those of the pre-crisis poor. However, it is likely that usage patterns of both poor and non-poor populations changed during the pandemic. In the absence of empirical evidence or theoretical insights on how to adjust for this, the use of prior patterns remained the best option available. However, it is unclear just how effective this was in accurately identifying the neediest and most vulnerable in the selected cantons.

Although satellite poverty mapping uses the best data available, the results may not meet traditional standards of statistical robustness. The number of household observations that provide the estimate of a given grid cell's average consumption level is likely to be relatively low compared with the image area's total population. In addition, the average level of income for a grid cell is not necessarily of great interest for the assignment of social protection programs. Further, areas of high inequality may still contain many poor households despite not having a low average income.

Satellite images may be available more frequently than many other data sources, but they hardly reveal changes over time. A census is typically undertaken only once a decade and most socio-economic surveys are conducted with multi-year lags. Models trained to predict items from such low frequency data sources can fill in temporal or spatial information gaps, for outreach and planning for instance, but the static nature of the built environment means that even frequently updated images reveal very little change. When using satellite images to identify low-income populations (geographic targeting), the analysis is only viable for overall levels, rather than for short- and medium-term poverty dynamics. For instance, although night lights can capture variation in economic growth, their highest resolution is over 100 times lower than that of very-high-resolution satellite imagery (375 meters vs. 30cm per pixel), making them unsuitable for detailed poverty estimation. Bouguen et al (2019) suggest that high-resolution images can identify the impact of household-level cash transfers, but only over an extended time horizon. Satellite-based estimates of poverty dynamics remain unavailable due to the lack of large-scale, dense, and GPS-tagged consumption data needed to train machine learning models.

For many policy tasks, the cost of high-resolution satellite images can limit their usefulness. Whilst lower resolution images are available worldwide free of charge, high and very high-resolution images often need to be commissioned from satellite imagery providers. For extensive areas, costs can be prohibitive, ranging from USD1 to USD30 per square kilometer for high-resolution images (Aiken and Ohlenburg, 2023b). Similarly, although some images may be available from past capture, recent images often need to be commissioned, implying additional costs. In addition, high frequency and high-resolution images must be paired with rich training data to leverage their potential for fine-grained estimation. These financial concerns and data constraints continue to limit the practical applicability of satellite image-based targeting, particularly in low- and middle-income countries.

Moreover, whether in normal or crisis times, selecting beneficiaries by geographical area may be politically unacceptable. Population groups, such as those identified by religion or ethnicity, often cluster geographically. As a result, even if areas were initially chosen with different criteria, geographic eligibility assignment can lead to actual or perceived preferential treatment of some groups over others. In Nigeria, satellite-based geographic targeting is being used carefully to be in line with socio-political realities on the ground: the NASSP SU established quotas at the LGA (admin 2) based on the objective incidence of poverty and vulnerability as well as the intention of the government to expand the program's geographic coverage. However, within the LGAs, wards are prioritized for support based on satellite-based poverty maps. Likewise, poverty maps will also help prioritize wards where social registers (the RRR in urban areas and the NSR in rural areas) need to be expanded.

Use of mobile phone data

The identification of intended beneficiaries with mobile phone-based data faces a different but no less formidable set of challenges—the most important of which is related to the association between phone numbers and individuals. Many social protection programs, and particularly those basing eligibility on the actual or estimated consumption/income of an assistance unit, are targeted at the household level. However, phone numbers are registered to individuals rather than households and the registration data does not include information on whether the registrant is head of their household. Since SIM cards may be shared by several household members, a phone number does not necessarily identify individual usage. Furthermore, both individuals and households may use multiple SIMs, so that phone numbers cannot be associated uniquely with households. The overall impact of multiple SIM usage is unclear, as richer households potentially spread their phone usage across devices: usage may therefore appear relatively low overall compared with households that share a single device and thus a single number. Multiple numbers also raise the prospect of duplicate applications from the same household, which can give an unintended advantage to those with multiple devices (and typically greater wealth). The analysis of phone and SIM sharing in Togo by Aiken et al (2022) sheds some light on these issues, but more evidence and analysis is needed. For better accuracy, extensive data that links a household roster, a subscriber database, and well-being information—possibly via a national ID—would be ideal.

The presence of multiple MNOs further complicates phone-based targeting. A complete picture of phone activity would require the merging of data across mobile networks. Different pricing schemes and payment plans impede direct comparability. In addition, securing data access agreements with several profit-driven actors represents a major organizational and political hurdle. Unfortunately, partnering with a single MNO in each geographic area, as was done in the second stage of the STEP-KIN roll-out, significantly limits access and is incompatible with universal service provision.

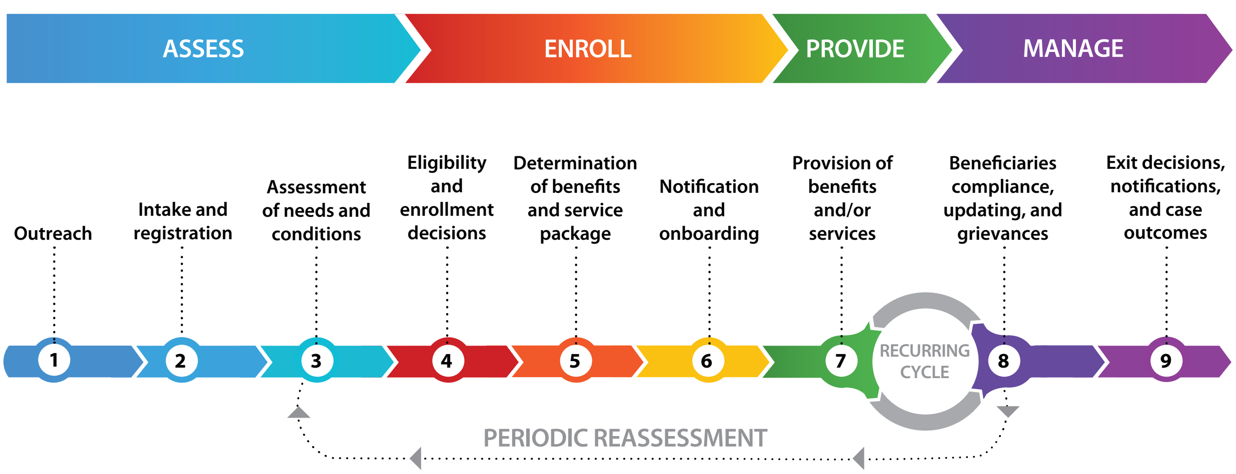
A particularly glaring drawback of mobile phone targeting is that some households do not regularly use a phone at all. Likely to be amongst the poorest members of society, these households face exclusion from phone-targeted programs, thereby compounding existing deprivation. Furthermore, female-headed households exhibit lower phone usage, adding a potentially detrimental gender dimension. Likewise, rural households have lower average phone access, further disadvantaging those whose physical location limits their access to public services and economic opportunity. Finally, when using SIM registration and population units, the identity of the registrant may affect intra-household resource allocation. Social protection programs where the recipient of the payment is female tend to result in a slightly higher allocation to females and children (Camiletti, 2020). If men are more likely to be listed as phone subscribers, intra-household allocation effects may be less than desirable. This was the case for DRC's STEP-KIN, for instance, in which nearly two thirds of the registrants were male. By contrast, the Novissi Model 2 achieved gender parity.

In conclusion, phone-based targeting is a promising new approach but remains relatively untested. Although Aiken et al (2022c) report relatively encouraging simulation results for Togo, a standard proxy means test may be preferable to a mobile phone approach. Additional evidence from a head-to-head comparison of methods in a randomized control trial or an *ex post* evaluation from a socio-economic household would shed additional light on relative targeting accuracy. In the case of DRC's STEP-KIN program, affluence testing criteria were applied to applicant phone numbers, but their targeting effectiveness could not be assessed empirically due to a lack of data linking household conditions and phone use.

2.3. INNOVATIONS IN OTHER ASPECTS OF SOCIAL PROTECTION DELIVERY

Innovation is currently taking place not only in targeting processes, but also along the entire social protection delivery chain (see Figure 3). This section reviews examples of innovations (pandemic-related and others) in four distinct areas: (i) outreach, (ii) intake and registration, (iii) payments, and (iv) monitoring and beneficiary management. Informed by the Sourcebook on the Foundations of Social Protection Delivery Systems (Lindert et al, 2020, Figure 3), this section provides a definition and overview of each area.

FIGURE 3. The Social Protection Delivery Chain



Source: Lindert et al 2020

2.3.1 Outreach

Without effective outreach, many of those meant to benefit from social protection remain unaware of programs, let alone any registration processes. In the context of the pandemic, social distancing and restrictions on movement required changes to traditional approaches such as paper-based registration. The emergence of novel data sources and approaches provides opportunities for further innovation in social assistance delivery. This subsection considers the phone-based approaches used in Togo and the DRC, as well as the use of satellite data to support outreach.

Key innovations

A key advantage of mobile-phone outreach is that it does not require a resident or household register. Based on registries of mobile phone numbers, MNOs send targeted messages (e.g., phone calls, text messages (SMS), and voice messages (interactive voice response calls (IVR)) directly to subscribers in designated areas. Users are assigned to the tower to which they are habitually nearest. While SMS had already been widely used in traditional approaches (e.g., UNICEF's RapidPro SMS dialogue tool), its usage has thus far been limited to communication with existing and potential beneficiaries who already provided their phone number to the program during the intake process. In that sense, using mobile data as a base registry widened the scope of outreach to intended beneficiaries. The way phone numbers were used to contact subscribers differ across the three country cases, as described below.

In Togo's Novissi Model 2, phone-based outreach played an important role in complementing traditional methods. First, CDR was used to estimate the welfare level of mobile phone subscribers. Based on this information, the program pre-selected mobile subscribers in target areas to receive an SMS about the program inviting them to initiate registration via USSD. This direct-to-household approach was complemented by more traditional outreach methods, such as an extensive 7-language radio campaign across 30 stations (Lawson et al, unpublished). Finally, community leaders helped distribute information about the program through local structures and networks.

Similarly, in DRC, the STEP-KIN program complemented traditional outreach channels and sent mass messages to all eligible phone numbers. This was vital in a context where no alternative registers were available. The list of eligible phone numbers of potential beneficiaries (known as "the whitelist") was generated by the contracted MNO using five exclusion criteria. This was a simpler approach than the screening mechanism used in Togo's Novissi Model 2. In conjunction with information points, posters, deployment of community ambassadors and radio discussions, all phone numbers deemed resident in the target areas (and not screened out by exclusion criteria) received SMS and IVRs. They were also contacted by the call center (Mukherjee et al 2023).

Unlike Togo and DRC, which only contacted pre-selected specific phone numbers in target areas, Nigeria's outreach strategy was purely geographic. An innovative pilot for the country's Rapid Response Registry (RRR) sent SMS blasts to all mobile numbers in metropolitan areas. Satellite-based poverty maps identified those poorer urban wards that were previously not part of the country's limited social protection system. Digital methods were then used to capture these invisible urban populations. Outreach was supported by radio and television transmissions, use of social media influencers, and local town criers in target areas. These were deemed necessary to build confidence in an environment with limited trust in digital communications such as SMS. The rapid registry will also play a key role in identifying urban beneficiaries under the NASSP SU program.

Satellite images offer an alternative data source to bridge geographic coverage gaps of social registries, helping address a potential under-subscription of households. In Ecuador and Costa Rica, governments worked with the social enterprise Prosperia.ai to improve outreach by guiding social workers to households that were not currently included in their social registry and may have been neglected. An experimental machine learning method combined social registry data with satellite images to create maps on which recipient households were marked. A machine learning algorithm was trained to recognize the likely presence of beneficiary households on unseen 50x50m image tiles with a high resolution of 50cm. An additional algorithm clustered the tiles to reveal 41 areas with significant outreach potential that could guide the dispatch of social workers to uncovered areas (Aiken and Ohlenburg, 2023a). Although this use was not fully operationalized, Prosperia.ai is due to deploy a similar system in Argentina.

Challenges and drawbacks

No single outreach channel, including mobile phones, can be expected to reach all intended beneficiaries. As outlined above, digital exclusion has many facets. Even those households that own and use a phone regularly may not be able to understand the outreach message due to illiteracy, disability, or language barriers. Recipients may also be wary of potential scams and may not trust the messages or payment transfers they receive. Experiences in both Togo and the DRC have demonstrated the importance of community outreach for raising awareness and building credibility for new programs. In Togo, despite multi-layer outreach strategies, only around 40 percent of mobile phone subscribers were aware of the Novissi program and attempted to register.[37]

Reliance on the private sector is an additional hurdle faced by governments in reaching intended beneficiaries. Mobile network operators, satellite data providers and technology companies (including social media) may not wish to take part in assistance programs due to strategic and financial considerations. These are likely to be insurmountable without investing in appropriate incentives or legislative action. The level of MNO participation directly influences the scope and success of outreach. Customers of non-participating MNOs cannot receive outreach messages and are therefore deprived of the opportunity to apply for programs. Although DRC STEP-KIN's limitation to a single MNO was a useful rationing device under emergency settings, it would be problematic for regular programs that are intended to serve the whole list of eligible subscribers, across operators. Moreover, when several operators are involved, a robust ID verification mechanism, ideally linked to a foundational ID system,[38] is important to thwart duplicate registrations.

With respect to targeting based on satellite imagery, the distant view from sky/space and the static nature of the built environment can restrict outreach.[39] Satellite imagery can trace the impacts of large cash transfers but generally fails to capture changes in the population and in living conditions

37 Aiken et al., 2021 based on Novissi Model 2

38 Also, as discussed in the previous section on targeting, it is also important to recall that the unit of phone-based outreach is the subscriber, which is close but different from individuals, and is also different from households in the absence of a household roster than links individuals to households.

39 On a separate issue, the dynamic potential of phone data is a double-edged sword, in terms of using regular vs. current locations depending on the program objectives. On the one hand, mobile network data can reveal population movement that may support social protection delivery during times of unusual internal migration. For example, aggregate population movements derived from phone data in the Gambia (Arai et al, 2021) showed substantial movement out of the capital city region to rural areas. Such information could inform adaptive social protection programs, including data-driven localization of outreach activities. In such times, the assignment of a phone's regular location, which uses location data over extended periods, will be inaccurate for some time after migration. Whether the use of residence and movement data is appropriate thus depends on the context and the program objectives.

over the short- to medium-term (Huang, 2021). Also, while satellites can capture the location, size and type of dwellings, such characteristics are relatively static—the only exceptions being rapid urban expansion commonly found in slum areas and the construction of temporary shelters. Another limiting factor of satellite-based outreach is the need for extensive training data, which requires a social registry with broad, accurate coverage that includes household geo-locations. Outreach activities supported by satellite imagery are thus most suitable for bridging coverage gaps in pre-existing, extensive social registries.

2.3.2 Intake and registration

Accurate information about applicants is essential for effective program administration all along the delivery chain, including during the intake and registration stage. In addition to the innovations made in Togo and the DRC, this section reviews examples from Southern Africa and Asia. Simply put, intake refers to the process of "initiating contact, engaging clients, and gathering information", while registration consists of "recording and verifying that information" (Lindert et al, 2020). The pandemic restricted regular data collection modalities involving physical contact, such as household visits and the verification of identification documents. Therefore, simplified alternative approaches had to be conceived. Outside periods of crisis, new computational tools are helping increase data usage and availability.

Key innovations

Although mass mobile messages can be used for one-way outreach, registration requires two-way interaction and the active participation of potential beneficiaries. Togo's Novissi Model 2 capitalized on the two-way communication potential of mobile phones for intake and registration, employing a USSD interface of text-based menus. To register, potential beneficiaries merely needed to provide their name and voter ID number, together with their consent for processing personally identifiable information for program purposes. The most interesting feature of this approach was that data retrieved from the voter registration list[40] was used to complete the applicant's profile automatically and facilitated the household-level targeting process. Given that voter ID was only applicable to individuals who possessed a physical card, this security feature guarded against fraudulent registration under the name of a different person (Lawson et al, unpublished).

DRC's STEP-KIN introduced a similarly streamlined registration process. The targeting process was completed through whitelisting based on simple exclusion criteria, and like in Togo, interested applicants submitted their data via a user-friendly USSD registration platform. In the case of DRC, however, the verification of applicants' information was minimal (i.e., whether they reside in target areas, whether the numbers are in the whitelists) and the program worked on a first-come, first-served basis. Registration closed once the applicant quota was reached.

The pilot for Nigeria's RRR also used phone-based USSD registration, but this was followed by in-person household visits to verify information and collect additional datapoints. Initial challenges with text-based mass registration led to the deployment of a voice-based option which subsequently enhanced response rates. Nigeria's RRR involves in-person household visits for the collection of some 42 variables, corresponding to about a quarter of all data on the National Social Registry. Although the initial objective was to reduce contact times and lower data collection costs during the pandemic,

40 The variables included voter's unique identification number, a security number available only in printed credentials, name, gender, home location, and occupation (Lawson et al, unpublished).

processing times in Nigeria were still longer than in the Togo and DRC (the latter were able to start paying tens of thousands of beneficiaries in the very first year after the start of the pandemic). It was more than 1.5 years after the onset of the COVID-19 crisis in 2022 that Nigeria was able to provide cash transfers to a portion of beneficiaries. They scaled-up the enrollment of more beneficiaries in November 2023.

Elsewhere, in an endeavor to improve data quality, Liberia and India have put into place systems to detect misidentification of households[41] and false claims. To improve its social registry, and avoid misidentification or misclassification of households, Liberia developed a ML algorithm using photos of housing structures and GPS coordinates collected during household visits (Del Bono et al, forthcoming). In India, an AI platform for the country's subsidized health insurance PM-JAY uses image analysis to spot duplicate applicant photos and to foil attempts to create ghost beneficiaries. To identify potentially manipulated images (that are not simply duplicated, cropped, or mirrored), the algorithm uses nodal points for a unique encoding of a person's face. Human operators then verify and confirm any suspect images (Jain 2021). Although not flawless, these low-cost methods can be useful for mitigating fraud.

Challenges and drawbacks

Digital exclusion is of course the most important barrier to the effectiveness of phone-based and other technology-driven intake and enrollment approaches that are not accompanied by a physical alternative. For example, only 72 percent of mobile phone subscribers were able to complete the Novissi 2 registration process once initiated, taking a total of four attempts on average (Aiken et al., 2021). In South Africa, the civil society organization "Black Sash" documented the various challenges that applicants faced when applying for the fully digital Social Relief of Distress (SRD) grant (Senona et al, 2021). For example, households often had to seek community support, by borrowing a suitable device from neighbors or getting help from acquaintances to navigate the process. Unfortunately, digital exclusion in the early parts of the delivery chain can break down the social assistance delivery chain altogether. Households that are not made aware of the existence of an assistance program are unable to register for it in the first place, or fail to receive payments despite their eligibility, and are therefore unable to benefit from the social protection coverage intended for them. Although fully digital programs may be attractive for administrative simplicity, the complex needs of beneficiaries suggest that systematic support along the entire delivery chain is necessary.

2.3.3 Payments

In this paper, payments refer to the provision of benefits to beneficiaries—usually cash, but sometimes also vouchers. The fast, reliable, economical receipt of payments is necessary to translate program costs into actual value for beneficiaries. During the pandemic, social protection programs were able to leverage advances in payment systems, such as mobile wallets, to deliver support: in the 120 social assistance programs supported by the World Bank in 2022, the share of digital payments reached 92 percent.[42] In some cases, specific circumstances required adaptations to payment systems, and these are discussed in further detail below.

41 In detecting duplicates, GPS coordinates in combination with high-resolution satellite imagery might work in rural areas, but only as an indication since more than one household may live in a given house. However, in dense urban settlements, GPS isn't precise enough to disambiguate households as most devices mismeasure by several meters.

42 Based on the estimates from 580 million beneficiaries from 120 cash transfer programs in 66 countries which are supported with the World Bank operations (World Bank, 2022b).

Key innovations

In Togo and DRC, automatic account creation facilitated payment delivery for those that were not existing clients of payment partners. Existing partnerships with MNOs for the purposes of targeting can also streamline payment delivery processes when mobile money accounts are used. Making mobile money accounts available to all program beneficiaries had several advantages: it eliminated the costs of opening the account (e.g., time spent, provision and verification of documents etc.) and provided training in its use, effectively lowering learning cost. This also had the effect of boosting overall financial inclusion. In Togo, Novissi's use of voter ID, including remote verification, further facilitated the process.

In the DRC, the simplification of Know Your Customer (KYC) requirements enabled the creation of new low-cost, quick-payment accounts for half of all STEP-KIN beneficiaries without requiring an ID document or physical verification. The central bank, as financial regulator, issued an instruction to authorize specific social protection accounts to be created in this manner by recognized institutions (Gentilini et al, 2021). Accounts were restricted and had low transaction limits but were sufficient for the purpose of social assistance delivery. Remote account opening was also allowed, albeit with even lower transaction limits. This regulatory flexibility fostered financial inclusion, creating accounts for those without any form of identification. In Togo, the use of voter IDs, which serve as official identification documents, provided sufficient information for the automatic creation of regular, unrestricted mobile money accounts.

Beyond KYC, changes in regulations were made in other countries during the pandemic, enabling non-bank institutions to serve as payment service providers. For example, Ecuador allowed the involvement of non-financial actors such as grocery shops and pharmacies to act as 'cash-in-cash-out' agents (Hammad et al 2021). Similarly, Angola introduced a new law on Payment Systems in 2020 (Law 40/20) allowing non-bank institutions to operate as Payment Service Providers. In Haiti, a new regulation had been issued granting stand-alone licenses for the provision of digital payments to non-bank payment service providers (World Bank 2022b). In South Africa, supermarkets acted as an important cash-out point for those COVID-19 grant beneficiaries without a bank account (Senona et al, 2021).

Challenges and drawbacks

Even with the innovations mentioned above, fundamental challenges related to the delivery of payments remain, such as lack of access points and limited connectivity. When expanding Togo's Novissi program to rural areas, the ability to cash out became a considerable barrier to social assistance delivery. As a result, an agreement was made with MNOs that a mobile money agent should be available no more than 5 km from a beneficiary's residence. Service fees are also an important hurdle in many countries, e.g., in Ecuador. Finally, connectivity, particularly in rural areas, continues to be a significant bottleneck to attracting payment service providers (e.g. in Egypt, World Bank 2022).

Emergency programs may require some flexibility with respect to anti-money laundering (AML) and combating the financing of terrorism (CFT) regulations. International counterterrorism financing obligations as well as the need to safeguard domestic revenue collection and avoid money laundering ultimately necessitate the verification of the identity of financial account holders. Although KYC requirements are important tools for AML/CFT, these policy objectives need to be balanced against access to social protection and emergency support. In the DRC, authorizing new accounts with small balances and transaction values is a useful and practical compromise, and still allows for the possibility of full authentication later. Lowering standards in emergency situations may be a sensible temporary policy solution but would need to be revisited for post-emergency settings. Such a "stepped" compliance approach can provide benefits from a financial inclusion perspective but requires greater supervision to safeguard against abuse. Thus, policy coordination among the relevant parties is indispensable.

Procurement or technical considerations can sometimes limit the ability of program administrators to make the optimal choice for beneficiaries. To ensure a fair process, an open and multi-provider approach is ideal, as this can reach more subscribers (and potential beneficiaries) in target areas. However, involving multiple providers may not always be feasible due to local contexts and the lack of incentives for payment service providers. In the case of DRC, given the absence of a mechanism to verify the identity of applicants from multiple carriers, STEP-KIN purposefully chose a single mobile operator per geographic area to reduce the risk of duplications.

Whilst creating basic accounts with limited KYC enhances the breadth of financial inclusion (greater account ownership), the impact on the depth of inclusion is limited (actual usage and benefits). The limited services available with basic accounts only represent a small part of the pathway to financial inclusion. Savings capacity and microlending, for example, may remain out of reach for many. The full array of financial services will remain conditional on the satisfaction of KYC/AML requirement and several other factors, including a strong national ecosystem of digital payments.

2.3.4 Monitoring and beneficiary management

Monitoring and beneficiary management refers to the activity of "continuously engaging and collecting information from the field or other sources (as other databases), which is then processed through a set of protocols, recorded, and used to make decisions." It also includes grievance redress, by which queries, suggestions, positive feedback, and concerns can be raised and addressed, problems with implementation resolved, and complaints addressed. Monitoring and beneficiary management often receives little attention during the early development and implementation of a new program. This was often true for temporary benefits under COVID-19 emergency programs, where the priority was rapid funds disbursement. This is indeed an area with great potential for the application of innovative approaches, e.g., for better communication, grievance redressal, re-certification, and fraud detection. This section reviews some emerging algorithmic innovations in this area, mainly relevant for regular non-crisis programs.

Key innovations

Technology-supported communications are an important area of innovation in the effective management of social protection programs. Virtual assistants (or "chatbots") can be used to manage common queries by providing frequently requested information and standard responses. This has the advantage of reducing the load on call centers and streamlining beneficiary management. Chatbots have been deployed in several countries to support users of social security institutions, including in Malaysia (ISSA, 2022) and various Latin American countries (ISSA, 2021). Technology-assisted communications can also lower accessibility barriers, as exemplified by Egypt's sign language chatbot for the hearing impaired (UNDP, 2020). With these systems in place, only non-standard queries that cannot be resolved by chatbots are channeled to human agents. This reduces the resources required for beneficiary management, including program staff. Advances in natural language processing (NLP), as exemplified by OpenAI's ChatGPT model, are likely to broaden the capabilities of AI agents to handle increasingly complex interactions and have a significant impact on government-to-person (G2P) interaction.

As a part of its efforts to address complaints related to eligibility, and make determinations more transparent to applicants, Prosperia.ai has developed a method to explain and communicate the results of proxy means tests for Latin America governments. To keep applicants and enumerators from gaming proxy means tests, details of statistical decision processes are usually kept secret. However, recent advances in the explainability of predictive models now allow for intuitive explanations to be generated for a given applicant, without divulging the underlying mechanics. Providing a visualization of each applicant's summary characteristics (housing, assets, occupation, socio-demographics) and comparing them with a reference group, helps identify areas that increase or decrease the welfare estimate (Carrillo et al, 2021). Such a tool may be especially useful for grievance and redress processes.

In India, the national health insurance scheme PM-JAY is using algorithms to identify fraudulent treatment claims. Ghost beneficiaries and false claims result in major losses to social protection systems. In this context, AI- and data-enabled systems hold significant potential. India's pilot system converts claim forms into digital text via optical character recognition software and then applies natural language processing models to spot suspicious patterns. India's model was trained for certain high-cost procedures that have significant fraud potential, such as knee replacement surgery. Another complementary approach being trialed is the use of social network analysis to identify collusion patterns, flagging cases that may require further investigation (Ohlenburg, 2022).

With respect to re-certification, South Africa's SRD 350 pandemic grant illustrates the potential of digital data using a more frequent and less costly eligibility test. Beneficiaries who received a bank transfer over the prior month were excluded from the grant, as it was deemed that they had a source of income which made them ineligible for assistance. The system relied on cooperation with the banking association, which checked beneficiary accounts and then submitted a list of those who had received payment above a certain threshold. The system was hampered by initial implementation issues, because a simple receipt of payment does not necessarily signify an income—as a result, this re-certification criterion was only applied to those who had registered an account with one of the larger banks (Matthews et al, 2020). There might also be data privacy concerns and beneficiaries may reject the back-end processing of their account for monitoring purposes. This could further discourage beneficiaries from using their accounts more generally and compromise outcomes related to financial inclusion. Nevertheless, regular use of such affluence tests as well as other exclusion criteria shows how data exchange protocols can be used to foster a more dynamic social protection system.

Furthermore, novel data sources, especially when combined with AI, can strengthen preparedness and early actions to make social protection more adaptive and responsive to diverse climate-related shocks. Over the past decades, there has been significant progress to link social protection systems with early warning systems and climate forecasts, including for drought, food security, and disaster preparedness and responses. For example, the Dominican Republic has used social registry data to estimate the likelihood of a household being vulnerable to natural disasters (UNDP-UNEP 2018). While there are countries with similar system linkages, the expansion or payout of social assistance benefits are not generally triggered by these forecasts. However, in Niger, targeting mechanisms and payout timings were adjusted based on the Water Requirement Satisfaction Index, which measures dryness via satellite data on rainfall and evaporation (Brunelin et al, 2022). Similarly, Bangladesh deployed anticipatory cash transfers ahead of severe flooding—these transfers appeared to have had a greater impact the earlier they were disbursed (Pople et al, 2021). These examples highlight the potential of novel data sources for facilitating early action, allowing households to avoid negative coping strategies such as asset sales that lower their earnings potential.

Finally, the use of business process automation in beneficiary management is likely to become more prevalent as the capabilities of automated decision-making evolve and AI tools gain wider acceptance in the public sector. While the evidence of business process automation for case work is limited, one municipality in Sweden, Trelleborg, provides an example of when automated decision making in social assistance was successfully implemented using a hybrid approach (Ranerup and Henriksen, 2022). In this case, according to several qualitative interviews, automated suggestions were made to human decision makers based on various documents received from applicants, allowing them to focus on more complex cases while waving through simple ones. However, cautionary tales of failed AI projects also exist. For instance, Lokshin and Umapathi (2022) cite three cases from the US and Denmark in which efforts by child protection services to algorithmically identify vulnerable minors were terminated, due to a relatively high error rate. Poor quality data that leads to flawed decisions is certainly a growing area of concern.

Challenges and drawbacks

In general, explainability remains an important challenge, due to lack of experience and a rapidly evolving technological landscape. Both the theory and methods in this field are in the early stages of development and documented applications in a social protection setting remain sparse. The tension between the transparency of public decisions and process integrity (maintained by guarding information) implies trade-offs as these objectives cannot always be fully reconciled. Nevertheless, a balance does need to be struck, and it may be better to implement a flawed solution imperfectly than not to provide a solution at all.

Fraud mitigation is a long-standing concern in social protection systems and is evolving rapidly in the digital age alongside the threat landscape. As fraud techniques emerge that take advantage of new technological possibilities, detection systems require continuous adjustments and updates. However, given the rapid pace of change, renewed efforts by administrations may be necessary to stay competitive, particularly considering the growing capacity of malevolent actors. During the pandemic, a variety of cases emerged across the globe that illustrate the need for fast adaptation.

In general, despite their enormous potential, AI and data-driven use cases for managing beneficiaries of social programs are still few and far between.[43] Digital technologies and DPI offer growing opportunities to make feedback and grievance mechanisms more responsive, speedier, and less costly than traditional approaches. With the use of business intelligence and a variety of data sources (not only program data but also non-program data such as social media) can enhance program monitoring and review. However, governments are only beginning to harness the potential of data generated by

43 Outside of social protection, there are also, albeit limited, use cases of AI and technologies to enhance citizen engagement. For example, Nigeria piloted an AI solution for citizen feedback to monitor project progress using mobile app which consisted of several features including AI-powered tag cloud, geo-referencing, image classifier, and image matching, and opinion mining and sentiment analyzer (World Bank 2020).

digital systems and devices (Center for Global Development 2019; Ohlenburg 2020).[44] For adaptive social protection systems, the technical readiness to forecast shocks is not a sufficient condition to trigger anticipatory actions to nimbly scale up or down social assistance—countries must also establish national financing, institutional, and political enablers.

44 Yet, an example is India's energy subsidy reform program, which offered a real-time rating option for the service providers—a sort of customer satisfaction feedback mechanism— using SMS and online portals.

3. Preliminary results and analysis

This chapter presents some of the preliminary outcomes of the digital programs deployed in DRC, Nigeria, and Togo during the COVID-19 crisis. It is structured around the following four topics: (i) implementation speed, (ii) administrative costs, (iii) targeting accuracy, and (iv) convenience for beneficiaries.

3.1 IMPLEMENTATION SPEED

To be effective, crisis response programs need to be designed and implemented quickly. Countries with established registries of potential beneficiaries with high coverage can easily achieve that (e.g., Chile and Turkey provide two excellent examples of rapidly deployed COVID-19 programs). However, as discussed above, in the cases of the DRC, Nigeria or Togo, the information on desired beneficiaries was largely or completely absent. To appropriately assess the speed of implementation, a fair comparison would be to examine the time taken to design and implement a new welfare targeting program under normal, non-crisis conditions. For example, the Georgia Targeted Social Assistance program, designed in 2005 to replace several privilege-based social assistance programs, took about 18 months: the design phase lasted some 8 months, with applications being received from the 6th month onwards, needs assessments from the 8th month onwards, and distribution of cash benefits after 18 months. The 2000 Armenian Family Benefits program required about two years to deploy. After the Philippines initiated a dialogue on cash transfer programs in 2006, it took one year to deploy the pilot phase (with 4,500 households) and another year to enroll the first set of 300,000 beneficiaries.

Compared to the cases above, in DRC and Togo, the time between the enabling political decision and the first round of payments was significantly shorter: 8 to 9 months for the Novissi Model 2[45] **and about 11 months for DRC's STEP-KIN.** The biggest time cost for DRC's STEP-KIN was securing the financing and the selection and contractual arrangements with the MNOs. In Togo, the estimation of the poverty map and the individual targeting model took the bulk of the period. Once the design and contractual arrangements with MNOs were concluded, outreach, application, needs-assessment, and payment transfers were done relatively quickly. Overall, the programs in DRC and Togo took only half

45 Model 1 of Novissi was deployed within one month of the first recorded COVID case.

of the time compared with the design and deployment of a traditional PMT-based social assistance program. In Nigeria, due to political economy factors, and the use of in-person visits, emergency cash transfers were deployed only one and a half years after the crisis began.

3.2 ADMINISTRATIVE COSTS

In normal times, the administrative costs of welfare-targeted cash transfer programs represent a small share of total program costs. Grosh et al (2022) report that share to be about 5–10% for cash transfer programs (conditional or not). Tesliuc et al (2014) report a range from 2–10% for last-resort cash transfer programs in the ECA region. Schnitzer and Stoeffler (2020) found that the share of administrative costs as a proportion of total program costs in Burkina Faso, Chad, and Niger was 0.4%-5.5%. Rosas et al (2016) report a cost of 7.7% for the implementation of the first phase of the Tanzania Productive Safety Net (PSN) program. Jamaica's PATH program registered a slightly higher share of administrative costs, at about 11% for 2018, to cover the delivery of conditional cash transfers, case management services and the cost of associated social workers (World Bank 2019).

Compared to welfare-targeted cash transfer programs operating under normal times, the emergency cash transfer programs implemented in DRC and Togo had a similar share of administrative costs as a proportion of total program costs, for a similar range of benefits, i.e., 5.7% in DRC and about 10% in the case of Togo Novissi Model 2. The composition of administrative costs for the STEP-KIN program shows that half of these costs were for MNO services (see Table 3 below). In the DRC STEP-KIN case, several factors contributed to low administrative costs including the following: the use of an existing implementation agency (the Social Fund), a digital intake strategy with low data collection cists, a simplification of program procedures triggered by emergency conditions (e.g., simplified KYC requirements), and fee waivers offered by the successful bidder for technical assistance and payment services.

TABLE 3. STEP-KIN program budget, direct versus operational cost

Cost component	Percentage
Direct costs (transfers to beneficiaries)	**94.3**
Indirect costs (administrative, operational cost)	**5.7**
Mobile money services (by MNOs) and technical assistance (by implementing partner)*	3.0
Operational costs (including Social Fund staff, central management information system, communication materials)	1.5
Banking and wire fees (from Social Fund commercial bank to mobile money service provider's bank account)	1.2
Total	**100.0**

Source: Mukherjee et al 2023
* Technical assistance included hotline and piloting a new system and platform for the registration and payment.

3.3 TARGETING ACCURACY

At the time of writing, neither Novissi Model 2 nor STEP-KIN had put into in place ex post assessments of their targeting approaches based on rigorous quantitative methods. In the case of STEP-KIN, a qualitative assessment and phone survey was conducted among program beneficiaries, suggesting that the program reached its intended target group. Most of the phone survey participants were poor and vulnerable, according to self-declared information: 39 percent were unemployed, 61 percent earned less than $100 per month on average (mainly from commerce activities (47 percent) or manual works (16 percent). Results of the survey suggest that the program principally reached people in the informal sector (Social Fund 2022).[46]

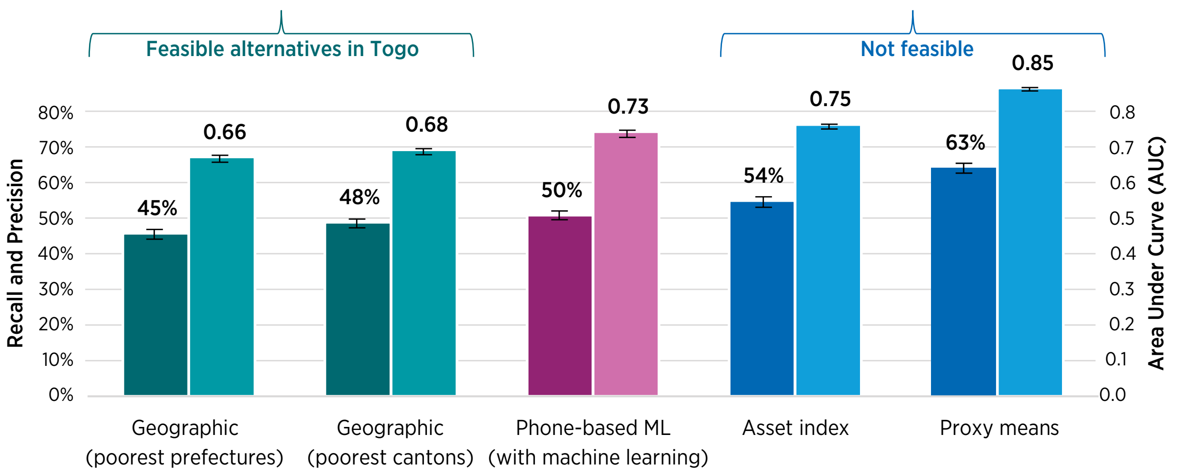
Although the targeting approach used by STEP-KIN was useful in an emergency context, records of program consultations highlight some important challenges, especially with respect to perceptions of program inclusiveness. This perception is illustrated by the following remarks made during the survey: “some rich people are living in poor neighborhoods” and “poor people may not have a phone”. Furthermore, the limited program budget was viewed as contributing to exclusion, with feedback received along the following lines: “there are other poor neighborhoods” and “some deserving people don’t have the mobile number of the MNO which participate in the STEP-KIN”. It is important to note that at the time of writing, targeting performance had not been assessed by quantifying the level of errors based on the household survey. Furthermore, the program found that around 10 percent of beneficiaries never withdrew their funds despite being active users of the mobile network and after repeated contact attempts (Mukherjee et al 2023). Additional measures like household visits could have been taken to validate the identity and legitimacy of these beneficiaries. This is particularly important in the context of regular programs.

In the case of Novissi Model 2, Aiken et al (2021) have simulated the performance of the actual targeting mechanism with two alternative targeting approaches using an asset index and a PMT score—all calibrated based on the 2018/19 household survey (see Figure 4 below). No such comparison could have been done for the 2020 program due to a lack of survey data. The darker bars show the share of true poor as a share of total poor and the lighter bars the precision of the program (share of total poor as a share of program beneficiaries). These measures are the closest to the measures regularly reported in targeting assessments. They show that the method used by Novissi Model 2 was moderately accurate, reaching about 54% of the true poor, compared to the PMT score, which is estimated to have reached 63% of true poor among the program beneficiaries. The 9-percentage points difference between the two targeting approaches is quite large.

Figure 4 below reveals that phone-based targeting performed roughly on par with an asset index, which can be seen as a simplified proxy means test (Aiken et al 2021). In this figure, the darker bars represent recall and precision (on the left axis) and the lighter bars stand for Area Under the Curve (AUC, right axis). Since neither proxy means test nor asset index data could be collected, the phone-based approach represented the most accurate feasible solution in this case, with the subscriber-level targeting approach seemingly supporting the geographic targeting approach. It is important to note that these are simulation results, whereas it is common in practice to encounter unforeseen complications that affect program roll-out, including targeting accuracy. As such, ex post data collected from actual beneficiaries would provide the best evidence. An upcoming evaluation of Novissi Model 2 is expected to provide such data.

46 The survey confirmed that recipients used money to meet their basic and daily needs, including food (46 percent), health and education (35 percent), reinvestment in their livelihoods (32 percent), and rent (12 percent).

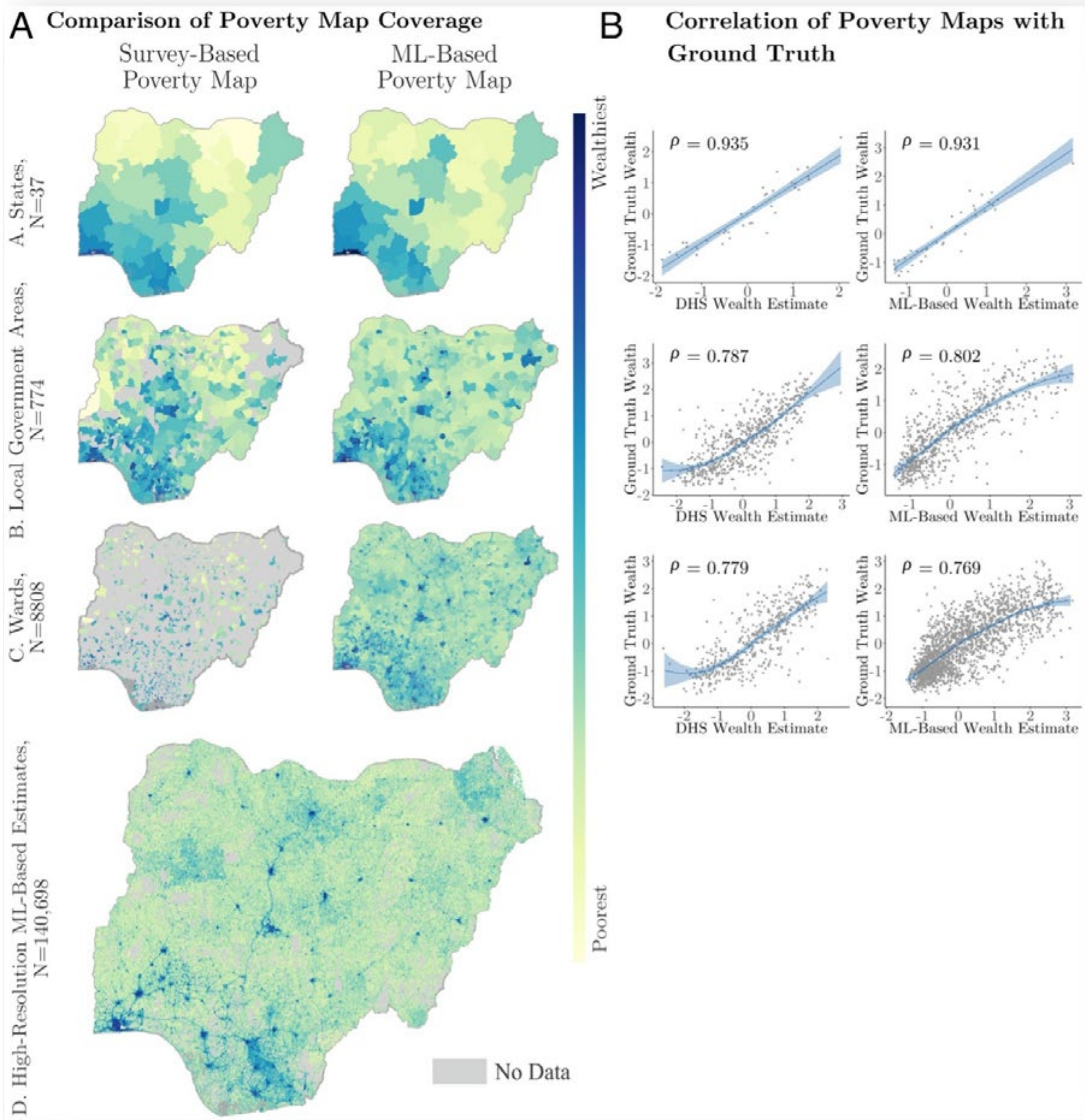
FIGURE 4. Targeting a hypothetical nationwide social assistance program



Source: Aiken et al. 2021

In the case of Nigeria, Smythe and Blumenstock (2022) compared a traditional survey-based poverty map and a machine learning-based poverty map to assess the coverage and accuracy of the two different approaches. Figure 5A below compares the coverage of two poverty maps, with gray areas indicating the areas where no household was interviewed for DHS in 2018 (these administrative units were not selected for the survey sample). While both maps have complete coverage at the state level, the survey-based map has more grey areas at the disaggregated levels of LGA and wards, as the DHS sampling strategy covered only 14 percent of all wards in the country. In contrast, the machine learning model can extrapolate welfare estimates in areas where the survey was not conducted, thereby increasing the resolution of the poverty maps at much more disaggregated levels – even compared to traditional small area estimation techniques. Figure 5B below compares the accuracy of the two maps. This comparison was made feasible thanks to two different data survey sources: a DHS survey was used to train a machine learning model, which was then validated by assessing the precision of the predictions in NLSS 2018/19 georeferenced data. At all levels (State, LGA, and ward), the analysis confirmed a strong correlation between the predictions and NLSS 2018/19 ground truth. It is important to note that the welfare measure used in the simulation was not consumption, but rather an asset/wealth index that, in general, is moderately correlated with consumption.

FIGURE 5. **Coverage and correlations of ML and benchmark poverty maps to NLSS-estimated ground truth poverty maps**

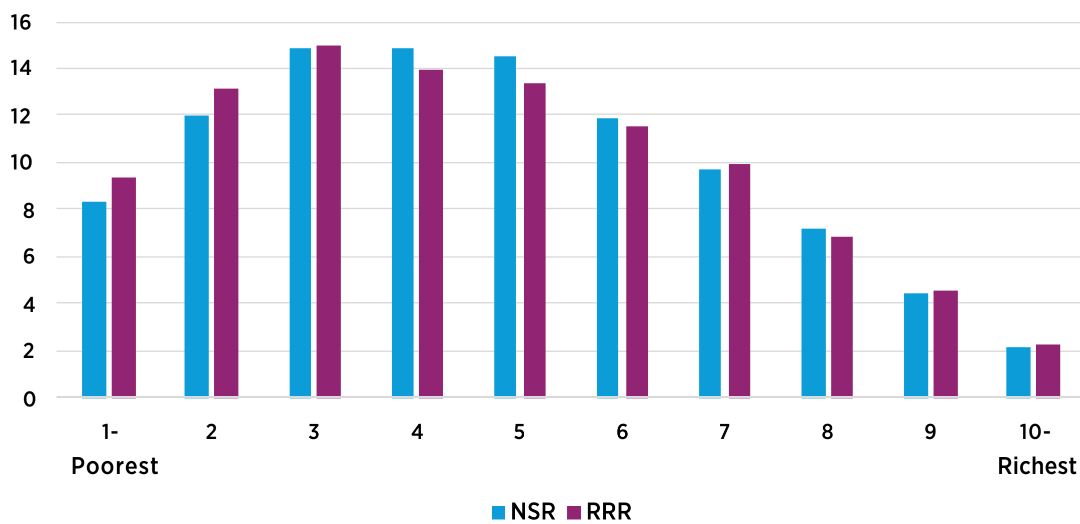


Note: A compares the coverage and estimates of traditional survey-based poverty maps (Left) and ML-based poverty maps (Right) at the three different administrative levels: state (row A), LGA (row B), and ward (row C). Regions without data are shown in gray. Row D shows the high-resolution ML-based estimates prior to aggregation. For privacy reasons, high-resolution poverty estimates are not generated for grid cells with fewer than 10 inhabitants. B compares the ML and survey benchmark (the DHS) wealth estimates of each administrative unit against the NLSS ground truth estimate of that unit's wealth. Pearson's correlation coefficients are reported across all relevant units. Fewer observations exist in B because not all LGAs and wards contain households that were surveyed in the DHS. All correlations are significant at P = 0.001.

Source: Smythe and Blumenstock (2022)

In Nigeria, both registries, the conventional NSR and the new RRR, were deemed to be similarly effective in identifying the poor and the vulnerable. A harmonized proxy wealth indicator was developed with the uniform set of variables collected for both registries. The proxy indicator shows that, although the registries were compiled through different processes and targeted in different areas (urban vs. rural), both registries have a similar profile of households in terms of wealth. They are both similarly effective in identifying the poor and the vulnerable (roughly the poorest 65% of the population), as is evident in the higher frequencies among the second to fifth deciles (Figure 6).[47] Under the ongoing scale-up (NASSP SU), the government plans to ensure that at least 80 percent of beneficiaries are from the bottom six deciles.

FIGURE 6. Distribution of households by wealth deciles (Nigeria)



Source: World Bank, mimeo

3.4 CONVENIENCE AND COST FOR BENEFICIARIES

The few studies that estimate the cost for beneficiaries in accessing cash transfers put the share of these costs (monetary costs and opportunity costs of time spent) in the 5-10% range as a share of total cost (Tesliuc, 2014). An important feature of the mobile-based emergency cash transfer programs in Kinshasa and Togo was the very low set up and transaction costs, if any, for beneficiaries. This was also reported in satisfaction surveys. STEP-KIN and Novissi Model 2 targeted those who already had mobile phones, and sufficient basic (digital) literacy for cashing out money using mobile wallets. However, as mentioned above, some of the poorest populations[48] without mobile phones or adequate literacy were excluded by this method.

47 In Nigeria, 40 percent were living in extreme poverty, and another 25 percent were considered poor and vulnerable to extreme poverty (Poverty Assessment, World Bank).

48 The digital exclusion is also more prevalent among females. For example, based on phone surveys, about 36 percent of STEP-KIN beneficiaries were females, showing the gender inequality.

4. Innovation examples after the COVID-19 shock

A brief survey of World Bank social protection practitioners was conducted to identify similar cases, carried out after the COVID-19 shock, to those discussed above. Two such pilots are presented below. The first is the Togo Social Safety Net and Basic Services Project which incorporated innovations into the more traditional delivery infrastructure/process. The second is the Malawi Canva Cash Transfers Pilot which integrated an expanded monitoring and evaluation framework from the design stage.

4.1 TOGO SOCIAL SAFETY NET AND BASIC SERVICES PROJECT

The Togo Social Safety Net and Basic Services Project illustrates a unique multi-stage approach to identifying and registering households using novel data sources, SWIFT-PMT and community validation. Building upon Novissi's remarkable technological leapfrogging during the pandemic, the Togolese government continued its journey of improving the effectiveness and efficiency of social protection delivery in data-poor and resource-constrained environments.

Since 2017, the National Agency for Grassroots Development (Agence Nationale d'Appui au Développement à la Base, ANADEB) implemented the Social Safety Net and Basic Services (Filets Sociaux et Services de Base, FSB) project, funded by the World Bank and the Togolese government. While cash transfers under the project were mostly deployed in rural areas, an additional financing from the World Bank and the French Development Agency (Agence Française de Développement, AFD) in 2021 extended the project's support to Greater Lomé to mitigate the lasting impacts of the pandemic in urban and peri-urban areas. In Lomé, the project targeted around 33,000 households providing each with cash transfers of 15,000 FCFA (ca. US$25) every three months for 18 months (i.e. a total of six payments totaling 90,000 FCFA or about US$150).

The targeting methodology in Greater Lomé, previously untrialled and developed jointly between the government and the two development partners mentioned above, relies on a novel multi-stage approach to identify and register households using a variety of mechanisms:

1. **Geographic targeting.** The initial step was to identify and prioritize the most vulnerable neighborhoods on the basis of satellite imagery and crowd-sourced data. In the absence of disaggregated consumption data at the neighborhood level, the project team used the Project

Targeting Index methodology (Finn and Masaki, 2020) to build a composite index and rank neighborhoods in Greater Lomé (admin-level 4) from the most to the least deprived.

2. Self-registration. Households living in the 45 most deprived neighborhoods were able to self-register in person. The government ran a public outreach campaign and set up registration sites in the targeted neighborhoods asking applicants to provide basic contact details. Self-registration increases the opportunity costs of registrants and introduces an element of self-selection that helps filter out the better-off households (Alatas et al. 2016b).

3. SWIFT Proxy Means Test (SWIFT-PMT). Due to the budget available, the program had to prioritize households with higher needs among those that pre-registered. The World Bank's Poverty and Equity practice developed a resource-efficient method for monitoring poverty called SWIFT (Survey of Well-being via Instant and Frequent Tracking) that requires a minimal questionnaire to estimate household consumption using traditional econometrics combined with machine-learning techniques to maximize out-of-sample performance (Yoshida et al. 2022). The team adapted this method to develop a PMT model for identifying the poorest households within previously identified neighborhoods. As shock-responsiveness was a key objective of the program, the SWIFT-PMT questionnaire included questions on assets as well as specific consumption items to better capture changes in welfare. The SWIFT-PMT was administered in person for most pre-registrants—although for a small sample of households it was first collected through phone calls or USSD before being verified in person.

4. Community validation. To confirm that households assessed as poor by the SWIFT-PMT were also considered poor by the community, ranked households underwent a participatory community validation process for the final beneficiary selection.

Although multi-stage approaches are not unusual for registering and prioritizing beneficiaries, the combined use of satellite imagery and crowd-sourced data, self-registration, modular and shortened SWIFT-PMT questionnaire, and community validation makes Togo's Social Safety Net and Basic Services Project a unique case.

To evaluate the approach, a consumption and expenditures survey was collected in-person for a sub-sample of households from preregistered and non-preregistered households, in and beyond targeted neighborhoods. This consumption survey will serve as 'ground-truth' to assess the targeting performance and cost-efficiency of the different stages and intake modalities used. Also, consent to link phone usage metadata with survey responses was obtained from households to explore how to potentially synergize this approach with CDR-based targeting, previously employed in Togo under the Novissi program.

Although the targeting methodology deployed by the Safety Nets and Basic Services Project in Greater Lomé has yet to be evaluated, it could offer multiple insights on the trade-offs between costs, accuracy, and timeliness when collecting data and assessing the needs and conditions of potential beneficiaries. These insights can support the design of the dynamic social registry currently being developed by the government (Registre Social des Personnes et des Ménages) to make sure it reaches the poorest and most vulnerable households, while ensuring that data stays up to date. Although limited in scale, the project tests the boundaries of what is practical and effective for a single program in terms of combining targeting approaches, computational methods, novel data sources, and robust evaluations.

4.2 MALAWI'S CANVA CASH TRANSFERS

The Malawi cash transfer pilot is another example of the use of novel data sources for social assistance delivery that can also apply to post-pandemic conditions. The pilot program was funded by Canva, a visual communications platform, and implemented by the international NGO GiveDirectly. The program provided cash transfers totaling around $600 per beneficiary ($50 per month, over one year) to people in TA Khongoni, rural Lilongwe, a region surrounding the nation's capital city. Preparatory work started in 2021, with 12,800 beneficiaries receiving payments in 2022 (Phase 1).

The pilot targeted those living under the international extreme poverty line of $1.90, i.e. around 80% of the population in the target areas. The incidence of poverty was higher than initially expected, likely due to long-term impacts of COVID-19 as well as the lean season.

The pilot adopted a similar approach to the Togo Novissi Model 2 and used novel data sources for geographic and individual targeting. Mobile phones and mobile wallets were then used as the key tool for implementing remote enrollment, verification, and payment. Outreach and communication about program eligibility was largely done through community meetings and in-person engagement. Since relying on mobile phones only would have excluded a significant share of the target population, particularly the poorest who may lack phones, the pilot complemented the remote digital approach with in-person targeting and enrollment alternatives to mitigate digital exclusion.

Three different approaches for individual targeting were used as follows:

1. **Saturation/Census:** Geographic targeting was done without an assessment of households. All households in the identified poor area were eligible for cash transfers and were visited through a census–sweep approach to registration. Beneficiaries enrolled in the program upon the verification of their national ID.

2. **Proxy Means Test:** A PMT[49] model (for estimated consumption) was calibrated based on the National Statistical Office's Integrated Household Survey from 2019-2020 (IHS5). Through household visits, answers to a PMT questionnaire containing about 30 questions were collected and subsequently used to estimate the welfare of the household. Those households with a lower welfare score than the threshold score (i.e., total consumption of $1.90 per capita per day) were eligible for the program. The eligible households received a repeat visit for the onboarding part of the program.

3. **Tele-targeting (TT) combined with in-person Proxy Means Test:** GiveDirectly conducted a representative household survey of the rural Lilongwe region for a sample comprising 15,000 households. This included a PMT survey of 12,000 households and a consumption module for 2,000 households. Consequently, GiveDirectly estimated consumption via a PMT score, which was then matched with subscriber CDR data (through an agreement with the MNO) to develop

49 Unusually, the PMT formula contains several food consumption items that improve the precision of the formula and ensure that the formula adapts to economic shocks, but consumption items have not been included in most PMT models so far because they are easy to misreport and hard to verify.

a second prediction model which estimated consumption using CDR and a machine learning algorithm (i.e., CDR-based PMT). After conducting in-person community meetings in target villages, subscribers were invited to register via a brief USSD survey, which screened individuals using the CDR-based consumption score (CDR-based PMT), their home cell tower location, and other self-reported information. Registrants in target areas who were over 18 years old, whose ID number was verified against the national ID database, and whose CDR-based consumption score was lower than the threshold of $1.90 daily were deemed eligible for the program. Following digital enrollment, GiveDirectly conducted in-person outreach in all areas where TT was implemented: program staff visited TT target villages to conduct a door-to-door PMT survey to i) identify and enroll individuals excluded by the digital approach, ii) measure inclusion and exclusion errors on the ground, and iii) assess the operational feasibility of combining both remote and in-person targeting and enrollment approaches.

The individual targeting methods described above were applied in three distinct districts of rural Lilongwe regions, with 24% of 12,800 total beneficiaries (3,052 beneficiaries) selected using the tele-targeting method. While the preparatory work of arranging data sharing with the MNO took about 7 months, the rest of the process was fairly rapid, including collecting the ground truth data, and developing the tele-targeting algorithm to produce CDR-based consumption scores (CDR-based PMTs), and self-enrollment. Out of about 19,400 users that registered via a USSD interface, 3,052 beneficiaries were enrolled within 56 hours after validation of their national ID numbers.

TABLE 4. Speed of different targeting methods (Preparation and average time required for registration and enrollment)

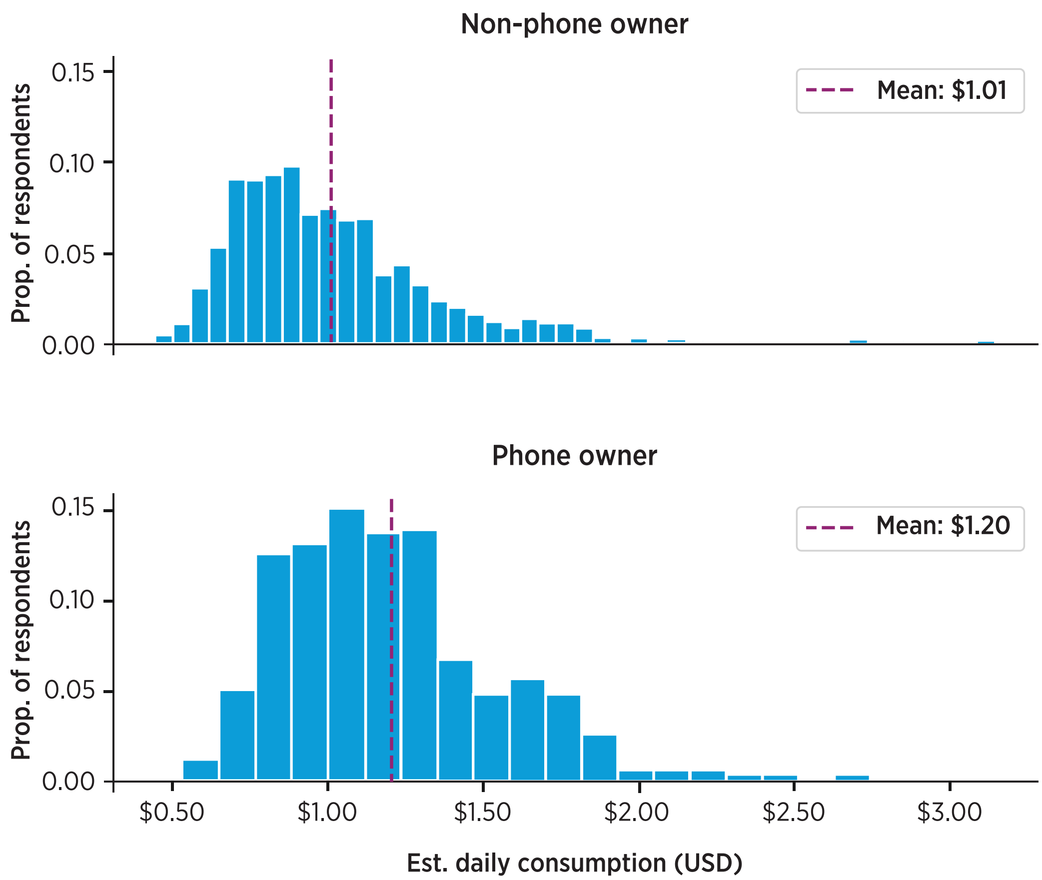
Targeting method	Preparation time	HH visit to registration	Registration to enrollment
1. Saturation	4 months	6 days	22 days
2. PMT	3 months	12.5 days	14 days
3. Tele-targeting	7 months	0 days*	0 days

*For tele-targeting, registration was through USSD platform/survey, without household visits.
Source: GiveDirectly 2023

For Phase 1 of the project, the share of administration and other indirect costs is estimated to be around 20% (80 cents go directly to the recipients for every 1 dollar in the total budget) but there is great variation across the targeting methods. The important takeaway is that the initial fixed cost for tele-targeting was significantly higher compared to saturation due to the cost of collecting ground truth data and developing the tele-targeting algorithm. The incremental cost per recipient would be notably lower if the program were to expand. Even in a setting where ~35.8% of households owned a phone, the variable cost-per-person-enrolled for combining tele-targeting and in-person PMT surveys (i.e., a hybrid remote and in-person approach) was still only around 50% of the variable cost-per-person-enrolled for just in-person PMTs for the entire population. It is likely that cost-efficiencies would increase in populations with higher rates of phone ownership. It is also important to note that the fixed cost for tele-targeting could be significantly decreased for future iterations of the program, given that once data-sharing partnerships and protocols are established, they can be used for multiple targeting efforts. Therefore, a higher number of ground truth surveys were collected for the purpose of training the tele-targeting algorithm than required for the initial tele-targeting process itself.

Based on estimated consumption, it is important to note that beneficiaries who owned a phone were on average 20% wealthier than those who did not (Figure 7). This confirms that extra effort is required to reach out to, and include, non-phone owners.

FIGURE 7. Share/Distribution of beneficiaries by estimated consumption (based on PMT)—non-phone owners vs. phone owners



Source: GiveDirectly, 2023

Using consumption data collected from 2,000 households, GiveDirectly tried to validate the CDR-based consumption score and estimate the targeting performance (see Table 5 below). Targeting errors were estimated at 17% for the tele-targeting algorithm and 13% for the PMT when targeting the bottom 80% of the consumption distribution—a cutoff that in this case coincides with the $1.90 poverty line. When targeting the bottom 30% of the distribution, the performance of the tele-targeting algorithm is similar to the one used in Togo's Novissi (58% vs. 53% inclusion and exclusion errors for each model respectively). The PMT questionnaire was based in part on the IHS5, using 17 questions from its consumption module. Although the evaluation of comparative targeting performance was not carried out with the same consumption module as the IHS5, the evaluation of a PMT that includes food consumption variables is novel and of methodological interest.

In general, the results show that the returns of targeting diminish as the proportion of the population that is eligible increases. The tele-targeting model performs only slightly better than random methods when targeting the poorest 80% of the population, suggesting that the tele-targeting model has limited benefits in settings where high poverty is prevalent. Although the PMT model performs slightly better than the tele-targeting model, it shows similar patterns of diminishing returns as the proportion of eligible beneficiaries increases.

TABLE 5. Results of Tele-targeting model (CDR-based PMT model) (top) and PMT model (bottom)

5A. Results of CDR-based PMT model

Proportion of Population Targeted	Pearson	Spearman	AUC	Accuracy	Precision	Recall
10%	0.23	0.28	0.64	84%	20%	20%
20%	0.23	0.28	0.64	73%	34%	34%
30%	0.23	0.28	0.64	65%	42%	42%
40%	0.23	0.28	0.64	61%	51%	51%
50%	0.23	0.28	0.64	59%	59%	59%
60%	0.23	0.28	0.64	59%	66%	66%
70%	0.23	0.28	0.64	65%	75%	75%
80%	0.23	0.28	0.64	73%	83%	83%
90%	0.23	0.28	0.64	84%	91%	91%

5B. Results of PMT model

Proportion of Population Targeted	Pearson	Spearman	AUC	Accuracy	Precision	Recall
10%	0.486659	0.567576	0.784146	86%	33%	33%
20%	0.486659	0.567576	0.784146	78%	45%	45%
30%	0.486659	0.567576	0.784146	73%	55%	55%
40%	0.486659	0.567576	0.784146	70%	63%	63%
50%	0.486659	0.567576	0.784146	70%	70%	70%
60%	0.486659	0.567576	0.784146	72%	76%	76%
70%	0.486659	0.567576	0.784146	74%	82%	82%
80%	0.486659	0.567576	0.784146	79%	87%	87%
90%	0.486659	0.567576	0.784146	88%	93%	93%

Notes:
Proportion of Proportion Targeted (Program coverage), from the bottom to top based on the consumption scores
Accuracy: Correct prediction over all predictions
Precision: True positives over total positives (true and false positives) where a positive in this case is being under the poverty line.
Note that exclusion error/inclusion error is calculated as 1 – Recall/Precision.

Source: GiveDirectly 2022

Building upon the initial pilot, Phase 2 will support twice as many beneficiaries as Phase 1, with an additional budget of $20 million to support the poor in rural areas (GiveDirectly 2023). With this expansion, the program is currently using saturation to enroll beneficiaries. Moving forward, combining remote and in-person targeting, and enrollment models may increase the efficiency and speed of delivering social cash transfers at scale, even in rural settings with low rates of phone ownership. To that effect, for the later phases, the organization will explore how to most effectively combine traditional and novel approaches to achieve better outcomes in terms of cost-efficiency and targeting performance.

In sum, the combination of methods need to be selected based on the context, including poverty and mobile penetration rates in the program areas, as well as the targeting performance of different methods (advantages and disadvantages). The poverty incidence is high in Malawi with an overall national rate of 72%, and 80% in the pilot areas. Digital exclusion is another important factor when targeting the poorest, particularly for a country like Malawi where mobile penetration is still low (34%). Separately and relatedly, targeting performance must be examined in greater depth, particularly when it comes to targeting the extreme poor in areas with less widespread poverty, such as the bottom 10%, 20%, and 50% of that population.

5. Conclusion and way forward

Although recent progress in non-traditional data, methods and systems for social protection heralds a new era of innovation for greater program adaptability and inclusiveness, it is important to be aware of the limitations of such approaches. This chapter distills key insights from the discussion above and looks ahead to identify important opportunities in this domain. The first part of the chapter summarizes the utility of novel approaches for emergency settings like the COVID-19 pandemic and other crises (such as natural disasters), as well as for regular social assistance programs. This is followed by an examination of current limitations and unknowns of such approaches. The chapter concludes with an assessment of policy implications and a set of recommendations.

5.1. EMERGENCY SITUATIONS

The most important advantage of non-traditional approaches in emergency settings is that they make welfare-targeted cash transfer programs possible even when there is no pre-existing social protection delivery system. All three countries reviewed in this paper (Togo, DRC, Nigeria) lacked a pre-existing social registry covering the target populations. DRC's STEP-KIN demonstrated that payments can be made in the absence of any existing information, purely based on MNO data. In Togo, mobile phone data from MNOs and a recent electoral register with a few relevant variables, such as profession and residence, facilitated individual level targeting and verification. In Nigeria,[50] the low coverage of the existing registry focused on rural areas, thereby limiting the usefulness of extending emergency support for the urban population. However, it is not known how well these simple approaches worked, due to the lack of a rigorous evaluation mechanism (see 5.3).

Second, digital systems enhanced ease of access and lowered costs, both for program implementers and successful applicants. As discussed previously, the relatively low administrative cost (which may even decline further over time) is a strong argument in favor of the programs reviewed. The question remains as to whether the cost of also providing support to the digitally excluded will eventually outweigh the cost advantages of digital delivery.

Third, the programs were created and implemented more quickly (in less than one year) than traditional targeted social assistance based on a physical footprint. Countries having strong service

50 The acceptability issues that surfaced for satellite targeting in Nigeria suggest that political economy considerations may render technically effective solutions unviable even during emergency times of considerable need.

delivery fundamentals, particularly dynamic and inclusive social registry, were able to deploy support even faster. The innovation cases reviewed show that securing the funding and creating the enabling conditions, such as agreements with MNOs and adjustment of financial rules and regulations, was more time consuming than the technical aspects of digital deployment (e.g., targeting methods, online application platforms, management information systems). Once this was completed, payments could be transferred quickly, and any scale-up or repeat programs would be simpler and faster to implement.

Fourth, the programs that used the mobile phone as the main or only service delivery tool will likely exclude a significant share of the population, i.e., those 15 years and over that do not own a mobile or do not have sufficient mobile literacy to register, enroll and receive payments without physical contact with the program staff. Not surprisingly, the proportion of adults without a mobile phone is larger in low-income countries, especially in FCV countries, and is higher among the poorest than for the rest of the population. Based on FINDEX 2020/21 data, mobile coverage was only 56% in LIC-FCVs (10 countries), 64% in non-FCV-LICs (7 countries), 83% in lower-MICs (41 countries), and 90% in the upper-MICs (36 countries). As discussed in this paper, only 34% of the households in the rural communities where Malawi's pilot was implemented owned a mobile phone (see Chapter 4.2).

The novel approaches examined here are applicable to a wide range of programs, including where restrictions on mobility are in place. The pandemic was a highly unusual situation with lockdowns[51] and social distancing restrictions, during which in-person efforts, such as outreach and registration, were not feasible. During other types of emergencies, such as natural disasters, humanitarian crises or conflict, displacement would be a more common scenario rather than containment at home, and therefore the geographic aspect may be limited in such cases. Nevertheless, mobile connectivity and cell tower localization may allow the adjustment of programs to suit different requirements. Satellite imagery and meteorology data have been used to monitor natural disasters (Bowen et al 2020) and may substitute or be increasingly adopted for geographic targeting for disaster-related support. Neither Togo's Novissi Model 2 nor DRC's STEP-KIN, designed as COVID-19 emergency programs, provide a directly replicable template for emergency social assistance more generally however, and would require significant changes for different scenarios.

In summary, the approaches reviewed here were useful in delivering support relatively quickly in a crisis, given the lack of established delivery systems and limited supporting infrastructure. The use of digital devices and solutions can offer a proportion of the population a convenient form of social assistance but does not provide a fully inclusive approach (as it excludes those who are without mobile phones, or have insufficient digital literacy to apply for and enroll in programs). In addition, the pandemic setting was atypical, and the applicability of these approaches and targeting techniques to other situations and settings remains unproven.

5.2. REGULAR OPERATIONS

Factors contributing to a successful mainstream social assistance program differ from those deployed in the context of a crisis. In non-emergency settings, the lead time required to develop

51 A synchronized economic shut-down accompanied by stay-at-home orders and social distancing meant that financial support needs were widespread and great.

a program is less of a consideration. On the other hand, the use of digital channels to lower costs and speed up delivery would be seen as important advantages. It remains to be seen whether the administrative cost of digital solutions would be lower for mainstream programs, given the more complex accessibility and beneficiary management requirements. On the other hand, adjustments to the design and implementation of digitally-delivered programs (such as a change in eligibility criteria) may be less complicated and costly to implement (e.g., due to less staff re-training and more efficient distribution and communications channels).

A major potential advantage of using these innovative approaches for mainstream social assistance programs is the ability to easily and cost efficiently expand outreach to a broader population, beyond the poor. This is particularly important considering the various crises that may result from climate change, natural disasters, and economic and political instability. As such, while traditional social assistance programs tend to focus on the poorest populations, there is a growing need in many countries to extend support to vulnerable and informal populations as well. The use of digital data collection, instead of data collection through physical contact with household members, can be useful in such a context. Combined with the use of novel data sources, digital registration and intake processes can also lower hurdles and costs for registrants, thereby potentially extending program reach and expanding a social registry's coverage.

An important benefit of phone-based estimates is continuous data availability, which can make assessments, as well as the social registries themselves, more dynamic. Instead of the multi-year gaps between official household surveys common in LMICs (Barca and Hebbar, 2020) and the frequency of updates of administrative data, phone data can be collected continually and evaluated at a much higher frequency. This could open the door to a steep change in social assistance responsiveness to economic conditions. Programs could expand horizontally and vertically with relative ease, when necessary, potentially laying the foundation for universal social protection and even automatic economic stabilizers that play an important role in aggregate demand management in high income countries.

5.3 LIMITATIONS AND UNKNOWNS

While the advantages and benefits of novel approaches are enticing for policymakers, it is important to review and discuss their limitations and unknowns. The issues discussed here are primarily related to outreach and targeting, as these were the focus of the case studies under review.

First and foremost, an entirely digital approach will necessarily mean some exclusion, the extent of which is still not fully understood. As discussed above, an approach based on mobile phone ownership will exclude many of the poorest populations (which represent a high proportion among LICs). Furthermore, the participation of MNOs will determine whether adopting phone-based outreach is feasible in the first instance. This dependence on the private sector may be problematic, particularly in an oligopolistic or monopolistic market, or when the market share of the participating MNO(s) is small or inadequate. In addition, the phone use patterns of individuals vs. households is not well known: targeting units are typically either individuals or households, but it is unclear to what extent either relates to phone usage as revealed by CDR. As outlined above, specific targeting is hampered by factors including the use of multiple SIMs (by persons or households), sharing of SIMs (between persons and households), fragmentation of usage across multiple providers, and a lack of distinction between business and personal use. Resolving these issues at the household level would require a complete listing of phone numbers per household, access to CDR from all prevalent providers, and linked household-level consumption data.

Registration, assessment, and provision of programs via digital devices is a major trend in social protection, but the correlation between digital exclusion and poverty poses a challenge to universal access and affects the core constituency of many programs. Effective approaches to the mitigation of digital exclusion, such as self-registration for those without devices, remain under-explored. To ensure fairness and inclusion, a system that integrates the parallel provision of traditional approaches for a subset of the population is required. The cost implications as well as the ultimate targeting outcomes of such a hybrid model need to be investigated in the context of practical applications: this would likely require complete survey sweeps to identify the level of exclusion in target areas.

The increased availability of geospatial, mobile, and other non-traditional data, coupled with the progress in machine learning modelling, are promising developments for creating frequently updated poverty maps for smaller geographic units. The accuracy of predictive models will depend on the correlation between geographic features and regional poverty, the availability of individual-level variables and their correlation with individual welfare, and the effectiveness of the inference approach. Collecting more geo-tagged household survey data could eventually improve the training of these models, which in turn will improve prediction. Combining geographic data with household-level information that allows inference of economic wellbeing, as done with CDR in Togo, may provide an alternative source of data for satellite-based poverty maps. Some of the unknowns with respect to the actual targeting performance of the programs reviewed here include the following:

- **Ground truth quantiles.** Remote poverty estimation with satellite images, as used in Togo's Novissi 2, may be less useful for household-level targeting than it would seem at first sight. For social protection targeting, there are three main issues to consider when training a model for satellite images on estimates of mean consumption per image time:

  - First, a quantile of the consumption distribution, rather than area mean consumption, is the relevant eligibility measure, and the presence of households below the implicit or explicit consumption threshold (such as the second decile of the national consumption distribution) may be weakly related to the area's mean level.
  - Second, even if one were to use a more tailored statistic, the limited number of geo-located households surveyed in a given area may be insufficient to yield adequate estimates for that area.
  - Third, in addition to population density, local inequality also matters in determining the likely number of targeted households but may not be picked up by this approach when using relatively low-resolution images. An empirical assessment of the validity of these concerns may require a census-style survey sweep of numerous areas across varied geographies.

- **Composite satellite indicators.** DRC's STEP-KIN program used a composite indicator of urban deprivation that successfully identified low-income neighborhoods, and the Safety Nets and Basic Services Project program (in Lomé, Togo) followed an analogous approach. In a similar vein to the point made above on quantiles, an equally weighted measure of several area characteristics may or may not be a good predictor of the number of poor persons or households. In the absence of geo-located ground truth data (observed consumption or income), such a composite of measurable indicators thought to be relevant is a pragmatic choice, but with unknown accuracy. Both composite indicators and remote poverty estimation may be useful rationing devices for resource-constrained programs. However, determining whether either is an effective geographic targeting method would require a comparison with detailed empirical data, e.g., small area estimation poverty maps (Elbers et al, 2003).

- **Multi-stage targeting.** The accuracy of targeting methods is commonly evaluated in ex ante simulations with survey data and related information, through empirical trials, or with ex post beneficiary sample surveys. Evaluations can be conducted for a single targeting method, or may compare different methods, but they are usually only made at one level of analysis, such as geographic areas or households. Novissi Model 2 and STEP-KIN, as well as the Lomé Safety Nets and Basic Services Project, used multi-stage approaches, and each of these stages suffered from targeting errors. The way in which the different stages interact when it comes to targeting errors is unknown, but it seems likely that they compound to some extent, rather than cancelling each other out. A complete dataset of geo-located household consumption linked with CDR (or any other data source used) would be needed to gauge the ultimate performance of a multi-stage system.

- **PMT innovation robustness.** Innovations in proxy means testing methods are being developed, for instance in the Lomé and Malawi cases. In particular, the use of consumption variables in the PMT as well as data collection via mobile phones are novel non-traditional practices. As components of the overall consumption, particular consumption items (such as meat) are not just direct components of the outcome, but also highly predictive of other components. Despite the strong correlation with target variables, PMT administrators have traditionally avoided such questions given that the information is difficult or impossible to verify, and easy to misreport.[52] The extent to which households can be trusted to report their expenses truthfully is an open and important question. Remote data collection, such as via phone interviews or at registration desks, essentially extends the lack of verifiability to the entire set of responses. Whether misreporting is an important issue in practice is worth exploring: additional variables promise more accurate models, and remote data collection substantially lowers enumeration costs, potentially opening the door to wider coverage and more frequent re-certification.

The quality of the training model depends in part on the availability of quality ground-truth data. If the objective of the poverty map is to estimate the consumption- or income-based poverty for small areas, the best exogenous variable of the model will be the household consumption measured in a household survey. Some of the poverty maps developed elsewhere have used asset indices derived from DHS surveys, which are less accurate than household consumption data derived from living standards measurement surveys. These indices are only a proxy of the true measure to be predicted, and they correlate positively but weakly with household consumption. Even where average income is available, the density of the data in most sources is too low for use with high resolution images. Corral et al (2023) find that models optimized on R^2 can make misleading geographic targeting choices, and that machine learning approaches trained on geo-located variables may not provide greater accuracy than those using census data. Another limitation is that estimates of average income may diverge from the quantiles of the consumption distribution relevant for program assignment. Although the

52 What we know from the PMT literature is that some household misreport the information that is hard to verify, if they believe that this misreporting will improve their odds of becoming eligible to the program.

quality and availability of satellite-based welfare estimates has been rising considerably, they remain relatively unproven compared with more established targeting methods.

Phone-based targeting appears to have lower accuracy than proxy means tests (PMTs). In a simulation for Togo, Aiken et al (2022b) found a 8–9% accuracy difference in an early evaluation, calling into question the suitability of phone-based targeting for large-scale deployment where a PMT is viable. The combination with satellite-based targeting is highly innovative, but the overall accuracy of this multi-stage process is difficult to ascertain and has not yet been fully investigated.

With respect to targeting for economic criteria in crisis settings, it is important to note that changes in economic welfare during an emergency are not well understood. Geographic and individual targeting used welfare estimates modelled on pre-emergency conditions, as no evidence of correlations between observable characteristics (satellite images or CDR) and living conditions was available. In addition, the normally used consumption measure of economic wellbeing is an annual concept that has limited applicability to emergency living conditions. Although these approaches can indicate disadvantaged areas and households, it is not clear how precise they are, nor how a ranking by need can be determined during a crisis. These uncertainties extend across countries and programs.

CDR-based methods face a key limitation when it comes to their 'targeting unit'. As discussed in this paper, whereas satellite-based geographic targeting is calibrated on the traditional household unit of social assistance assignment, phone-based methods target SIM cards, which do not correspond with household units. In the absence of an accurate user register and a household roster that links individuals to household units, targeting accuracy at the SIM level is an ill-defined concept: with SIM cards, it is not known who precisely is being targeted, nor how to measure their economic wellbeing.

At the time of writing, the targeting accuracy of these novel approaches is not known—investing in well-designed *ex post* evaluations is sorely needed. The targeting accuracy statistics reported in DRC and Togo, and in the Lome and Malawi (rural Lilongwe) pilots, are all based on simulations using different welfare proxies rather than the actual measured welfare of the recipient households. Simulation studies based on sample populations in living standard measurement surveys are the most common tool for assessing targeting accuracy. Although it is useful, and essentially irreplaceable, for program design, this approach implicitly assumes that (i) the targeting instrument can be applied to the full population at scale and (ii) the responses to the targeting survey will be the same as to the sample survey. These major assumptions merit empirical testing to verify that the simulation and actual outcomes coincide, and that the policy thus operates as expected. Adding a question about program receipt to a sample population survey is the best ex post evaluation approach in this regard. Randomized control trials (RCTs) offer a different evaluation approach, namely assessing the impact of different programs on certain outcomes of interest. The impact of varying a program's targeting methods on consumption can offer important insights into their relative merits. However, the limited scale of RCTs, due to resource intensity, means that they cannot be a direct substitute for ex post, representative sample surveys that can assess population-level targeting outcomes.

A key attraction of novel data sources is the avoidance of costly, complex and time-consuming household surveys, which are conducted infrequently in many countries. However, the vision of frequent low-cost data updates may not be realistic in every context. CDR and some satellite data are held by commercial operators who may charge for access or even refuse access all together. For satellite imagery, much depends on the resolution needed for sufficient accuracy. Expenses for data science work, IT infrastructure, transaction costs, and the operation of a parallel provision system also need to be considered. Pilot projects will offer an important starting point for estimating large-scale program costs.

From an operational perspective, in addition to the technical aspects discussed above, robust good practices are also needed for the development of scalable and sustainable business models and standards for the use of mobile phone data (including DRC), geospatial data and other novel forms of data. When using such data for policy interventions, establishing suitable arrangements for data analysis and building collaborative partnerships with MNOs, government agencies and other partners, are important considerations (Milusheva et al, 2021).

5.4 SUMMARY AND RECOMMENDATIONS

Social protection is at a threshold of transformative change due to growing digitalization, the rise of AI and machine learning, and the need for adaptive and universal systems. This section draws out several insights that emerge from the comparison of programs and their applicability in different contexts. A key theme is the complementarity between traditional and non-conventional approaches. Although novel data sources and methods have been shown to serve as possible substitutes for established social protection systems during the pandemic, these methods targeted a broad set of vulnerable populations and not necessarily the poor. In addition, their effectiveness remains unclear due to a lack of robust evaluations. On the other hand, rather than offering outright alternatives, novel approaches can enhance and supplement traditional delivery chains.

The next generation of cash transfer programs is likely to combine both traditional and novel approaches for service delivery and eligibility determination. Many of the phases of the delivery chain are likely to shift towards digital delivery via mobile devices, while more traditional modalities will be maintained as a backup to provide alternative channels for those beneficiaries that cannot use digital means. Likewise, use of novel data and estimation techniques are likely to complement the more traditional, now mainstream, modalities of using representative household surveys and/or administrative data and regression techniques to calibrate the eligibility criteria and to apply these criteria across the whole set of program applicants.

The spread of novel data sources and techniques will likely be gradual. The recent innovations documented in this paper hold great promise for both service delivery and targeting, but their effectiveness and efficiency has not yet been thoroughly evaluated, and some difficult issues remain. Use cases are still few and far between and challenges related to exclusion and data protection, among others, need to be overcome. This suggests that the integration of these novel approaches is likely to be an incremental process rather than a sudden shift, and will work in tandem with traditional approaches.

The digital programs implemented in DRC and Togo were deployed swiftly and in a user-friendly manner, at moderate administrative cost. However, these programs were relatively small, and the approaches used may be more suitable for emergency contexts than for mainstream social assistance programs. Nonetheless, there is significant potential for digitization in the delivery chain that includes, but is not limited to, outreach, registration, and eligibility determination. The programs reviewed here can be interpreted as precursors of a digital-first shift for social assistance. Integration of novel data and processing technology along the delivery chain is progressing apace, leading to a greater use of digital technologies for social assistance. In that process, any use of mobile data must respect user consent, privacy, and security. Many countries have yet to adopt framework legislation to govern the use of mobile data for social protection purposes. Much work remains to be done to ensure that technology-enabled systems become tools for inclusion and universal social protection, and do not lead to further exclusion and retrenchment.

From the present analysis, several policy recommendations emerge to support the path towards effective, equitable, ethical, and inclusive social assistance. First, a strong political will and commitment is of course an important pre-requisite. Other recommendations are set out below and include making targeted investments, mitigating digital exclusion, improving data collection and analytics, establishing data protection safeguards, and bridging existing knowledge gaps. Some of these recommendations apply more broadly and are not necessarily linked to the usage of novel data and estimation techniques.

1. Invest now and invest more in adaptive social protection

Timely investments in state capacity are required to enable adaptive social protection, and to strengthen anticipatory mechanisms for better crisis preparedness. Experience during the pandemic demonstrated how functional social protection mechanisms, which typically serve as the backbone of regular social assistance programs, can be leveraged for the purpose of emergency support. The integration of novel technology into broad-based social protection systems will require substantial and conscious investment in at least four areas: (i) public infrastructure—both traditional and digital—including foundational ID, data integration/interoperability, social registries, household surveys, and payment systems, (ii) personnel capacity with the aim of attracting, training and retaining digital experts, which will require change management in public administration, (iii) preparation of the legal and regulatory conditions that ensure to roll-out programs in a timely, ethical and effective manner, and (iv) revisiting the institutional framework and arrangements with MNOs and other relevant actors that can enable or facilitate the usage of novel data sources and artificial intelligence for social services. This would include establishing the necessary partnerships [53] with the private sector and civil society to access data, tap into digital ecosystems, and collaborate with communities able to mitigate digital exclusion. These four areas should be readied and tested during non-crisis times, anticipating future shocks that would require a nimble and quick scale-up of social assistance measures.

2. Mitigate risks of digital exclusion through combined approaches

Novel data sources and processing methods are convenient and beneficial for most people with access to phones, lowering administrative costs, expanding outreach, and speeding up responses. However, no targeting approach can be 100% inclusive—and when using mobile and digital approaches there will always be some degree of digital exclusion. Satellite images provide actionable insights, but are limited by their long-range, relatively static perspective. Call detail records can be a useful tool for reaching

53 For a more structured process for data access and analysis, establishment of institutional arrangements to support scale and ensure sustainability is needed. For a scalable approach that beyond a single use case or program, establishing some form of institutional arrangement that can establish a structured process for data access and analysis will be key. Considering where the onus of responsibility lies for various steps in the process, from data access to management or handling and analysis will be critical not only to ensure sustainability but also to build trust across stakeholders, including beneficiaries. Such an arrangement could be leveraged to raise awareness, avoid duplication of efforts across social assistance payment streams and establishing a nodal point of contact for coordination to ensure that the data being leveraged is not mishandled and made vulnerable to bad actors.

those with mobile phones, but coverage gaps mean that not everyone can be reached. Therefore, to reach all those in need, multiple and diversified channels should be used, and support provided for those excluded from digital communications. For universal, equitable and ethical social protection, alternative access mechanisms need to be in place, even though traditional manual approaches are more costly and time consuming. These mechanisms are necessary to reach the most vulnerable, who are disproportionately affected by various shocks and are typically the focus of regular social assistance programs. In addition, the rise of smartphones across the world means that the days of using CDR for welfare targeting are already numbered. Therefore, social protection systems must constantly integrate new technologies and continuously adapt to changing contexts.

3. Improve data collection, data analytics and data systems

While novel data sources have significant potential, administrative data and public data systems will continue to play an essential and irreplaceable role. As discussed in this paper, countries with integrated and inclusive registries were able to respond to COVID-19 more efficiently (e.g., Turkey, Chile, Jordan, Pakistan). In many other countries, existing databases were used to accelerate data integration and interoperability efforts (e.g., Morocco using their informal health insurance database, Bangladesh using their garment sector registry).

Improving data collection for traditional household surveys can be an important foundation for both traditional and innovative approaches. Although new data sources may partly obviate the need for slow and costly data collection at the applicant and beneficiary levels, household survey data remain essential. The use of ML and AI based on novel data sources continues to rely on traditional household surveys for ground truth data in calibrating targeting models.[54]

More frequent household survey data linked with mobile data may enable the estimation of dynamic welfare models and machine learning. Similarly, regular and more extensive geo-located household surveys can help realize the potential of satellite-based methods. As demonstrated by the Togo case, collecting mobile phone information and CDR data (of all adults or heads of households) for the sole and explicit purpose of matching it with mobile data in the future, could facilitate the estimation of changes in welfare resulting from shocks, and help re-calibrate social assistance programs (Aiken et al, 2023).[55] The same data could help improve the accuracy of phone-based methods, especially if data sources can be linked to mobile data in an ethical and privacy-preserving manner in line with the principles of informed consent.

4. Establish strong data protection safeguards in social protection systems

Data protection has multiple facets including fairness, transparency, accuracy, and accountability consent, and security. Although these apply to all social protection systems, they are especially salient for those that leverage digital technology and novel data. A failure in any one aspect may compromise the entire system. In addition, many of these issues will be difficult to tackle once the system is up and running. Embedding data protection principles into the very design of social protections systems is essential for trustworthy, fair, transparent, and equitable programs.

54 The cost of sample surveys is insignificant compared with their benefit for social protection systems, whether they follow a traditional or novel approach. No less importantly, they inform many other crucial areas of public policy and there are technical advancements (ex. estimating welfare with a smaller number of variables, remote data collection) to ease data collection. Frequent data collection, instead of the often-long intervals between rounds in many LIC/MICs, ought to become the norm.

55 A caveat is that there is no evidence on how this could affect mobile phone usage behavior.

5. Enhance evidence-based decision making through monitoring and evaluation programs

Program monitoring and evaluation (M&E), together with data analytics, will help fill knowledge gaps and promote evidence-based decision making. From the present review of country approaches, several questions emerge that need to be answered before innovations can be adopted more widely across social protection programs, namely (i) the comparative accuracy of phone-based and satellite-based methods (and their combination) versus proxy means tests and community-based targeting, (ii) the demographics of SIM card usage, including use of multiple SIMs per household, (iii) data requirements for accurate satellite-based geographic targeting, (iv) reliable measurement and estimation methods of welfare during crises and shocks, and (v) evidence of using such estimations to trigger emergency programs for adaptive social protection.

ANNEX
Appropriateness and minimum conditions of different targeting methods

Method	Geographic targeting	Demographic targeting	Means testing	Hybrid means testing	Proxy means testing	Community-based targeting	Self-targeting	Lottery
What is it?	Location determines potential eligibility for benefits When working in isolation, all the population living in the area of intervention are considered eligible However, it is commonly used as a first phase in targeting, to allocate caseload, with some other method used to further reduce the pool of the eligible	Uses age or other demographic characteristic to determine eligibility for benefits Can be applied in isolation or as an additional criterion in mixed methods When used alone, sometimes referred to as universal child grants or social pensions	Compares a measure of the income, consumption, and/or wealth of the social assistance unit (family, household) with eligibility thresholds Often, but not always, verifies a substantial portion of the information with independent sources	Measures and verifies sum of income, consumption, and assets, as in a means testing Imputes the value of other flows where they are not verifiable Imputations often fairly simple—for example, marginal productivity per hectare in agriculture, or unit of livestock, or assumption of a few days a month of low wage employment for day laborers although more sophisticated estimations may be used	Uses easy-to-verify characteristics or proxies (for example, composition of the social assistance unit, size, quality or location of its dwelling, its assets) to predict money metric well-being Weights derived from statistics/ econometric models of various degrees of sophistication	Uses organized local-level groups composed by local leaders, civil society, and government officials; group members are from and are very active in the community, and they decide who in the community should benefit	Anyone may participate, but some element of the program makes it unattractive to the less needy For example, a low wage may be offered in exchange for work on community infrastructure or service projects Nutritious but less preferred foods (broken rice, coarse flours, less attractive packaging)	A random process to ration program resources or slots for enrollment among individuals, households, or communities that are all eligible
When is it appropriate?	When differences in poverty, vulnerability, or implementation capacity have a sharp geospatial gradient May work best when people don't move too often or easily between the delineated areas	To address right base approach When the program is focused on biological or social vulnerabilities of children or elderly When age or family structure are highly correlated with poverty or vulnerability to poverty	When income, consumption, or wealth are relatively easy to verify—for example, through data matching with other government-held records When labor markets are highly formal	When a moderately high share of incomes—but not all—can be verified When labor markets have moderate informality	When informality is too high to make means testing or hybrid means testing viable but household specific rankings are still desired	To address program administrator's myopia and lack of knowledge about the community To promote community engagement and improve accuracy at lowest level	When the program design is conducive	To address the fact that program's budget is not enough for covering the total number of eligible claimants When it is difficult to rank among many similar claimants, or when such rankings would not be socially accepted

Minimum conditions	Data to build geospatial analysis on indicators such as well-being, poverty, social development, access to services, infrastructure, climate, soil, and so forth, including big data	Good coverage of identification documents to verify age For poverty reduction, monetary poverty or vulnerability must be highly correlated with the predefined category	High levels of literacy and documentation that can be used as proof of declared information; high capacity levels of staff to properly collect the information required and to digitize the self-declared information Has effective verification process, including home visits and/or interoperability of government databases	High levels of literacy and documentation that can be used as proof of declared information High capacity levels of staff to properly collect the information required and to digitize the self-declared information Has effective Verification process, including home visits and/or interoperability of government databases Plausible models can be built to impute income from household assets or informal labor	Administrative capacity to interview on demand potentially eligible applicants and/or to conduct survey sweeps in high poverty areas	Requires a strong, small, and cohesive community structure (hard to use for larger groups where knowledge of one another is limited) Requires effective outreach and capacitation to local actors that will be running the process and supporting program implementation	Subsidies: clear dichotomy in place so that the selected goods are not attractive to the nonintended population but available to the intended population Temporary employment; type of work or benefit amount or goods is not attractive to the nonintended population	Have a transparent process to randomly select the beneficiaries
Pros	Simple to apply and does not create social tensions among close neighbors, though it may across jurisdictions Errors of exclusion are transparent to the political process	Corresponds with notions of deservingness in most places Unlikely to be Stigmatizing Relatively simple to implement and for people to understand	Accurate metric for well-being when its development follows basic standards and minimum conditions It is sensitive to quick changes in well-being, either idiosyncratic or covariate	Reliable metric for predicting full well-being when its development follows basic standards and minimum conditions It is somewhat sensitive to quick changes in well being; the formulae for imputations for informal incomes may need to be adjusted in response to covariate shocks	A statistically plausible and replicable method to rank households when informality is high	Benefits from the locals and their knowledge of the community to identify the population of interest Generates local level buy-in because the locals feel they are part of the process; improves acceptability of the program	Little administrative effort given to eligibility (other aspects of running public works programs are demanding, as are the logistics of food distribution)	Transparency Fast and inexpensive; replicable in any environment, both rural and urban Requires minimal administrative capacity

Cons	There are likely to be clear errors of exclusion as poverty, vulnerability, or other forms of need will exist in excluded territories Can encourage migration to the program's treated areas	May be only mildly correlated with money-metric measures of welfare	High requirements for data available for verification Relies on qualified administrative staff	Relies on quality of data available for verification Relies on qualified administrative staff Contains some statistical error from the imputations	Contains statistical error; formulae may be more or less precise depending on context and statistical methods Difficult for people to understand; lack of transparency Insensitive to quick changes in well-being	Local actors can have own preferences and consequently bias the process toward certain groups Different communities may implement guidelines differently May reinforce existing power structures and replicate patterns of exclusion Social tensions within the population and local actors can arise	May not reduce the number of people desiring to participate enough to meet a budget or implementation capacity; additional layers or filters or other eligibility criteria may be needed	Not to be used when population is heterogeneous enough so that the difference between two eligible people does matter and there are reasonable means to rank more finely
Shock responsiveness	Often used to respond to natural disasters May be used to select areas for resilience building activities in disaster-prone areas, or in association with early warning system	Not shock-responsive by nature, but if high coverage programs exist, benefits can be increased quickly	It is sensitive to quick changes in well-being, either idiosyncratic or covariate	Moderately sensitive to quick changes in well-being	Insensitive to quick changes in well-being, and consequently not suitable for addressing shocks and for shock response without adaptations	It seems that communities often incorporate local knowledge of which households are suffering idiosyncratic shocks such as illness, accident, or unemployment; for covariate shocks like natural disaster, may involve a new assessment exercise to update information	People may change their calculus about self-targeting as their need increases or decreases—for example, a family that prospers may no longer report for public works, or one that loses its preceding job or faces a downturn may seek additional days of public works	Can be used for some shock responsive programs by ensuring that those who feel affected would apply and would have equal chance of entering in the program

Source: Grosh et al. 2022.

References

Aiken, Emily, Guadalupe Bedoya, Joshua Blumenstock, and Aidan Coville. 2022a. "Program Targeting with Machine Learning and Mobile Phone Data: Evidence from an Anti-Poverty Intervention in Afghanistan." arXiv:2206.11400. https://arxiv.org/abs/2206.11400.

Aiken, Emily, Suzanne Bellue, Joshua Blumenstock, Dean Karlan, and Christopher R. Udry. Estimating Impact with Surveys versus Digital Traces: Evidence from Randomized Cash Transfers in Togo. No. w31751. National Bureau of Economic Research, 2023.

Aiken, Emily, Suzanne Bellue, Dean Karlan, Chris Udry, and Joshua Blumenstock. 2022b. "Machine learning and phone data can improve targeting of humanitarian aid." Nature 603: 864–70. https://www.nature.com/articles/s41586-022-04484-9.

Aiken, Emily, Viraj Thakur, and Joshua E. Blumenstock. 2022c. Phone Sharing and Cash Transfers in Togo: Quantitative Evidence from Mobile Phone Data. In ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies (COMPASS) (COMPASS '22), June 29-July 1, 2022, Seattle, WA, USA. ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3530190.3534796.

Aiken, Emily and Ohlenburg., Tim 2023a," Novel Digital Data Sources for Social Protection: Opportunities and Challenges", GIZ.

Aiken, Emily and Ohlenburg, Tim. 2023, Slides for "GIZ Guidance Tool for the Use of Alternative Social Protection Systems". GIZ.

Batana, Yele Maweki, Alexandra Jarotschkin, Akapo Konou, Takaaki Masaki, Shohei Nakamura, and Mervy E. Viboudoublou Vilpoux. 2021a. "Profiling Living Conditions of the DRC Urban Population: Access to Housing and Services in Kinshasa Province." Policy Research Working Paper No. 9857. Washington, DC: World Bank. https://openknowledge.worldbank.org/handle/10986/36633.

Batana, Yele Maweki, Takaaki Masaki, Shohei Nakamura, Mervey E. Viboudoulou Vilpoux. 2021b. "Estimating Poverty in Kinshasa by Dealing with Sampling and Comparability Issues." Policy Research Working Paper No. 9858. Washington, DC: World Bank. https://openknowledge.worldbank.org/handle/10986/36631.

Blumenstock, Joshua, Gabriel Cadamuro, and Robert On. 2015. "Predicting poverty and wealth from mobile phone metadata." Science 350, no. 6264: 1073–76. https://www.science.org/doi/10.1126/science.aac4420.

Bouguen, A., Huang, Y., Kremer, M., & Miguel, E. (2019). Using randomized controlled trials to estimate long-run impacts in development economics. Annual Review of Economics, 11, 523-561.

Bughin, Jacques, Jeongmin Seong, James Manyika, Michael Chui, and Raoul Joshi. 2018. "Notes from the AI Frontier: Modeling the Impact of AI on the World Economy." report, McKinsey & Company, Washington, DC. https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-aifrontier-modeling-the-impact-of-ai-on-the-world-economy#.

Canva 2023 Uplifting people from extreme poverty: the next step in our journey with GiveDirectly https://www.canva.com/newsroom/news/givedirectly-update-2023/.

Camilletti, Elena (2020). "Social Protection and Its Effects on Gender Equality: A literature review", Office of Research – Innocenti Working Paper, WP-2020-16, December 2020.

Carillo, A., Cantú, L.F., Tejerina, L. and Noriega, A. 2021. "Individual explanations in machine learning models: A case study on poverty estimation', Prosperia.

Center for Global Development. 2019. "Towards Real-Time Governance: Using Digital Feedback to Improve Service, Voice, and Accountability." CGD Note, Center for Global Development, November. https://www.cgdev.org/sites/default/files/towards-real-time-governance-using-digital-feedback-improve-service-voice.pdf.

Chi, Guanghua, Han Fang, Sourav Chatterjee, and Joshua E. Blumenstock. "Microestimates of wealth for all low-and middle-income countries." Proceedings of the National Academy of Sciences 119, no. 3 (2022): e2113658119.

Corral, Paul; Henderson, Heath; Segovia, Sandra. 2023. "Poverty Mapping in the Age of Machine Learning". World Bank Policy Research Working Paper 10429. https://openknowledge.worldbank.org/server/api/core/bitstreams/628af49e-8f46-4679-a04a-0a9f00c8684c/content.

Digital Public Goods Alliance 2023 https://digitalpublicgoods.net/.

GiveDirectly 2022 "Canva partnership tackling extreme poverty in Malawi one year on" https://www.givedirectly.org/canva/.

GiveDirectly 2023 PPT "Internal Learning Synthesis: Malawi Khongoni Pilot—Teletargeting Retrospective"

Elbers, Chris & Jean O. Lanjouw & Peter Lanjouw, 2003. "Micro-Level Estimation of Poverty and Inequality," Econometrica, Econometric Society, vol. 71(1), pages 355-364, January.

Fathom. 2020. "Fathom-US 2.0 Flood Hazard Data." https://www.fathom.global/product/flood-hazarddata-maps/fathom-us/.

Fisker, Peter Simonsen; Gallego-Ayala,Jordi Jose; Malmgren Hansen,David; Pave Sohnesen,Thomas; Murrugarra,Edmundo. 2022. Guiding Social Protection Targeting Through Satellite Data in São Tomé and Príncipe (English). Social Protection and Jobs Discussion Paper; no. 2212 Washington, D.C.: World Bank Group. http://documents.worldbank.org/curated/en/099135010252263269/P176471047cadc0240ba5d08ef8a2bc86b3.

Gentilini, Ugo; Almenfi, Mohamed; Orton, Ian; Dale, Pamela. 2020, 2021, and 2022. Social Protection and Jobs Responses to COVID-19: A Real-Time Review of Country Measures. World Bank, Washington, DC. © World Bank. https://openknowledge.worldbank.org/handle/10986/33635 License: CC BY 3.0 IGO.

Gentilini, Ugo. 2022. Cash Transfers in Pandemic Times: Evidence, Practices, and Implications from the Largest Scale Up in History. Washington, DC: World Bank. © World Bank. https://openknowledge.worldbank.org/handle/10986/37700 License: CC BY 3.0 IGO.

Grosh, Margaret; Leite, Phillippe; Wai-Poi, Matthew; Tesliuc, Emil. 2022. Revisiting Targeting in Social Assistance: A New Look at Old Dilemmas. Human Development Perspectives;. © Washington, DC: World Bank. http://hdl.handle.net/10986/37228 License: CC BY 3.0 IGO.

Gualavis, Melany; Newhouse, David. 2022. Integrating Survey and Geospatial Data to Identify the Poor and Vulnerable Evidence from Malawi. World Bank Policy Research Working Paper 10257. https://documents1.worldbank.org/curated/en/099622212082223165/pdf/IDU1023c6619164f6141491b97f1a1153a5a2eb7.pdf.

He, Y., Thies, S., Avner, P., & Rentschler, J. 2021. Flood impacts on urban transit and accessibility—A case study of Kinshasa. Transportation research part D: transport and environment, 96, 102889.

Huang, L. Y., Hsiang, S. M., & Gonzalez-Navarro, M.. 2021." Using satellite imagery and deep learning to evaluate the impact of anti-poverty programs". No. W29105. National Bureau of Economic Research.

ILO. 2021a. World Social Protection Report 2020-22: Social Protection at the Crossroads—In Pursuit of a Better Future. Geneva: ILO

International Social Security Association (ISSA). 2021. The application of chatbots in social security: Experiences from Latin America. Blog post. https://ww1.issa.int/analysis/application-chatbots-social-security-experiences-latin-america.

International Social Security Association (ISSA) 2022. Artificial intelligence in social security institutions: The case of intelligent chatbots. Blog post. https://ww1.issa.int/analysis/artificial-intelligence-social-security-institutions-case-intelligent-chatbots.

Jain, N., Using AI and ML in National Health Insurance Scheme of India (Anti-fraud framework), presentation slides Indo German Programme on Universal Health Coverage, 17 January 2021.

Lawson, Cina; Koudeka, Morlé; Cárdenas Martínez, Ana Lucía; Alberro Encinas, Luis Iñaki; Karippacheril, Tina George. 2023. Novissi Togo: Harnessing Artificial Intelligence to Deliver Shock-Responsive Social Protection. Social Protection and Jobs Discussion Papers; 2306. © World Bank, Washington, DC. http://hdl.handle.net/10986/40405 License: CC BY-NC 3.0 IGO

Lindert, Kathy, Tina George Karippacheril, Inés Rodriguez Caillava, and Kenichi Nishikawa Chavez. 2020. "Sourcebook on the Foundations of Social Protection Delivery Systems." Washington, DC: World Bank. https://openknowledge.worldbank.org/handle/10986/34044.

Lokshin, Michael. and Umapathi, N. 2022. AI for social protection: Mind the people. https://www.brookings.edu/blog/future-development/2022/02/23/ai-for-social-protection-mind-the-people/.

Lowe, Christy. 2022. "The digitalisation of social protection before and since the onset of Covid-19: opportunities, challenges and lessons". ODI Working Paper.

Marin, Georgina and Palacios, Robert. 2020. "The Role of Digital in the COVID-19 Social Assistance Response." G2PX. The World Bank Group. http://documents.worldbank.org/curated/en/099830009302217091/P1731660f8c52f062092ac00d53c648bac7.

Matthews, T., Gronewald, C. and Moolman, B. 2020. "It's a lifeline, but it's not enough: The COVID-19 social relief of distress grant, basic income support, and social protection in South Africa", Black Sash.

Metz, Anna, Marin, G., Marskell, J., Clark, J., and Karpinski, K. 2022. A Digital Stack for Transforming Service Delivery ID-Payments and Data Sharing (English). Washington, D.C.: World Bank Group. http://documents.worldbank.org/curated/en/099755004072288910/P1715920edb5990d60b83e037f756213782.

Mukherjee, Anit Nath; Bermeo Rojas, Laura Ximena; Okamura, Yuko; Muhindo, Jimmy Vulembera; and Bance, Paul G. A. 2023. Digital-first Approach to Emergency Cash Transfers: STEP-KIN in the Democratic Republic of Congo (English, also available in French). Social Protection and Jobs Discussion Paper; no. 2302 Washington, D.C.: World Bank Group. http://documents.worldbank.org/curated/en/099935104272316767/IDU05debcd9500bf004a580a48b0c6c201068bdc.

Ohlenburg, Tim. 2020. AI in Social Protection—Exploring Opportunities and Mitigating Risks. Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ).

Pople, A., Hill, R., Dercon, S., & Brunckhorst, B. (2021). Anticipatory cash transfers in climate disaster response. CSAE Working Paper.

Senona, E., Torkelson, E. and Zembe-Mbakile, W.. 2021. Social protection in a time of COVID: Lessons for basic income support, Black Sash.

Smythe, I., and Blumenstock, J. 2022. Geographic microtargeting of social assistance with high-resolution poverty maps. PNAS Research Article 2022 Vol.119. https://www.pnas.org/doi/full/10.1073/pnas.2120025119.

Tesliuc, Emil, Lucian Pop, Margaret Grosh, and Ruslan Yemtsov. 2014. Income Support for the Poorest: A Review of Experience in Eastern Europe and Central Asia. Directions in Development: Human Development. Washington, DC: World Bank. http://hdl.handle.net/10986/18886.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society Series B: Statistical Methodology, 58(1), 267-288.

United Nations Development Programme (UNDP) (2020) Artificial Intelligence is leaving no one behind: Accessible Chatbot for deaf and hearing-impaired persons. Blog post. https://www.undp.org/arab-states/stories/artificial-intelligence-leaving-no-one-behind-accessible-chatbot-deaf-and-hearing-impaired-persons.

World Bank. Multiple years. World Bank Open Data. https://data.worldbank.org/.

World Bank. 2009 and 2023. Classification of Fragile and Conflict-Affected Situations. https://www.worldbank.org/en/topic/fragilityconflictviolence/brief/harmonized-list-of-fragile-situations.

World Bank. 2020. Artificial Intelligence in the Public Sector: Maximizing Opportunities, Managing Risks. Equitable Growth, Finance and Institutions Insight. © World Bank, Washington, DC. http://hdl.handle.net/10986/35317 License: CC BY 3.0 IGO.

World Bank. 2022a. Poverty and Shared Prosperity 2022: Correcting Course. Washington, DC: World Bank. doi:10.1596/978-1-4648-1893-6. License: Creative Commons Attribution CC BY 3.0 IGO. https://doi.org/10.1596/978-1-4648-1893-6.

World Bank. 2022b. Digital cash transfers: current state and scale up potential. Mimeo.

World Bank. 2022d. Deep Dive into the Ecosystem for the Delivery of Social Assistance Payments: Türkiye Case Study (English). Washington, D.C.: World Bank Group. http://documents.worldbank.org/curated/en/099010011072245762/P17316601c8b0401b0ab340214079f623b6.

World Bank. Mimeo. Distribution of households by decile, by NSR and RRR (As a part of Technical Assistance Program).

Yoshida,Nobuo; Takamatsu,Shinya; Yoshimura,Kazusa; Aron,Danielle Victoria; Chen,Xiaomeng; Malgioglio,Silvia; Shivakumaran,Shivapragasam; Zhang,Kexin. 2022. The Concept and Empirical Evidence of SWIFT Methodology (English). Equitable Growth, Finance and Institutions Insight Washington, D.C.: World Bank Group.

https://documents.worldbank.org/en/publication/documents-reports/documentdetail/099547109302235758/idu04a3a086c0a2da04853084b10e855253105f9.

worldbank.org/g2px

WORLD BANK GROUP
Social Protection & Jobs

G2Px
PAYMENTS THAT EMPOWER