



# Factors affecting the usefulness of existing social protection databases in disaster preparedness and response

Valentina Barca and Clare O'Brien

**How can social protection systems be used in disasters, as a complement to, or substitute for, humanitarian assistance? Oxford Policy Management led a research project investigating this question, looking at the role of social protection in both mitigating the impact of large-scale shocks and supporting households after a crisis hits. We identify factors that can help and hinder effective disaster response, and consider how social protection actors collaborate with others working in humanitarian assistance and disaster risk management (DRM). This policy brief specifically focuses on the characteristics of existing social protection databases that enhance or limit their potential use in emergencies. It draws on examples from across five of the project's case studies: Lesotho, Mali, Mozambique, Pakistan and the Philippines.**

Many social protection interventions hold data on the households or individuals they support. These interventions might include cash transfer programmes (such as child benefits, old-age pensions or poverty-targeted cash assistance), public works programmes, social insurance or assistance with expenditure on health care.<sup>1</sup> Information about a wider population group – including potential future recipients or households who have been assessed but classified as ineligible – is also often collected, but not always stored and maintained. Moreover, recent years have also seen a rapid acceleration in integrated approaches to data and information management for social protection, in order to provide a coordinated and harmonised response to the multi-dimensional vulnerabilities faced by individuals across a life-cycle.

Stakeholders are recognising that, where a social protection system has pre-positioned data, it may be able to contribute to better preparedness for disasters – and to improved shock-response. For example, providing top-ups to social protection beneficiaries at a time of emergency ('vertical expansion') involves using existing data. Similarly, 'horizontal expansion' (reaching a wider group of beneficiaries) could benefit from existing data on potential beneficiaries, while new programmes could also 'piggyback' on existing social protection datasets, rather than collecting data from scratch.

Specifically, there is a potential (subject to the challenges set out in this brief) for such data to improve the efficiency of emergency response by promoting better understanding of households' vulnerability, reducing duplication of data by multiple agencies, improving timeliness of identification of individuals for assistance (leading to timely support if acted on promptly), and increasing coverage and predictability of assistance for households. Whether or not this improves programme impact depends on how the data is used (see Section 3) and on its ability to identify and 'target' those in need in any given shock, compared with alternative approaches – for example the collection of new data in the aftermath of a disaster (see Section 2).

In this Brief we explore the factors affecting the usefulness of social protection data – focusing on non-contributory social protection<sup>2</sup> alone as this was the focus of our research across the five case study countries – in four sections:

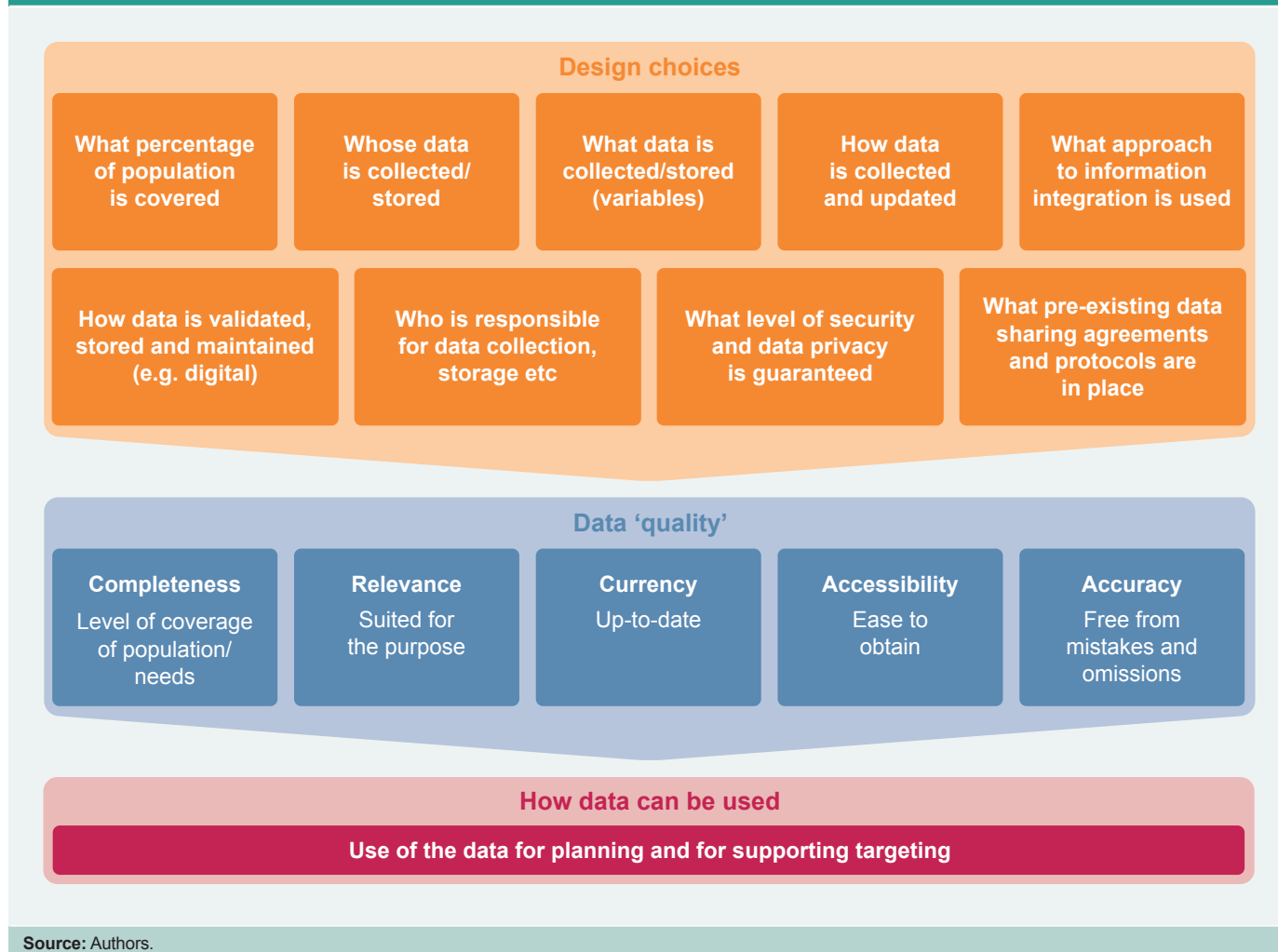
1. We first provide some clarification on different types of social protection databases, as this has very different implications for shock-responsiveness.
2. We discuss the five dimensions of social protection 'data quality' that enhance or limit their use in emergencies: completeness, relevance, currency, accessibility and accuracy. We explain the meaning of each of these terms, providing examples for each, and briefly summarise the design choices (who collects the data, how often, how it is stored etc.) that affect data quality in each of the dimensions.
3. We explore what can be done with this data for disaster preparedness and response.
4. We provide some final conclusions and recommendations.

<sup>1</sup> Not all social protection programmes have this type of data. For example, school feeding programmes may not maintain records of the precise children supported. A subsidy programme may have no information on the households or individuals who benefit from it.

<sup>2</sup> In this briefing note, the term social protection will be used to refer to non-contributory social protection.

The nature and quality of social protection databases and information systems is so varied that it is meaningless to ascribe a generic role to their use in emergencies (see Section 1), and inappropriate to assert that they will always be of use: such a role can only be identified with reference to the particularities of the database(s) in the country and context under review. Figure 1 below summarises our conceptual framework for determining this role. It illustrates how design choices for social protection data collection, storage and management affect data quality, which drives the suitability of the database for planning and for supporting targeting. This framework holds true for use of that data for standard social protection provision, and is put to the test when analysing use of the data for a different purpose: emergency preparedness and response. This is further explained in the sections below.

**Figure 1: Design choices for social protection data collection, storage and management that affect data quality**



## 1. Social protection databases come in many shapes and sizes

Social protection data on individuals and/or households is often digitally held within a database, a static container that helps to organise, store and retrieve large amounts of data easily. Databases are also referred to as ‘registries’ or ‘data repositories’, all de facto interchangeable terms. They come in many different shapes and sizes primarily depending on: a) the type/s of programme/s they serve and their design (e.g. targeting approach, coverage); b) countries’ needs and context; c) whether they play a role at programme level or at national/integrated level; and d) the chosen approach to data collection (see also Section 3). Table 1 below summarises the main approaches to storing social protection data. It should be noted many of these are complementary and coexist within the same country as they perform different functions. Moreover, these databases are only one part of a wider information system – most often served by a tailored software that transforms data retrieved from the database (and elsewhere) into usable and useful information. At programme level, this is referred to as a Management Information System (MIS).

Table 1: Main approaches to storing social protection data

Name	Number of programmes covered		Electronic records available		Brief description
	One	Multiple	Beneficiaries	Non-beneficiaries	
Paper records	●				Registration records held manually. Data which is not stored digitally is not easily retrievable.
Programme database (beneficiaries only)	●		●		Electronic data is retained only on beneficiaries. Data on households/individuals who are registered but not enrolled are not digitised.
Programme database (beneficiaries and non-beneficiaries)	●		●	●	Comprehensive data on both beneficiaries and non-beneficiaries who have applied or been registered for a particular programme (successfully or unsuccessfully) is retained and maintained.
Integrated beneficiary registry		●	●		A database which integrates data from existing databases of several different social protection programmes (not necessarily from all programmes in the country). It provides a consolidated overview of data collected by different programmes, focusing on beneficiaries alone.
Integrated beneficiary and non-beneficiary registry		●	●	●	As above, but also integrating information on non-beneficiaries from existing social protection programme databases.
Social registry		●	●	●	A database created through a national and coordinated data collection effort that is then drawn upon by different social protection and other programmes (not necessarily by all programmes in the country). Contains data on potential beneficiaries of social assistance, most often containing socio-economic variables.
Virtual social registry		●	●	●	A technical approach to developing a social registry by sourcing data from existing government databases, making these interoperable: i.e. ensuring that these can 'talk to each other' by using a Unique ID (e.g. National ID number) as identifier for the household or individual.

Source: Authors, see also Barca (2017). Note: this table presents core typologies, large variations also exist within these.

All our case study countries have several **programme databases**, varying in size depending on the programme they serve, compiled and held by either the government or non-governmental agencies. Most retain digital data on both beneficiaries and non-beneficiaries, but, for example, Mozambique's Programa de Subsídio Social Básico (PSSB) cash transfer programme has an electronic database of beneficiaries only. Beyond these programme-level databases, Mali is in the process of developing an **integrated registry** (the Registre Social Unifié) that includes information on beneficiaries and non-beneficiaries from across several recent social protection and humanitarian programmes, and has plans to move towards a social registry (centralised data collection) approach. Pakistan, the Philippines and Lesotho all have a **social registry** developed on the back of one programme but not intended for the exclusive use of that programme (see also Figure 2):<sup>3</sup>

- In Pakistan the National Socio Economic Registry (NSER) covers 27 million households (85% of population) and serves over 30 federal and provincial social sector programs. It was developed under the Benazir Income Support Programme (BISP), which covers 20% of population.
- In the Philippines the Listahanan database covers 15.3 million households (77% of population) and serves some 60 programs from many government agencies. Listahanan was developed under the Pantawid Pamilyang Pilipino programme, which covers 21% of population.
- In Lesotho the National Information System for Social Assistance (NISSA) covers some 117,000 households (22% of population) and has been piloting integration with selected programmes. NISSA was developed under the Child Grant Programme, which covers 5% of population.

## 2. When is social protection data fit for use in emergency contexts?

We briefly discuss the five key dimensions of 'data quality' for social protection databases (and related information systems) that enhance or limit their potential uses (see Section 4): completeness (coverage), data relevance (sometimes classified as 'appropriateness'), data currency, accessibility and accuracy (with a focus on usability).<sup>4</sup> While this framework could be used to assess usefulness of a database for standard social protection delivery, we use it here to assess usefulness to enhance preparedness and response to shocks.

### Completeness (coverage)

**Completeness refers to the number of records compared with what would be perceived as a full set of records – 100% population, or 100% of those in need.** An existing social protection database may assist an emergency response if the data covers all of those affected by the shock, or at least a high enough proportion. Distinctions need to be made between data on beneficiaries and registered non-beneficiaries:

- In the case of databases that cover beneficiaries only, it is extremely rare that they would cover 100% of the population in any given area, as beneficiaries are almost always subsets of the population (e.g. the 'poor', older people, children, etc.). Programmes are also sometimes targeted geographically, exacerbating the coverage problem (e.g. when areas with no/low coverage are affected by a shock) – see also Box 1.
- Different databases have very different coverage of data on non-beneficiaries – and none can claim 100% coverage. Countries that have developed social registries serving several social protection programmes tend to have higher coverage of non-beneficiaries than those who have fragmented programme databases, as these have an explicit objective to retain data for use across several programmes and over time.

Figure 2 below compares and contrasts social protection databases in four of the research countries, showcasing completeness of social registry data where one exists (the green oval) and completeness of beneficiary data for one or more non-contributory social protection programmes (the red oval).<sup>5</sup> It also introduces an arbitrarily sized group of households potentially affected by any given shock, to show how existing data can be used for vertical expansion of the social protection scheme (top-ups to existing beneficiaries), horizontal expansion (temporary expansion of the caseload), or piggybacking on the social protection database to provide an emergency response.

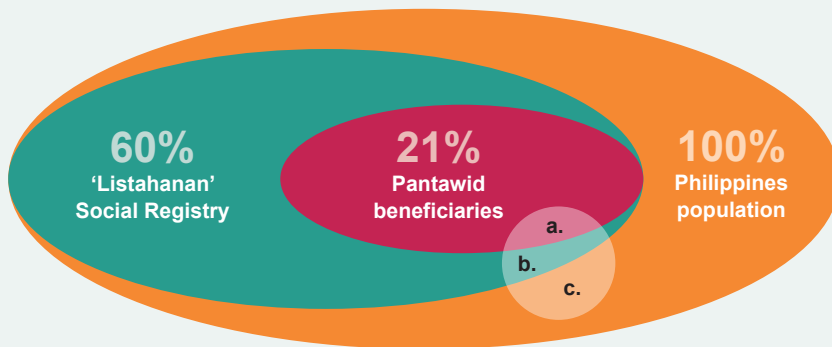
<sup>3</sup> It is important to note that each country calls their system a different thing, and names are not used consistently across countries, meaning it is important to focus on function and set-up (as per Table 1), not names.

<sup>4</sup> This framework is adapted from Wang, R. Y., & Strong, D. M. (1996) Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems* 12(4).

<sup>5</sup> Note that each country is likely to have several such programme databases.

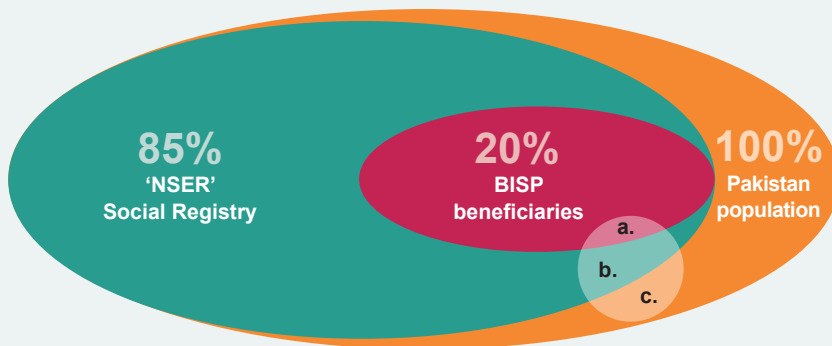
Figure 2: Coverage of selected social protection databases: four countries compared

**Philippines**



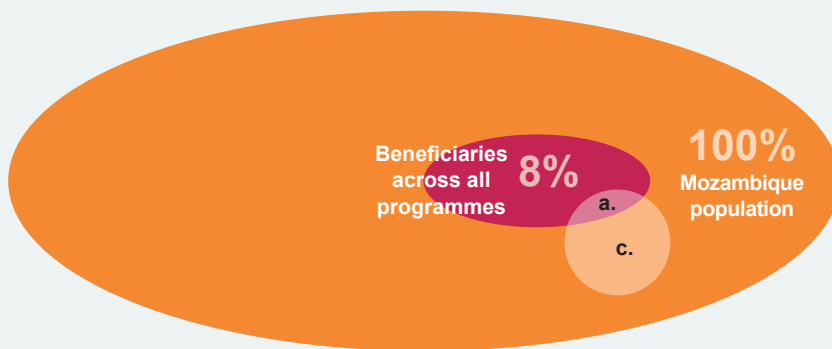
- a. Households that can be reached through **vertical expansion** or **piggybacking** (on the beneficiary databases)
- b. Households that can be easily reached through **horizontal expansion** or **piggybacking** (on the social registry)
- c. Households less easily reached through **horizontal expansion** or **piggybacking** (not covered by existing social protection databases)

**Pakistan**

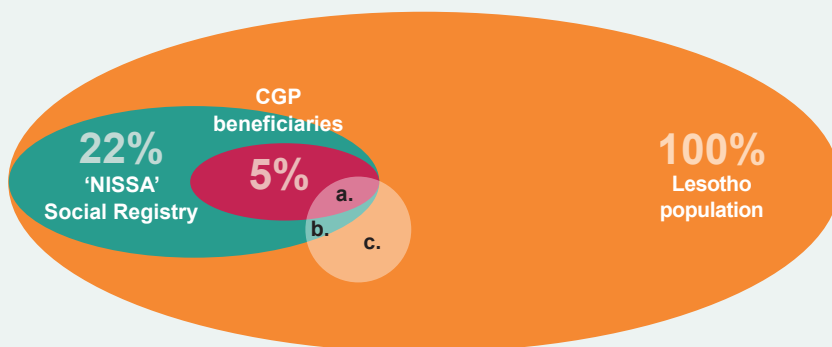


Households potentially affected by a shock

**Mozambique**



**Lesotho**



**Source:** Percentages refer to percentage of households covered by each. Please note these diagrams do not represent the totality of social protection databases in each country. Also note that Mozambique has no green circle as no data on non-beneficiaries is retained, while data on beneficiaries is held across several programme databases.

### Box 1: Lesotho's NISSA database – observations on completeness

The completeness of the NISSA database in Lesotho is a function of two aspects: its geographical reach and the percentage of households included within each geographical area:

1. **Geographically**, NISSA covered Lesotho's 10 districts at the time of the El Niño crisis in mid-2016, but only 36 of the 64 community councils (subdistricts) within them, some initially selected by lottery in the absence of sufficient resources to introduce the Child Grant Programme (which launched the database) to the entire country at once.
2. **The percentage of households** covered in each community council has changed over time. The initial data collection approach, used since 2010 in 33 community councils, aimed to collect data on 100% of households (with reviews showing coverage, in practice, was around 70%). Under a revised approach, that had been used in three community councils by mid-2016, a detailed survey is now only administered to households that the community identifies as poor (around 50% of households). This is not a failure of the 'new NISSA': it is simply that the underlying data collection approach was redesigned for the purpose of improving effectiveness and efficiency for other CGP programming, as well as its acceptability to communities (by starting with the community selection of poor households, rather than with a computerised poverty test). The redesign was not intended for use of the database in a crisis to identify previously non-poor households.

Source: : Lesotho Case Study.

### Relevance

**Data is relevant if it contains the variables required for the intended purpose.** Data collected for the provision of long-term social protection (i.e. another purpose) may not always be relevant in an emergency if it does not contain variables that comprehensively identify households in affected areas, and ideally that assess their needs:

- Where data on non-beneficiaries is adequately stored and maintained (see Table 1), it rarely contains a full set of operationally relevant variables – bank account numbers, full contact details, identification documents, geolocation, biometric data, recipient and alternate details, etc. – as these are almost always collected in the last stage of data collection during enrolment.
- Data on programme beneficiaries may contain more relevant and comprehensive variables, but is inherently limited to a smaller – and often specifically targeted – group, being those who have been deemed eligible for specific social protection programmes. That data will therefore be less complete (see section above).

Different types of social protection programmes have different information requirements (primarily linked to their approach to targeting), potentially affecting the relevance of their databases in a crisis. For example, a universal social pension (such as Lesotho's Old Age Pension) will need to retain less information than a poverty-targeted grant (such as BISP in Pakistan and Pantawid in the Philippines). A school feeding programme can piggyback on school enrolment lists in order to plan its resources, while a subsidy requires no database as such.

Typically, databases serving programmes that provide poverty-targeted social assistance (in cash or in kind) collect and store many variables. However, the types of variables collected for determining eligibility of households for targeted social protection programmes may be useful to assess chronic poverty, but may not be the best suited for identifying households who are vulnerable to disasters or other shocks, nor affected by a specific shock. For example, in Lesotho agencies providing emergency assistance commonly seek information on households' dietary diversity at the time of the shock as an indicator for prioritising levels of need; this variable will not be in a database of long-term well-being such as the NISSA. Pakistan is attempting to tackle some of the difference in variables by including data on climatic vulnerability in its new proxy means test, while also making efforts to provide geographic coordinates for all registered households.

### Data currency

**Data currency is the degree to which data are current (up to date), representing households' real circumstances at the required point in time.** It is of course impossible for standard social protection data to reflect the reality after a disaster, meaning some form of post-disaster revalidation is always required. The relevant factor is how up-to-date existing data is overall. For a registry to be fully effective – for both standard social protection provision and for emergency response – it should aim to offer dynamic inclusion of newcomers (e.g. migrants, new-borns), dynamic exclusion of those who have died or moved away, and dynamic management of transitory shocks (e.g. natural disaster, crop failure, unemployment, sickness, pregnancy) (Barca, 2017).

Given the nature of poverty and vulnerability, any system that bases targeting on an outdated and static snapshot will likely face serious challenges in providing support to those most in need,<sup>6</sup> leading to errors of inclusion and exclusion.

There were indications in our research countries that programmes can struggle to guarantee adequate data currency. Our research also showed that data currency depended on:

- **The specific approach to data collection:** mass national census surveys (as those for Listahanan in the Philippines, NISSA in Lesotho and NSER in Pakistan) are often scheduled every 2-3 years, but are rarely carried out within this timeframe – primarily because of the large budget and capacity commitment needed. For example, data collected for Lesotho’s NISSA database in 2010 had not been updated as of 2016. Ongoing, on-demand registration partially addresses this concern (while posing other challenges, including lower coverage, Barca 2017), which is why Pakistan is now piloting on-demand data collection to assess its feasibility given existing capacity.
- **Approaches to information management:** beneficiary data – used for implementation of specific programmes – tends to be more current than the underlying databases on potential beneficiaries, as these are ‘alive’ and updated through continuous contact. Whether these updates are then loaded onto the underlying database (e.g. the social registry) affects the currency of data overall. In Pakistan, for example, BISP data was perceived as more reliable and up-to-date than NSER data as it was being continuously used.
- **The underlying approaches to targeting:** for example, poverty targeted programmes require more frequent data updating, given the transitory nature of poverty.
- **The type of shock:** different shocks affect the extent to which household well-being is affected. Conflict or rapid onset disasters, for example, may cause widespread internal displacement, split up households and significantly change their material circumstances.

## Accessibility

**Accessibility refers to the ease for potential users** – most likely national or local government agencies and departments, or their partners – to obtain the data. Our research has shown this can vary widely depending on:

- **Who the users are:** information is power, meaning the sharing of data that has been collected at a high cost cannot be guaranteed. Challenges accessing data in our research countries were particularly acute for local levels of administration and non-government actors, where they had not collected it. Data sharing may be facilitated where relations between institutions have been built over time, and hampered where data has a political association.
- **What processes and authorisation levels are in place for data sharing:** for example, in Pakistan, pre-positioning NSER data for use by humanitarian actors and local authorities would require advance partnerships and memorandums of understanding with the national government (see Box 2). Direct access within Lesotho to the NISSA database during the El Niño crisis was confined to the capital, Maseru, among a limited number of staff in the Ministry of Social Development. This posed a challenge since the identification of specific households for disaster response was led by local rather than national authorities.

### Box 2: Pakistan’s IRC experiment

From its experience in operating cash transfers in Pakistan, the IRC has identified beneficiary selection and registration as both the largest time driver for programme start-up in humanitarian situations and a significant cost-driver for programme delivery.

In 2016, IRC conducted a small pilot research project in Sindh province to compare the efficiency of new data collection through community-based targeting (CBT) versus the use of the NSER as a pre-positioned database. Overall results suggest that this latter approach results in a higher level of operational efficiency (but roughly similar cost-efficiency): it could be administered in 16 days, more than twice as fast as CBT’s 35 days – but only if relevant MoUs for data sharing are already in place (the wait period to receive the data was 44 days).

Source: Pakistan Case Study.

<sup>6</sup> Costa Rica carried out a study to determine the length of validity of variables used for targeting. The conclusion was: ‘the variables associated with income and occupation have a higher dynamism and ought to be updated every 1 to 1.5 years, while the variables associated with housing and ownership of goods have a lower dynamism, so that an updating every 3 to 3.5 years is recommended’ (Irrázaval 2004).

- **Whether or not data is maintained and stored digitally:** Mozambique's PSSB only retains non-beneficiary data on paper, not easily accessible at a time of crisis.
- **What type of data interfaces are provided:** for example, web-service interfaces with allocated authorisation levels across all relevant stakeholders can enable instant access (and better secure data), versus ad-hoc data sharing via CD, email or USB stick (common practice in the research countries).
- **Capacity:** capacity in terms of (i) human resources to use the database (ii) availability of computers (iii) availability of electricity and network connection all affect accessibility. This is a particular issue at local levels of administration, where both staffing and resource constraints are most acute.
- **Existing provisions for data security and privacy:** Social protection programmes collect substantial amounts of personal information from potential beneficiaries, posing a risk of misusing or losing such information, potentially exposing households to further vulnerability. This may be a legitimate reason to restrict access.
- **Type of shock:** the challenges of accessing a database are compounded in a conflict or rapid onset natural disaster (e.g. destruction of infrastructure). In conflict situations, security concerns around the sharing of personal information are particularly worrying (see below).

### Data accuracy and usability

**Data is considered to be accurate if it is free from errors and omission.** Accuracy means that a high level of confidence can be placed on the data, affecting its wider credibility and ultimately its usability (which is also determined by the state of users' understanding of the data itself – e.g. data definitions, documentation and metadata).

On one hand, perceived accuracy is a function of the quality of data collection, verification (e.g. through supervisions and cross-checks with other databases) and validation (e.g. applying internal consistency checks, removing duplicates, ensuring standardisation of contents and formats, etc.). On the other, it is very much a function of the perceived trustworthiness of the institution responsible for collecting and housing the data: agencies may be hesitant about the quality of data collected by others.

### Box 3: Summary of design choices that affect data quality (and therefore its use) across these dimensions

We have touched on many design choices for standard social protection data collection, storage and management that affect data quality across each of the five dimensions cited above, and therefore the potential use of that data in an emergency context. It follows that adjustments to any of the design choices below can alter the usability of the data. We briefly summarise the main aspects here and in Figure 1 above. These are all affected by the purpose the data originally serves (e.g. the types of social protection programmes it is needed for):

- **What percentage of the population is covered:** as discussed, different social protection databases vary widely in their coverage of national population (e.g. beneficiary database of small targeted programme versus national social registry);
- **Whose data is collected and stored:** e.g. individual vs household level data, beneficiary vs non-beneficiary data, data on the 'poor' and 'vulnerable' or specific age groups (e.g. for a social pension) vs wider sub-sections of population;
- **What data is collected and stored:** e.g. extensive dataset on household socio-economic condition vs limited dataset with key variables; operationally relevant data collected at enrolment (e.g. bank account) vs basic data for eligibility determination; biometric information or not, etc;
- **How data is collected and updated:** any database is only as good as the data that it collects, hence the adage 'garbage in, garbage out'. For many social protection databases, the two most frequent approaches to data collection are: a) on-demand registration, which relies on households going to an office (or accessing an application/website) to apply; b) census-survey registration, which entails all or selected households in an area being interviewed at selected intervals. These both offer significant advantages and disadvantages (Barca, 2017) – meaning mature social protection systems tend to go with a combination of both and legislate a maximum of two years between data updates;



### Box 3: Summary of design choices that affect data quality (and therefore its use) across these dimensions (continued)

- **Approach to information integration:** e.g. whether data is linked to other databases (social protection and beyond) for data sourcing/verification/removing duplication, ideally through the use of a unique identifier (National ID or other);
- **How data is validated, stored and maintained:** what procedures are in place to guarantee data integrity, affecting trustworthiness;
- **Who is responsible for data collection, storage and management:** e.g. whether performed in-house by the lead social sector Ministry, through municipal offices, through other state institutions (e.g. Statistics Office) or through contracted third partners, affecting trust in its integrity;
- **Level of data security/privacy guaranteed:** depends on existing legislation and provisions, including adherence to international standards such as the United Nations Guidelines for the Regulation of Computerised Personal Data Files and ISO 27001;
- **What processes and authorisation levels are in place for data sharing:** e.g. web service versus batch processes; ad-hoc versus regular through Memorandum of Understanding.

There are trade-offs in these choices. For example, allowing organisations to feed their own updates into the database (e.g. if a programme implementer finds out about a change in the circumstances of one of its beneficiary households) can improve data currency, but may reduce trust in data validity if other users are not convinced about the quality of data collection by that organisation/programme.

### 3. What to do with the data when you have it

As discussed in the introduction, existing social protection data could be used for:

- planning purposes (e.g. estimating caseloads, aligning emergency response design, etc.);
- targeting before or after a shock.

In this brief, we focus on targeting as it is the most controversial: different options have different trade-offs. Regarding the use of beneficiary databases (e.g. for vertical expansion or piggybacking), there are three main options for targeting:

1. **Target all beneficiaries of a social protection programme within the crisis affected geographical area.** This can be pragmatic and improve timeliness in severe shocks where most households are likely to have been affected, especially the poorest and most vulnerable. It could lead to inclusion errors, however. This approach was used by WFP, UNICEF alongside government in the Philippines after Typhoon Haiyan – setting in place additional data collection to support affected households who were not in the Pantawid programme database.
2. **Target some of the beneficiaries.** For less severe crises, or where there is variation in impact, it may be appropriate to take the beneficiary list as a start from which household needs are verified.
3. **Target no beneficiaries.** Our research suggests that in some cases where beneficiary lists have been provided to organisations delivering emergency response, those households have been de-prioritised on the grounds that they already receive some help. This appears to have happened during the response to the El Niño drought in Mozambique in 2016, resulting from a lack of clear directive to agencies as to what to do with the list, leading to discretionary and inconsistent decisions.

As for the use of data on non-beneficiaries (e.g. for horizontal expansion or piggybacking), options could include:

1. **Drawing on variables in the database to select households.** This presupposes that households worst affected by a disaster can be identified from a certain variable, e.g. their wealth ranking; their geolocation (if GPS coordinates are collected and can be overlaid with hazard vulnerability maps); or the characteristics of a household member. If the criteria, and the circumstances in which they would be used, were planned in advance, such an approach would effectively pre-identify a cohort of households for assistance, enabling immediate targeting in a disaster. Pre-enrolment (e.g. collecting operationally relevant data) might also be possible as part of preparedness planning, and might enable a timely response in the manner of ‘no regrets’.

2. **The records could be used as a starting point for establishing a list of households, to whom further screening would be applied post-disaster** in order to identify which households have been worst affected.

This might enable a more rapid validation process.

However the data are used, gaps in coverage and quality mean complementary methods must be used to identify other households in need. Agencies must ensure that those excluded from a programme, or a registry, are not systematically excluded from the shock response.

#### 4. Conclusions and recommendations

Based on the evidence from the case studies analysed, this brief has provided a framework for assessing the potential role of prepositioned social protection data to respond to emergency contexts: an objective this data was rarely originally intended to serve. Our core conclusion is that – before using existing data at any cost – it will be essential for every country to carefully assess the main design choices of existing social protection datasets to assess how these affect their completeness, relevance, currency, accessibility and accuracy. It is also important to consider whether these aspects can be strengthened through adjustments to the way the data is collected, stored and managed.

A set of specific recommendations for governments and their counterparts emerge from this analysis:

- **Pre-emptively map completeness, relevance (e.g. variables covered) and data currency of all social protection databases** to assess the trade-offs of using existing data vs collecting new data in the aftermath of a crisis.<sup>7</sup> This would include: a) understanding overlaps and potential of each and every social protection database (not just flagship programme, e.g. including social insurance); b) mapping these against other potential data-sources (e.g. ID database, civil registry, municipal records, etc.).
- **Pre-emptively assess the extent to which the individuals or households targeted by existing social protection programmes overlap with those who are most vulnerable to shocks** (including coverage of these populations in disaster-prone areas), acknowledging that income and asset poverty is just one dimension of vulnerability. Importantly, the overlap may differ for different types of shocks: e.g. in slow-onset drought prone areas, chronic poverty may be strongly correlated with vulnerability to drought, while a similar relationship is less obvious for conflict.
- **Consider the prioritisation of disaster-prone areas in any planned expansion of coverage of existing databases** (perhaps even up to 100% of population in those areas), especially in contexts where shocks are recurrent (e.g. droughts, floods, cyclones). However, this can be very expensive, and the cost-effectiveness of such an approach would need to be considered. Moreover, it is essential to be mindful of good communication so as not to raise expectations among all households registered that they will receive assistance.
- **Consider the collection of extra variables (for some or all households) that are not immediately required by a social protection programme, but that may be operationally relevant in an emergency.** This necessarily has a cost, requiring additional time, money and resources to the implementer, and the inconvenience to the household of answering more questions. Whether or not this is an advisable preparedness measure depends on the relative costs and benefits of alternative data collection methods.
- **Pursue strategies to improve currency of social protection data, including adequate planning of the capacity and financial resources needed**, for example by: a) considering on-demand data collection approaches, b) ensuring periodic census survey efforts happen every 2-3 years at most; b) updating through integration of data from other sources.<sup>8</sup>
- **Ensure that protocols and agreements are in place to govern arrangements for timely access to data where appropriate.** In considering data sharing with third-party actors, it is essential to adhere to international data transfer and information privacy protocols, which legislate the collection, transfer and storage of information so as to ensure the integrity of data and confidentiality (see also Barca, 2017 not forthcoming). Moreover, it is essential to consider how information from national databases can be made available at local level countrywide, with staff trained and able to access it, if disasters are handled by local authorities.

<sup>7</sup> This would include an analysis of any existing emergency targeting systems and databases, as a key comparator.

<sup>8</sup> This could be done: a) from local to central level, allowing municipalities/lower implementation levels to update information whenever they are notified of changes (needs some level of control as this poses a risk); b) from other administrative databases such as the civil registry, which can be used to update information on household births and deaths; c) directly from citizens, online or on mobile apps (e.g. in Chile and Australia).

- **Ensure some form of revalidation exercise post-emergency, to reflect changing household conditions** (e.g. to track down displaced households and those with no ID/token, replace named carers for newly orphaned children, etc.). For example, this took 3 weeks after Typhoon Haiyan in the Philippines.
- **Ensure that, when giving top-ups to regular beneficiaries (vertical expansion) or horizontally expanding to households on a social registry, there is also a system for reaching other affected households** (area 'c' in Figure 2). For example, after Typhoon Haiyan, both WFP and UNICEF also offered cash benefits to non-Pantawid households separately through NGOs.
- **Ensure that even countries with more up-to-date, accurate, comprehensive, relevant data develop protocols and guidance as to how to use the data in the event of a crisis.** This is not to be left in the conundrum faced in Lesotho, Mozambique and the Philippines: do you purposely exclude existing beneficiaries because they're already covered, or purposely include them because you're sure they're vulnerable? This requires coordination, and an assessment of needs and benefit values across programmes – it is an issue that data alone cannot solve.

To conclude, no matter the convenience of using existing social protection databases for implementers, the most important consideration must be the impact of the overall response on communities affected by shocks.

## About the project

The Shock-Responsive Social Protection Systems study is a research programme (2015 to 2017) led by Oxford Policy Management (OPM), in consortium with the Overseas Development Institute (ODI), the Cash Learning Partnership (CaLP) and INASP. Its aim is to strengthen the evidence base as to when and how existing government social protection systems can better respond to shocks in low-income countries and fragile and conflict-affected states, thus minimising negative shock impacts and reducing the need for separate humanitarian responses. The research is funded by UK Aid from the UK government, as part of the UK Department for International Development's (DFID's) Humanitarian Innovation and Evidence Programme (HIEP). This policy brief was co-funded by the Australian Department of Foreign Affairs and Trade (DFAT), building on research it has commissioned on 'Data and Information Management for Social Protection'.

## Further information and references

This policy brief builds on the project's full case studies in Lesotho, Mali, Mozambique, Pakistan and the Philippines. These are all accessible on our project webpage. It also builds on DFAT-funded work on the role of integrated data and information management for social protection (Barca, V., 2017. Integrating data and information management for social protection: social registries and integrated beneficiary registries. Canberra). The views expressed are those of the authors and do not necessarily reflect the UK or Australian government's official policies.

*Suggested citation:* Barca, V. and O'Brien C. (2017) 'Factors affecting the usefulness of existing social protection databases in disaster preparedness and response. Policy Brief'. Oxford Policy Management, Oxford, UK.



Oxford Policy Management



In partnership with:



### Valentina Barca

Senior Consultant – Poverty and Social Protection Portfolio

Tel: +44 (0) 1865 207 300

Email: [valentina.barca@opml.co.uk](mailto:valentina.barca@opml.co.uk)

<http://www.opml.co.uk/projects/shock-responsive-social-protection-systems>