# WE HAVE THE DATA, LET'S USE IT BETTER

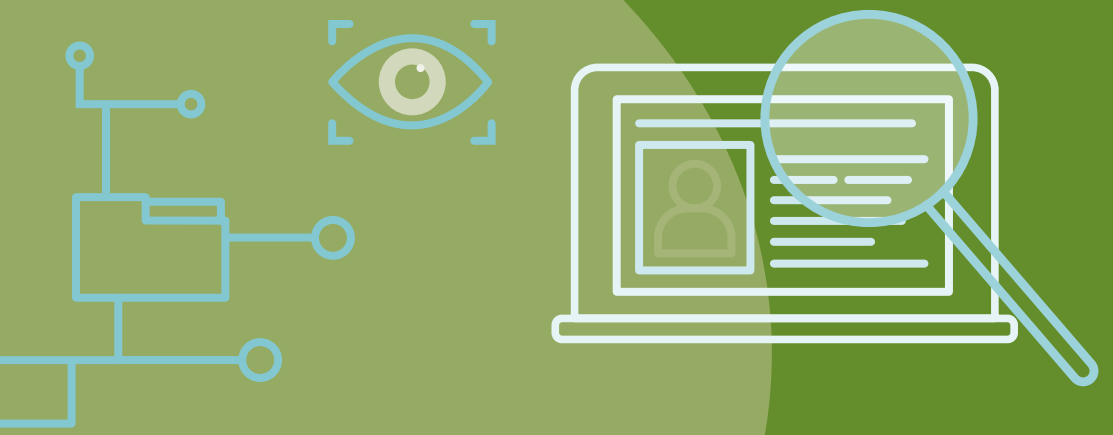
## Pushing the Boundaries of Social Protection
## Administrative Data Analysis and Use

Valentina Barca, Madhumitha Hebbar,
Charles Knox-Vydmanov, Ida Brzezinska

October 2023

giz Deutsche Gesellschaft
für Internationale
Zusammenarbeit (GIZ) GmbH

# CONTENTS



# 1 INTRODUCTION

## 1.1 SCOPE

**Data has the potential to impact millions of lives, as it has become critical to government decision making** (**World Bank, 2021**). Building a strong evidence base requires governments to tap into a range of data sources (**OMB, 2016**): population-level administrative data such as national identification systems and civil registries; programme-specific administrative data; general purpose surveys such as census surveys; data from specific studies and evaluations; and emerging non-traditional data sources such as mobility data and sensor data. Such an evidence base can inform all kinds of government policy and decisions, including those relating to social protection.

**This paper focuses on the better use of administrative data in the social protection sector** by social protection practitioners working in social sector ministries and partner institutions. Administrative data is "information created when people interact with public services", such as schools, health clinics, social protection systems, civil registration systems, or tax systems, and collated by the government (ADR UK, n.d.). The delivery of social protection[1] is increasingly managed digitally via programme management information systems (MISs) or integrated systems, such as social registries and integrated beneficiary registries (**Leite *et al.*, 2017**; **Chirchir & Barca, 2020**; **Lindert *et al.*, 2020**). While administrative data generated by these obviously has an immediate use – namely, identifying and delivering support to the most vulnerable – its potential is not limited to operational management alone.

**This paper investigates which *additional analytical and statistical uses* could usefully be made of social protection administrative data, beyond supporting routine operations and benefit/service delivery – especially when data is strategically analysed[2] and linked to other data.** In short:

- Given that a huge amount of data is being collected and stored, how can we ensure that insights and information from this data are not 'wasted' and are put to full use for policy and decision making?

1 Social protection encompasses social insurance, social assistance, active labour market policies and social care services.

2 Note that some of this analysis is regularly conducted and systematised by international organisations, including to measure progress against the Sustainable Development Goals, e.g. via the World Bank Atlas of Social Protection Indicators of Resilience and Equity (ASPIRE) database and the International Labour Organization's (ILO's) Social Protection Monitor. This publication focuses on national and international uses.

- How can we best leverage this data to further inform our understanding of poverty, vulnerability, social mobility, and the role and impacts of government social protection policies and programmes in addressing these?

**The primary focus is on implications for low- and middle-income countries.** However, the paper draws on examples from countries at different levels of economic development. Social assistance examples are prioritised, although examples from other pillars of social protection (e.g. social insurance) are also provided.

**This paper discusses analysis that is relevant at both a programmatic level and a social protection system level.** In some cases, the innovative use of administrative data can shed light on questions relating to the design and implementation of a given programme. In others, it can support the analysis of the performance of the social protection system as a whole.

**The analysis in this paper is based on a literature review of relevant academic and grey literature,** including unpublished operational documents from countries in which the authors have worked over the years. It also draws on the authors' broader experience, as well as key informant interviews and email/documentation exchanges with colleagues from a wide range of institutions, coordinated via the Social Protection Inter-Agency Coordination Board (SPIAC-B), to whom we are grateful.

**This paper is organised as follows.** Following a short section that provides the overall framing for the paper, and another in which the benefits and limitations of using administrative data are schematically set out, the paper systematically explores the key options for making better use of social protection administrative data – starting from routine analysis and moving towards more complex uses for evaluation purposes. In doing so, short country case studies are used to bring examples and best practices to life. A final section concludes the paper, drawing out overarching insights, recommendations and preconditions for the better use of administrative data.

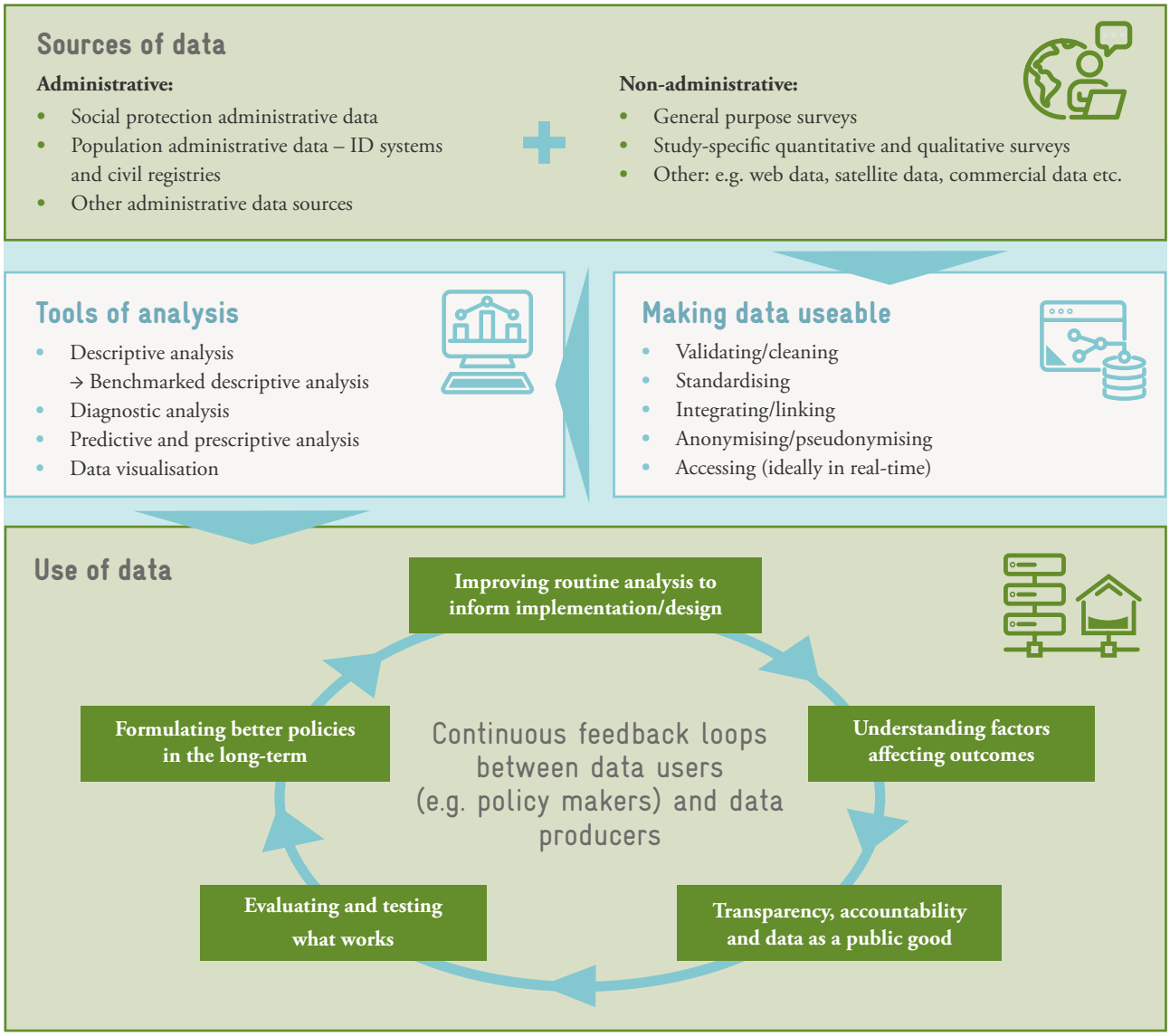
## 1.2 FRAMING

**The framing for this paper is provided in Figure 1,** which visually summarises the following:

- **Sources of data:** The sources of administrative data from the social protection sector include both 'core' data (e.g. demographic and socio-economic data used to determine eligibility) and transactional data emerging from operations (e.g. timestamps). These can be analysed as standalone data or combined with other sources of data (administrative and non-administrative) to further enhance insights.

- **Making data useable:** Depending on the information system managing social protection data in each country, the quality and accessibility of data may vary. Key processes for enabling more strategic analytical use of data include: *validation and cleaning*; *standardisation*, so that variables are clearly understood and align with other data sources; *integration/linking* with other sources (via unique identifiers or other means such as fuzzy matching; see Box 1); *anonymising*/*pseudonymising* to ensure data protection and privacy; and ensuring *accessibility*, ideally via real-time for dashboards and routine monitoring and evaluation (M&E).

**Figure 1** Conceptual framing




**Source:** Adapted by authors based on Eurostat (2018), Symons (2016), OECD (2019)

## BOX 1 What do we mean by 'linking' data?

Linking social protection administrative data to other data sources (such as national survey data or admin istrative databases from other parts of government) involves connecting data in order to gain a more complete and accurate understanding of a particular subject or population. It can happen in different ways, with implications for the ease and depth of analysis. For example, there are large differences between:

- Manual/ad-hoc versus automatic linkage of data, e.g. via interoperability or data integration (e.g. via data warehousing, data federation, or data virtualisation)
- Full 'matching' of records for specific households or individuals, which requires the existence of a unique ID for all individuals, versus less precise methods, such as:
  - 'Fuzzy'/probabilistic/statistical matching
  - Aggregated comparisons: simple comparisons between aggregated data generated from different sources (e.g. sections 2.1 and 2.2)

- **Tools of analysis:** Depending on the sources of data available and the nature of inquiry, a range of analytical techniques can be applied to the data:
  - *Descriptive analysis* uses administrative data to report simply on what has happened. This can be done using very basic statistics, such as (disaggregated) percentages and ratios; simple correlations, showing the relationship between two variables; and cross-tabulated statistics, enabling the comparison of two variables.
    → *Benchmarked descriptive analysis* pushes understanding further, by providing meaningful benchmarks for the descriptive analysis (including from external data sources).
  - *Data visualisation* complements the other tools of analysis as a tool for communicating results from one or more analytical techniques. It can take many forms, such as interactive dashboards, interfaces, graphs, and maps.
  - *Diagnostic analysis* seeks to understand why and how a certain phenomenon has happened (or not).
  - *Predictive analysis* applies mathematical and statistical modelling and techniques (including machine learning) to historical administrative data to predict what might happen in the future, while prescriptive analytics is the process of using data to determine an optimal course of action.

**NOTE,** diagnostic and predictive analysis are not covered in this review.

- **Use of data:** All of the tools of analysis discussed serve the ultimate objective of translating data into information that can serve different purposes. This can be visually represented along a policy cycle continuum, noting that there are significant overlaps between each aspect:
  - Improving the routine analysis of administrative data to inform programme implementation and provide minor fixes to design (as issues emerge)
  - Digging deeper into the 'why' via diagnostic analytics aimed at understanding factors affecting outcomes and patterns raised via routine analysis
  - Striving for transparency and accountability, while ensuring data is used as a public good
  - Leveraging administrative data alongside other sources to evaluate and test what works
  - Gaining insights to support the formulation of policies in the medium to long term.

3 For many of these benefits/advantages to materialise, careful design of administrative forms and underlying data is required.

## 1.3 BENEFITS AND LIMITATIONS OF USING ADMINISTRATIVE DATA FOR RESEARCH AND ANALYSIS

**Compared to other types of data (e.g. statistical survey data), using administrative data for research and analysis purposes has many potential benefits, as well as clear limitations.** These are summarised in Table 1 and need to be kept in mind while engaging with the contents of this paper.

**Table 1** Potential benefits and limitations of administrative data compared to other data sources, especially surveys

| Potential benefits[3] | Potential limitations |
|---|---|
| • Less expensive, because it is available as a 'side-product' of routine operations<br>• Less 'intrusive' and burdensome for target population<br>• Enables the collection of information that is more difficult to collect with a high degree of accuracy in a social survey context (e.g. the exact start and end dates of a job, receipt of a benefit or contribution to a scheme)<br>• Quickly reflects programme/operational changes (compared to survey data, which is usually collected at select intervals)<br>• Allows the tracking of individuals and households over time (longitudinal/panel data)<br>• May be usefully matched via unique identifiers to other data sources (including other admin datasets), expanding the range of analysis<br>• May offer a large 'observation' size allowing for disaggregated analysis and analytical approaches (e.g. quasi-causal methods) which take advantage of variations in the experiences of groups within large samples; may also cover the entire population of interest, with obvious advantages for the study of small subgroups and rare events<br>• Provides information on groups that may be least likely to take part in primary social science research (e.g. individuals from disadvantaged social groups)<br>• Particularly useful for studying issues that individuals might be reticent to disclose to a primary researcher (e.g. mental health problems, substance abuse); receipt of benefits is also often misreported in surveys<br>• Does not trigger 'Hawthorne Effects' (where subjects of an experimental study attempt to change or improve their behaviour simply because it is being evaluated or studied) | • The 'denominator problem', by which the participation-based nature of administrative data limits inferences (data often only includes population groups with certain characteristics)<br>• Data only available when the client is participating in the programme; less is known when they leave or do not enter<br>• Potential for bias due to attempts to game the system (to receive benefits/services)<br>• Risk of low quality data, including errors, omissions and overall 'messiness' of the dataset (especially if the MIS is not adequate) – likely to involve substantial data management<br>• Fragmented data when social protection programmes are managed by different institutions that do not coordinate and share data among themselves[4]<br>• Typically only includes input/output type data; not capable of measuring outcomes<br>• Risks to data privacy and protection |

**Source:** Adapted from Hotz *et al.* (1998), Tonder (2008), Connelly *et al.* (2016), Feeney *et al.* (2015), Cole *et al.* (2020)

4 If not addressed, this may generate duplicate beneficiary information when programmes are aggregated at the national level, due to individuals participating in more than one programme.

# 2

# MAKING BETTER USE OF ADMINISTRATIVE DATA: THE OPTIONS

**This section explores the key options for making better use of administrative data.** It is organised using the framing in Figure 1 and evolves from more simple and routine forms of analysis, which should be first streamlined as best practice, towards more complex and once-off exercises, which can be explored as a country's information systems mature. Many of these uses are touched on in the 1957 International Labour Organization (ILO) Resolution on Social Security Statistics (ILO, 1957) (see Box 2).

### BOX 2 Resolution concerning the development of social security statistics (1957)

As far back as 1957, the Ninth International Conference of Labour Statisticians adopted the *Resolution Concerning the Development of Social Security Statistics*, which stated that: "Each country should encourage the development of a system of social security statistics adequate to serve the following major objectives:

i. to provide basic data for the administrative control of the social security schemes and the appraisal of their operational efficiency

ii. to provide a basis for the appraisal of the financial structure of the schemes, for actuarial valuations and for short- and long-term forecasts

iii. to provide a means of appraising the social security system as an instrument of social policy and in particular to provide a basis for the evaluation of the level of social security protection afforded to the various population groups

iv. to provide general information on social security

v. to supply data for international comparisons in the field of social security

vi. to supply data from administrative and accounting records to satisfy important outside needs in the fields of labour, economic, health, demographic and other statistics"

## 2.1 IMPROVING ROUTINE ANALYSIS TO INFORM PROGRAMME IMPLEMENTATION AND RE-DESIGN

In many low and middle-income countries, despite digitised administrative delivery systems, data within these systems is hardly used, or used only periodically by statistically skilled M&E officers for reporting purposes. In a few others, social protection information systems[5] automatically generate a set of routine reports based on the data that is collected by the system. Most often, however, these have a limited focus on inputs and outputs (e.g. number of beneficiaries and their geographic distribution, average benefit amount).

**This section showcases how routine data analysis can be used for more insightful monitoring and assessment of performance, thus informing improvements in design and delivery in terms of some of the key outcomes that social protection programmes seek to achieve,** such as: (i) coverage, and the inclusiveness of coverage for different sub-populations, (ii) adequacy, (iii) timeliness, (iv) cost and cost-effectiveness.[6]

**Deeper and more insightful analysis of existing data requires analysis against a wide range of:**

- **Disaggregated data and comparisons across different subgroups,** to support relevant interrogation of the administrative data – as and when required: Typically social protection agencies analyse and report outcomes as aggregates, which masks disparities across subgroups based on individual and service characteristics. Disaggregated administrative data, i.e. data that can be broken down and analysed by these characteristics, is the first step to identifying differential barriers and needs, ensuring that the most vulnerable are not left behind. Disaggregation categories that are particularly relevant are summarised in Table 2.

**Table 2** Examples of most relevant disaggregation categories

| Information disaggregated by people's characteristics | Information disaggregated by service characteristics |
|---|---|
| • Age group<br>• Gender<br>• Disability status<br>• Occupation/ employment status<br>• Race, ethnicity or other identity markers<br>• Household income<br>• Household size | • Geographical unit (e.g. district, county, block)<br>• Region served (e.g. urban or rural; prone to recurrent shocks or not)<br>• Office or facility that provided the benefit/service<br>• Individual supervisor or caseworker<br>• Amount of assistance provided, mode of service delivery etc. |

**Source:** Adapted from Barca (unpublished 2014) based on Hatry (1999) and Grosh *et al.* (2008)

5 For example, programme/scheme MISs or integrated beneficiary registries acting as data warehouses.

6 Note: these are just a sub-set of desirable outcomes for achieving universal social protection: e.g. here we are not discussing comprehensiveness (the extent to which different risks are addressed).

- **Relevant benchmarks,** to make the data more meaningful and to understand trends over time, in different geographies, and against relevant targets and denominators (to understand relative, rather than 'absolute', performance)

**Table 3** Examples of most relevant benchmarks

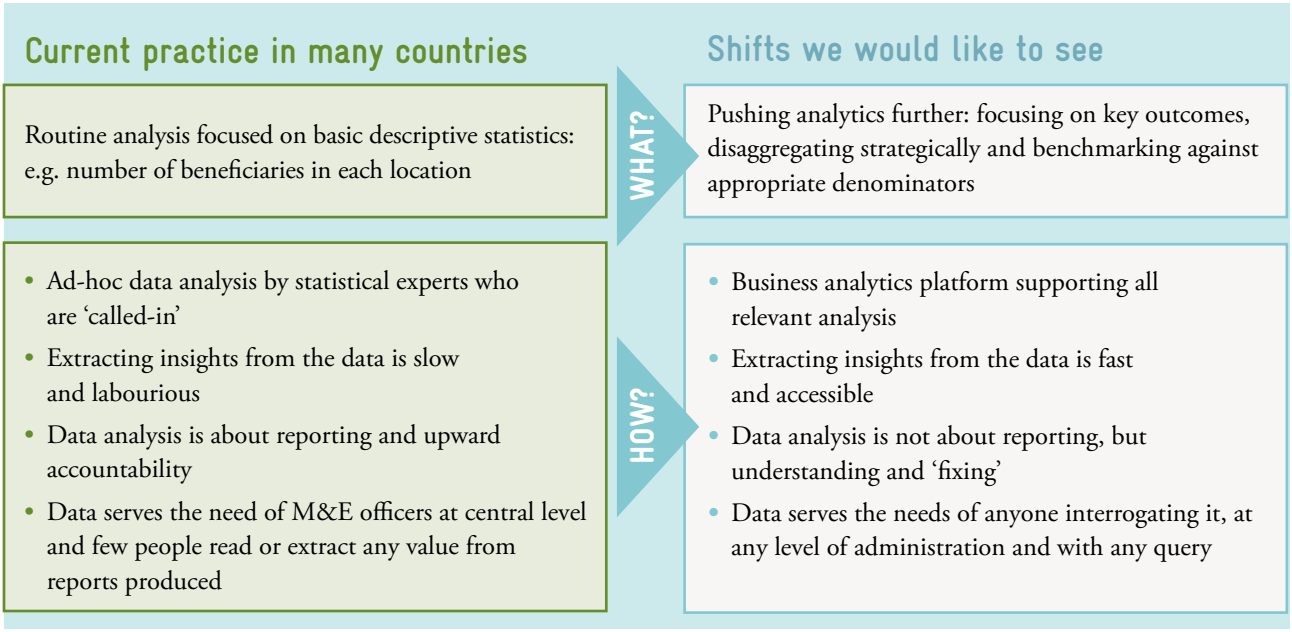
| Information benchmarked against sectoral data | Information benchmarked against relevant external denominators |
|---|---|
| • To previous performance<br>• To ministry/agency targets<br>• Among categories of users<br>• Among geographical areas<br>• Among organisational units<br>• By type and amount of benefit/service or specific vulnerability faced<br>• To the performance of similar programmes in other countries | • Relevant denominators will change depending on output of interest, but include comparable data from census, household income and expenditure surveys and labour force surveys. Also included in ILO (1957) recommendation.[7] |

**Source:** Adapted from Barca (unpublished 2014) based on Hatry (1999) and Grosh *et al.* (2008)

**The standardisation of such analysis can be further enabled via the creation of data analytics/ business intelligence platforms, which are automated to some extent, yet offer flexibility** for users to pursue the analysis they need. How this is done in practice can vary widely and is beyond the scope of this paper, although a great example from Türkiye is discussed in Box 3. These key shifts, compared to current practice in many countries, are summarised visually in Figure 2.

The subsections below further explore how the disaggregation of data and benchmarks discussed above play out with regards to key outcome areas of interest.

**Figure 2** Desired shifts in practice




**Note:** Many of the considerations discussed here, as well as insights deriving from routine analysis, can also feed into longer term policy re-design (discussed in sections below).

7 Social security statistical data "should also include relative measures which show relationships with relevant social, demographic and economic data, thus providing significant informative indications of the progress towards the achievement of most of the major objectives" (ILO, 1957).

### BOX 3 Türkiye's ISAIS

Türkiye's social assistance system, which comprises over 50 programmes, is operated through an advanced and interoperable Integrated Social Assistance Information System (ISAIS), which supports the integration of several functions along the delivery chain. However, despite its strong operational and administrative capacity, ISAIS initially had limited reporting and analysis capabilities. This made the retrieval of basic programmatic data difficult, time-consuming, and disruptive.

To address this, UNICEF supported the Ministry of Family and Social Services to develop a Data Warehouse and Business Intelligence (DW&BI) system and incorporate it into the ISAIS. The BI system transformed administrative data into interactive, real-time dashboards that are easy to access for both technical users and senior managers, and the DW enabled the retrieval of big data through multiple clusters and tables. These dashboards provide real-time descriptive data for all nationwide social assistance programmes with various breakdowns in relation to province, district, gender, age, and number of children. They also offer a forecasting module with time-series analysis to estimate the number of beneficiaries and total budget required for the selected periods.

The government of Türkiye has used the BI system to monitor and expand social assistance programmes, both during routine operations and during the earthquake response in February 2023. For example, for the design, scale-up and monitoring of the biggest social assistance programme, Turkey Family Support Program, the BI system provided comprehensive real-time data, supporting improved decision-making and programming. The system also allowed the government to quickly identify beneficiaries, allocate resources, and roll-out programmes in response to the earthquake. UNICEF is continuing to support the system for a second phase to integrate early warning systems with prescriptive analyses, integrated with various data sources such as weather and population. This will improve the resilience building aspects of the programmes before, during, and after disasters and shocks. Overall, the BI system has improved decision-making, effectiveness and transparency, enabling comprehensive monitoring and a risk-informed approach.

## 2.1.1 COVERAGE AND INCLUSIVENESS

**The first outcome of interest that can be better measured through better data analysis is the coverage and inclusiveness** of social protection interventions, i.e. the extent to which these interventions reach intended categories, especially those who face the highest barriers to access, a key target for achieving universal social protection (**ILO, 2021**).

**Analysis of coverage using disaggregated data is essential to understanding how well social protection programmes serve different sub-populations** (e.g. different age groups, women, people with disabilities, minorities, etc). Explicitly analysing how these different dimensions of disaggregation can combine and reinforce each other is essential to promoting socially and economically inclusive programming (**OECD, 2018**). Disaggregation against service characteristics can provide further insights into trends that may need addressing (e.g. over-coverage in rural versus urban areas).

However, **many of the statistics that can be provided by social protection administrative data alone** (e.g. number of beneficiaries in each disaggregated category), **are of limited analytical value, as they are not measured against policy-relevant denominators** acting as benchmarks.

In other words, disaggregation alone does not provide a *relative* measure of coverage and inclusiveness against desirable outcomes.

**Benchmarking administrative data to information from other national data sources – and embedding this in routine reporting – can often have more analytical value.** This can support analysis in relation to specific programme or policy objectives (e.g. the percentage of intended recipients reached) as well as international comparisons (e.g. coverage rates relative to internationally comparable benchmarks).

**Such an analysis of coverage and inclusiveness involves comparing numerators from administrative data (i.e. number of recipients/contributors) to denominators from relevant survey data** (including census, household income and expenditure surveys and labour force surveys). This strategy is particularly useful when surveys do not include questions on the receipt of relevant social protection benefits or when they misreport the programme receipt (**Dubois & Ludwinek, 2014**).

**Ideally, these denominators would be incorporated into data analysis platforms to automate such analysis.** This automation is more straightforward in places where relevant national surveys (such as income and expenditure and labour force surveys) are conducted routinely (e.g. annually). In many countries, such surveys are conducted infrequently. While population censuses are usually conducted every 10 years, the production of population projections and vital statistics can provide data for the intervening years.

**A wide variety of denominators are available for benchmarking social protection coverage.** Examples of these denominators are listed in Table 4, including some key considerations. Box 4 provides country examples of the use of denominators. Ideally, analysis should involve comparing multiple levels of disaggregation to relevant benchmarks. For example, the number of female pension recipients in rural areas could be compared to the rural population of older women according to the census data.

It is worth noting that benchmarking against denominators from other data is easier for some types of schemes than others. For example:

- This approach is relatively straightforward for schemes with easily verifiable eligibility criteria (such as for universal old age or child benefits), as one can be fairly confident that beneficiaries fall within the group indicated by the denominator.
- In the case of poverty-targeted schemes, denominators have to be specially developed based on the analysis of survey data. Exclusion errors associated with poverty-targeted schemes mean that the benchmarking of the number of recipients against the target population should be interpreted with some caution.[8]
- Given the significant challenges associated with measuring disability prevalence, care is required when comparing the recipients of disability benefits with the population of persons with disabilities (as per census or other surveys). When making this comparison, it is also important to consider whether or not the denominator reflects the population of persons with disabilities actually eligible for the scheme (e.g. a scheme may focus on severe disability, or particular types of disability). In this context, and considering the substantial variation in the measurement of prevalence across countries, comparing receipt of disability benefits to the total population is sometimes considered a more useful benchmark for cross-country comparison (**UNPRPD, 2021**).

8 For example, they often require the development of special denominators and percentage of beneficiaries based on microsimulations that rely on household survey data.

**Table 4** Examples of denominators for benchmarking coverage and inclusiveness of social protection schemes

| Type of denominator | Example of indicators | | Relevant data source for denominator | Considerations |
|---|---|---|---|---|
| | Numerator | Denominator | | |
| Households | Number of households receiving a cash transfer | Total number of households | Household survey; population census | Limited insight on coverage of intended population |
| | Number of households receiving a cash transfer | Number of households in target group (e.g. below poverty line) | | Does not shed light on targeting accuracy |
| Age eligible population | Number of children receiving a child benefit | Number of children aged 0–17 years | Population census and projections | More complex in places where age of eligibility may vary (e.g. early or late retirement age) |
| | Number of individuals receiving an old age pension | Number of older persons aged 65 and over | | |
| | Number of mothers receiving maternity benefits | Number of children born per year | | |
| Occupation and employment status | Number of workers actively contributing to a social insurance scheme | Working age population (e.g. aged 15–64) | Population census and projections | More meaningful in places where data is disaggregated by age and sex (given lower female labour force participation and variations by age group) |
| | | Labour force | Labour force survey, household survey | |
| | | Number of workers legally covered by schemes (e.g. employees in enterprises with more than 10 employees) | | |
| Persons with disabilities | Number of individuals receiving a disability benefit | Number of people with disabilities | Household survey; population census; demographic and health survey (DHS) | Measures of disability prevalence are inadequate in many countries. |
| | | Total population | Population census and projections | Prevalence measures may not reflect scheme eligibility. |
| Gender | Number of women receiving old age pension | Number of women aged 65 and over | Population census and projections | |
| Location | Number of households in rural areas receiving a cash transfer | Number of households in rural areas | Population census and projections | |
| Race, ethnicity and other identity markers | Number of children of ethnic minorities receiving a child benefit | Number of children of ethnic minorities | | |

**Source:** Compiled by authors; also see Methodological Annex in *World Social Protection Report* (ILO, 2021)

## BOX 4 Strategic uses of data disaggregation and benchmarks to inform an understanding of coverage and inclusiveness: country examples

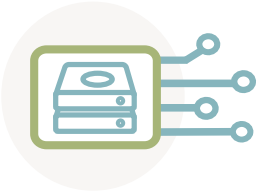
- In **Brazil**, administrative data from the social registry Cadastro Único and the conditional cash transfer (CCT) programme Bolsa Família is combined with data from the Pesquisa Nacional de Amostra de Domicílios (PNAD) National Household Survey to assess under-coverage of people from the lowest-income groups. This forms the basis for expanding registration through an active search strategy that proactively reaches out to under-served populations through door-to-door or community-level registration (**World Bank, 2019a**).
- In **Moldova**, data from the MSAS database that supported a new guaranteed minimum income programme was linked with the nationally representative Household Budget Survey to assess programmatic trends vis-a-vis the broader population. This systematic assessment helped to recognise and address over-coverage and under-coverage of specific categories: e.g. very low coverage of Roma populations and 1–2 person households with no children, and very high coverage of large families with children, compared to national statistics. The reasons for these disparities were further explored in qualitative research and the issue redressed via ad-hoc policy measures (**Barca & Carraro, 2013**).
- In the **Philippines**, the CCT Pantawid Pamilyang Pilipino Program (4Ps) specifically tracks coverage of indigenous peoples in geographically isolated and disadvantaged areas, families in need of special protection, and homeless street families. It is also benchmarked against targets that the Philippines has committed to in various national strategies (**World Bank, 2019b**).
- In **Uruguay**, the Integrated Information System of the Social Area (SIIAS) consolidates data across 57 programmes from 16 public institutions. Innovative disaggregation categories have been added over time to analyse the inclusion of women and transgender people (**UNDP, 2016**).
- In **South Africa**, age-disaggregated analysis of coverage benchmarked over time helped identify a continued decline in the take up of the Child Support Grant (especially for children 0–1 years) so that remedial measures could be taken accordingly (**SASSA, 2021**).
- In **Jamaica**, the Programme of Advancement Through Health and Education (PATH) disaggregated data for its CCT by gender, which enabled the introduction of a tiered and gender differentiated payment system for school children, in response to findings that the dropout rate among older boys was higher than among girls from the same cohort (**Paes-Sousa, Regalia, & Stampini, 2013**).
- The **ILO**'s World Social Protection Database (which underpins the analysis in its flagship World Social Protection Report) includes a number of indicators comparing administrative data to a variety of denominators, primarily for the purpose of cross-country comparison (**ILO, 2021**).

## 2.1.2 ADEQUACY

**The second outcome of interest is the adequacy of social protection benefits,** especially where these are cash-based. This is another key dimension of universal social protection, as it guarantees benefit levels that ensure effective access to essential goods and services, and secure protection against poverty, vulnerability and social exclusion, enabling beneficiaries to live with health and dignity, as per **ILO Recommendation 202 (ILO, 2012)**.

**Comparing benefit levels to benchmarks based on survey and economic data can provide an indication of benefit adequacy.** The nature of such analysis is influenced by the way in which the benefit level is formulated, which falls into two main categories:

- **Flat rate benefits**, where the benefit is the same for all recipients: This single benefit level makes analysis relatively simple, as no disaggregation is required. However, it may be debated whether or not information on such benefits should be considered a form of administrative data, given that there are pre-determined scheme parameters (rather than deriving from programme operation).
- **Benefits that vary between recipients:** Notable examples include (i) benefits that vary according to factors including poverty levels, household composition, level of disability and age of recipients/household members and (ii) contributory schemes, where benefit levels are calculated based on the level and duration of previous contributions. In these cases, it is important that data on benefit levels is disaggregated, for example, by reporting:
  - Average benefit paid
  - Highest and lowest benefit paid
  - Distribution of beneficiaries by the size of benefit (e.g. number of beneficiaries within a given range of benefit levels); see example from the Philippines in Box 7



**The disaggregation of data on benefit levels is particularly important when benefit levels vary between recipients.** Nevertheless, disaggregated analysis is also possible when benefits are flat rate, for example, by comparing the benefit level to denominators disaggregated for different groups (such as the value of a child benefit relative to average earnings of the population in the poorest quintile).

**There are a wide variety of denominators which can be used to benchmark benefit adequacy.** Some important distinctions include:

- The extent to which they are focused on **absolute levels of adequacy** (e.g. national poverty lines based on a minimum food basket) or on the **relationship to average incomes or wages** (e.g. wages, median income, previous earnings)
- The extent to which they are **normative** (e.g. national poverty lines, minimum wages[9]) or **contextual/descriptive** (e.g. median income, GDP per capita, wages, previous earnings)

The appropriateness of a given denominator will be influenced by the objectives of a given scheme (e.g. poverty reduction versus income replacement), and also the purpose of analysis (e.g. measurement against national benchmarks versus international comparison).

9 Note, even where a denominator is normative (i.e. a poverty line representing a minimum level of income/consumption that all people should be able to achieve), its formulation will often be undertaken with reference to the analysis of survey data.

**Table 5** Examples of denominators for benchmarking adequacy of social protection schemes

| Denominator | Source | What does it indicate? | Considerations |
|---|---|---|---|
| National poverty line | Household surveys | In theory, the national poverty line can be conceived as a nationally agreed definition of what is considered a minimum level of income (or consumption). There is significant variation in how countries define their poverty lines. | In practice, definition of national poverty lines can be arbitrary or influenced by political processes. |
| | | | Variations between methodologies for determining national poverty lines mean they are of limited value for cross-country comparison. |
| | | | Reliance on household survey data means that poverty lines may only be readily available infrequently and retrospectively. |
| International poverty line | Internationally defined, and calculated using PPP conversion (exchange) rates | International poverty lines are presented in PPP dollars and designed to provide a consistent measure of poverty for cross-country comparison. | There is significant debate about the elaboration (and level) of these poverty lines and the usefulness of such an undertaking. |
| | | | Different lines exist for different country income groups. The 'jump' in the poverty lines between these groups can lead to significant differences in the benchmark used for countries with similar levels of economic development. |
| GDP/GNI per capita | National accounts, combined with population data | GDP/GNI indicate the level of economic development of a country. One way of conceiving the use of this denominator is that it indicates adequacy relative to the redistributive potential of a country. | Data on GDP per capita is readily available from national accounts (and projections). |
| | | GDP/GNI per capita is also an indicator of average income, but is less useful than other indicators (e.g. median income) as a measure of average incomes within the population/labour market. | The use of GDP per capita can be problematic in countries with large sources of foreign income (e.g. various Pacific Islands). GNI per capita is arguably a better indicator, but is less routinely reported and readily available. |

| | | | |
|---|---|---|---|
| Survey-based indicators of average income/consumption (e.g. median income, average income of target group) | Household surveys | Household surveys indicate adequacy relative to the prevailing level of income/ consumption within the population (or a specific part of the population of interest). | Median income is not often routinely reported in household survey data analysis, and its availability at the international level is quite limited.[10] |
| | | Median income provides an alternative to measures of mean income, which can be distorted by high levels of inequality. | Indicators such as average consumption of target group require dedicated analysis of household surveys. |
| Wages | Labour force surveys; household surveys | Wages indicate the adequacy of benefits relative to earnings within the labour market. Where appropriate, this may be narrowed down to relevant sectors in the labour market. | Most useful for benefits conceived as a form of income replacement. |
| | | | Labour force and household surveys only collect data on average wages for employees in enterprises (so this denominator does not include the incomes of self-employed or contributing family workers, which constitute a large share of employment in many LMICs). |
| Minimum wage | Government laws and regulations | The minimum wage indicates adequacy relative to the normative benchmark of a minimum wage. | Requires understanding of the rationale for determination of the minimum wage, especially when making international comparisons. |
| Previous earnings | Social security data | Provides an indication of the replacement rate of a benefit provided. Previous earnings can be measured as final salary, an average of a number of years and, in some cases, average wages within the labour market. | Mostly used to measure the adequacy of contributory benefits. |

**Note:** PPP = purchasing power parity; GDP = gross national product; GNI = gross national income; LMIC = low- and middle-income country

**Source:** Compiled by authors; also see Methodological Annex in *World Social Protection Report* (ILO, 2021)

**Defining what might constitute an adequate benefit requires careful consideration and is often politically contested.** Key considerations include:

- What is the function of the benefit? For example, the benefit might be a form of income replacement due to lost earnings from work, compensating for a defined set of costs, or a supplement to the income of poor households.
- How does the benefit interact with other benefits, services and sources of support? For example, it is important to determine whether or not recipients receive other benefits, the extent of out-of-pocket expenses for accessing health care, and whether or not benefits are assumed to be supplemented by income from work or financial support from families.

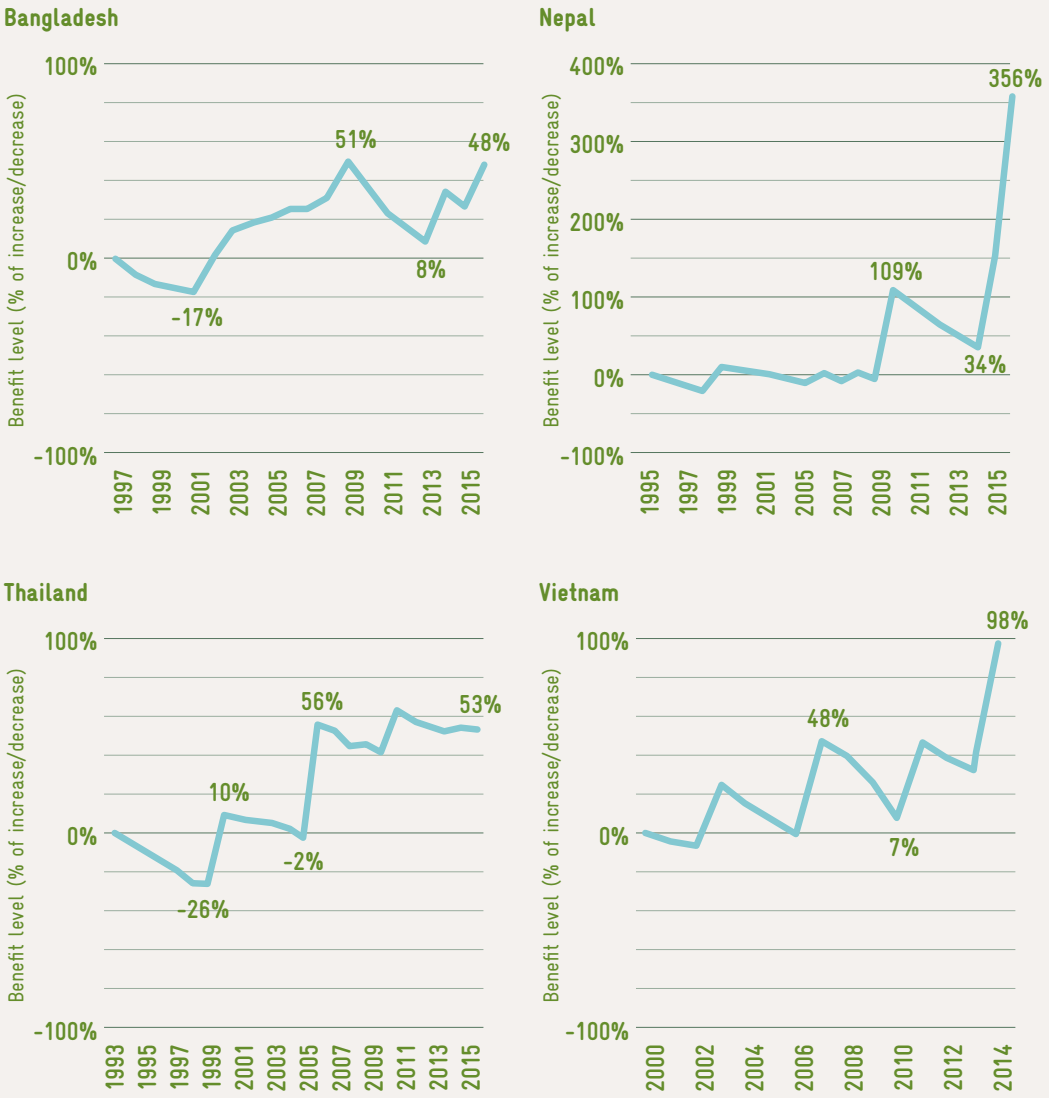
**Another important dimension to the analysis of benefit adequacy is changes over time.** Benefit levels should ideally be assessed relative to changes in prices (inflation) to determine changes in their real value (and, thus, inform the need for adjustments), even though this is rarely done. They can also be assessed relative to changes in wages or average incomes.

10 The World Bank's Poverty Calculator includes data on median income in PPP$ from most recent household surveys (https://pip.worldbank.org/poverty-calculator).

## BOX 5 Strategic uses of data disaggregation and benchmarks to inform an understanding of adequacy

Routine reporting of benefit adequacy is most common for contributory schemes in which benefit **replacement rates** are calculated by an individual's previous earnings and/or average wages in the labour market.

There are few examples of social protection implementing agencies routinely benchmarking the adequacy of benefits (beyond cross-country analytical work and databases, e.g. the World Bank's ASPIRE). Where this analysis exists, it is most commonly undertaken by independent researchers/academia, international organisations, and civil society organisations. As one example, a study of old age income security in four countries in Asia and the Pacific analysed the change in the real value of social pension benefits over time. The study found that, despite a lack of formal indexation, real values had increased since the schemes were introduced as a result of ad hoc benefit increases (**HelpAge International, 2017**):




**GDP per capita** is commonly used to benchmark the adequacy of social protection benefits across countries. This has been applied to social pensions (**HelpAge International, 2018**; **Palacios & Sluchynsky, 2006**; **Tran, 2021**), disability benefits (**UNPRPD, 2021**; **Development Pathways, 2019**) and child support (**McClanahan & Gelders, 2019**).

### 2.1.3 TIMELINESS AND PREDICTABILITY

**Measuring timeliness against set delivery targets, and the predictability of application processes and benefit receipt, are also key indicators of programme efficiency and effectiveness** (**Lindert *et al.*, 2020**), as well as a right established under ILO Recommendation 202. These can be relatively easily programmed into the information systems serving social protection programmes by leveraging digital time-stamps. Nevertheless, these indicators are seldom reported – or systematically analysed to understand reasons for delays and improve performance.

**Specifically, transactional data such as time-stamps can help to automatically capture and report on one or more of the following metrics,** including differences across relevant disaggregated data and benchmarks (e.g. compared to previous performance, agency targets, across categories of users, or among geographical areas or organisational units):

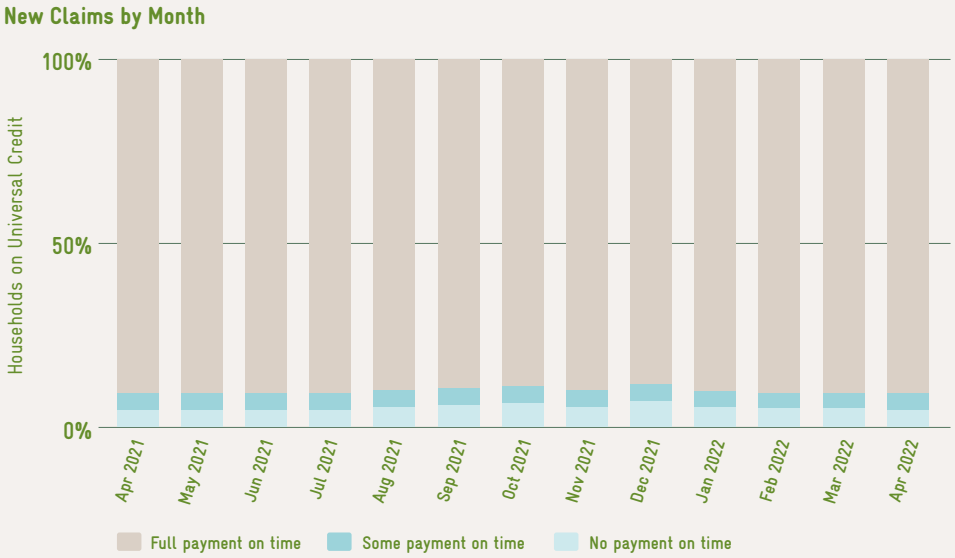
- Number of working days from the:
  - date of application to date of enrolment notification
  - date of enrolment to payment, in case of once-off programmes
  - start of the month to payment, in case of recurrent payments
  - date of update request to beneficiary information updating
  - date of grievance reporting to date of redressal
  - date of a certain shock/trigger and payment
- Percentage of applicants who are notified of eligibility status or enrolment decisions according to service standards (such as < 30 days)
- Percentage of beneficiaries who predictably receive payments within agreed service standards (such as < 7 days from agreed date)
- Percentage of beneficiaries with outdated data according to agreed standards (such as updated > 2 years ago)

**The analysis of disaggregated data may also help to uncover differences across subgroups and/or organisational units.** For example, certain decentralised offices may be struggling more than others to maintain delivery schedules and may require ad-hoc support.

#### BOX 6 Strategic uses of data disaggregation and benchmarks to inform an understanding of timeliness: country examples

- In the **United States**, the Social Security Administration's (SSA's) **Time Reference for Management Information** application is responsible for maintaining a series of tables on temporal information, such as case processing time and survival periods, for social security insurance and disability insurance recipients. The data is also published regularly, for instance, the number of **days from application to first payment** for the Social Security Disability Insurance. The SSA introduced the quick determination of disability process to reduce processing time for the severely disabled and this is tracked separately (**Koudou & Dely, 2022**).

- In the **United Kingdom**, the Department of Work and Pensions publishes the timeliness of payments of the Universal Credit (a benefit payment for persons of working age) every month. The metric is published separately for all claims and new claims.




Source: https://stat-xplore.dwp.gov.uk/webapi/jsf/dataCatalogueExplorer.xhtml

- In **Costa Rica**, administrative data from the Family Intervention Plans programme is used to monitor time taken for approval of referrals by service providers. Between January 2015 and June 2016, 56% of referrals took more than the standard of 30 days (**Lindert *et al.*, 2020**).
- In **Argentina**, the Federal Social Insurance Council has monitored the time taken to process pension applications since 2008. The average processing time has declined drastically from 538 days in 2008 to 15 days in 2022.
- In **Moldova**, the MSAS database identified areas where applications were either too few or where there were delays in processing applications. In the first case, the problem was eased by the substitution of absent social assistants with other personnel in the town hall. In the second case, extra equipment and personnel were deployed to reduce the backlog (**Barca & Carraro, 2013**)
- In **Georgia** legislation sets clear deadlines for the registration of households in the Database of Socially Vulnerable Households, as well as the determination of eligibility for and payment of Targeted Social Assistance and Child Benefits. These are actively monitored by the administrative system and local offices regularly receive warnings when deadlines are approaching.

### 2.1.4 BETTER MANAGEMENT OF PEOPLE AND RESOURCES, WHILE RESPONDING TO USER NEEDS AND PREFERENCES

**More broadly (e.g. beyond timeliness alone), social protection administrative data can be used to support the more efficient and effective management of people and resources,** similarly to private sector data-driven performance management. This includes incorporating a stronger focus

on user needs and preferences. Tailored analysis of data can support many functions including:

- Allocation of human resources (e.g. where demand is higher and/or service delivery is lagging)
- Planning and budgeting
- Measuring performance via key performance indicators (KPIs) and possibly building performance incentives around these
- Better serving people by responding to their needs and preferences
- Picking up on possible mistakes and/or fraud

### BOX 7 Using administrative data for better management of people and resources

- In the **Philippines**, the 4Ps conditional cash transfer tracks the mode of payment (cash cards versus over-the-counter payments). Based on this data, the government has progressively expanded the network of cash-out points available to beneficiaries (Lindert *et al.*, 2020)
- In **Morocco**, the Caisse de Dépôt et de Gestion (CDG) Prévoyance, the country's social insurance agency, is disaggregated by service channel (digital versus physical) to help the agency identify and remedy low levels of digital adoption by certain categories of users.
- In **Moldova**, during the roll-out of the new Guaranteed Minimum Income Benefit, the comparative performance of social assistants across the country was measured using administrative data. Key data points included the overall number of households each was able to enrol and the timeliness of the enrolment process, among other things. Low-performing assistants were supported with additional training and interviews to understand what was driving the backlog.
- In **Brazil**, finances for decentralised government units responsible for Cadastro Unico tasks are determined based on a Decentralised Management Index. This index is calculated on the basis of a series of datapoints, many of which can be derived directly from administrative datasets, e.g. the rate of actualisation of registry data, the rate of compliance with conditionalities, and the usage of the Unified Social Assistance System (SUAS).

## 2.1.5 COSTS AND COST-EFFECTIVENESS

**Data on social protection expenditure can contribute to improving implementation and informing medium-term policy development in various ways.** Three important areas include:

1. **Implementation:** Data on expenditure can potentially inform more scheme-specific analysis of questions such as the cost-efficiency of schemes, and whether budgets are being effectively executed.
2. **Policy development:** Disaggregated data on the scale and shape of social protection expenditure can inform policy discussions on the appropriateness of the existing distribution of expenditure. For example, how does the level of social

protection expenditure compare to other sectors, and does the distribution of investment between different schemes and population groups reflect perceived needs? This data can also contribute to analysis of the cost-effectiveness of the social protection system (or specific schemes). For example, levels of investment in social protection can be compared to the impact on indicators such as poverty and inequality.

3. **Assessing sustainability:** Data on expenditure may also contribute to analysis of the financial sustainability of the social protection system. However, data on current and historical expenditure usually need to be accompanied by projections of future costs to be meaningful (for example, through actuarial analysis or social budgeting exercises).

**From a practical perspective, there are two major challenges to collecting and interpreting social protection expenditure data:**

- **Such data may be drawn from different ministries and levels of government.** First, data on budgeted allocations and actual expenditure is collated at the level of ministries of finance linked to the elaboration of (and reporting of) national budgets. Second, implementing agencies may generate more detailed expenditure data on their operations, although the extent to which this is systematically monitored or reported also varies.

- **The ways in which this data is classified and reported (including disaggregation and the nature of classification) varies significantly,** both across institutions and from country to country. Ideally, what is required is a clear definition of what is – and what is not – social protection expenditure, and the key different types of social protection expenditure. However, being a relatively new sector in most low- and middle-income countries, it is uncommon to find clearly elaborated national-level statistical classifications for social protection.[11]

**In other words, data on expenditure requires significant efforts to make data more useable and useful,** as per this paper's framing (Figure 1), both in terms of standardisation and integration/linking across data sources. Resolving this issue requires standardised classifications such as the International Monetary Fund's (IMF's) Government Finance Statistics Manual (GFSM) in reporting government expenditure (explained briefly in Annex 1).

**In principle, expenditure data can be used to assess cost-effectiveness and cost-efficiency, however, in practice this is challenging.**

**In relation to *cost-effectiveness*, White *et al.* (2015) describe a variety of indicators that can be used to measure cost-effectiveness relative to outcomes including reductions in poverty and inequality.** These usually involve the comparison of levels of expenditure on social protection programmes or systems, with associated reductions in poverty or inequality measured by household surveys. In practice, however, the use of such measures has been relatively limited in social protection policy.[12]

***Cost-efficiency* is understood to relate to "ensuring that planned outputs of the right quality are produced at the lowest possible cost"** (White *et al.*, 2015, p. 11). This is an important topic of

11 International organisations also tend to vary in how they classify social protection data. Two important initiatives include the Asian Development Bank's (ADB's) Social Protection Indicator, and the World Bank's ASPIRE database, however, while these provide an important lens on social protection expenditure, they have some limitations. For example, they are not part of a comprehensive and consistent overall system of classification of expenditure. This means there is more risk of expenditure being double or triple counted across different sectors (e.g. health and education) when combined with other data sources. Data reported by ILO as part of its social protection databases somewhat suffers from this issue, by aggregating expenditure data with different definitions (ILO, 2022).

12 This is largely due to issues of comparability between different schemes and systems which may have different objectives. For example, social assistance and social insurance schemes vary in the extent to which reduction of poverty and inequality is a primary objective (compared to, say, providing income replacement). Such indicators are likely to be more useful in comparing the cost-effectiveness of different programmes in a given country. Additional issues emerge in cross-country comparison, not least in terms of how poverty lines are defined (White *et al.*, 2015).

discussion in social protection policy, often orientated around the question of the scale of administration costs. Two notable indicators used for cost-efficiency include:

- **Cost-transfer ratio (CTR).** The share of administrative costs in the total programme budget (administrative costs/total programme costs) or the administrative cost of making a one-unit transfer to a beneficiary.
- **Total cost-transfer ratio (TCTR).** The total cost, including transfers, of delivering a one-unit transfer to a beneficiary (total programme cost/total value of transfers) (**Lindert *et al.*, 2020**).

**These indicators have been used in the analysis of cost-efficiency for a range of different social transfer schemes in low- and middle-income countries** (**Lindert *et al.*, 2020**; **White *et al.*, 2015**). Creating such indicators relies on relatively granular data on different kinds of expenditure, which are likely to only be available at the level of the implementing agency for a given social protection programme.[13]

**Nevertheless, despite the potential for measuring cost-efficiency, there are significant challenges to doing this in practice.** A major challenge involves defining and tracking what constitute administration costs for a given programme. Notable issues include how to treat costs that are related to specific periods of programme implementation (such as set-up costs) and how to measure administrative expenditure that relates to multiple schemes (such as the time a social worker may spend on undertaking targeting for a cash transfer scheme). Challenges also relate to the interpretation of cost-efficiency indicators. An important issue is that cost-efficiency indicators show little about the quality of the benefit or service delivered. Administrative costs may be high in one programme because a higher quality of service is being delivered, meanwhile, they may be low in another because beneficiaries need to take on significant private costs to access a benefit (**Lindert *et al.*, 2020**).

**These issues suggest that administrative data on cost-efficiency should be treated with some care.** A pragmatic approach involves routinely collecting data on different kinds of costs incurred by a given programme that can be included within a broader dashboard of monitoring indicators. Nevertheless, care should be taken not to over emphasise the meaning of such indicators, which may well be more useful for tracking changes over time, than providing absolute benchmarks against other programmes. These indicators should also be complemented by other forms of analysis, such as specific research or field monitoring, and with investigating issues such as the private costs of programme beneficiaries.

## 2.2 BETTER UNDERSTANDING THE 'WHY': INVESTIGATING FACTORS AFFECTING OUTCOMES

Routine 'descriptive' analysis of administrative data can go a long way to understanding the 'what' of delivery, illustrating key trends and patterns, including against key outcome areas, as discussed in section 2.1 above. **This section goes one step further, briefly discussing the potential for more diagnostic analysis,** seeking to understand the 'why' and 'how', by investigating the factors affecting outcomes and gaining insights into the causes of patterns observed in the data.

13 It should be noted that classifications of expenditure at the total government level (such as those described in the IMF GFSM) do not include adequate data on levels of administrative cost.

**In many cases, administrative data can be the starting point for such analysis, complemented by ad-hoc data collection, both qualitative and quantitative.** The key difference with routine uses of data is that diagnostic analysis cannot be automated within a business intelligence platform: it requires skilled staff to interpret the patterns observed in the descriptive statistics and interrogate these further via additional methods, which may include:

- **Data drilling:** Drilling down into a dataset to reveal more detailed information about which aspects of the data are driving the observed trends, picking up anomalies, outliers.
- **Data mining:** Hunting through large volumes of data to find patterns and associations within the data. This can be conducted manually or via machine learning technology.
- **Correlation analysis:** Examining how strongly different variables are linked to each other.
- **Design of ad-hoc complementary research, both qualitative and quantitative:** Light-touch and low cost 'targeted' research to answer very specific questions, usually with highly purposive sampling drawn from the administrative data itself.

**The examples in Box 8 speak to the potential for such diagnostic analysis to be further incorporated into the toolkit of social protection practitioners.** In practice, this could be achieved via skilled in-house expertise or strategic partnerships with national bureaus of statistics or universities.

## BOX 8 Diagnostic analysis, interrogating the 'why' of patterns observed in administrative data: country examples

- In **Mexico**, administrative data from the conditional cash transfer programme Oportunidades was used to determine who was likely to drop out of the programme. The analysis indicated that wealthier beneficiaries had a greater likelihood of dropping out, suggesting that the conditionality indeed acted as a screening device (**Álvarez *et al.*, 2008**).
- In **Moldova**, the comparison of administrative data and survey data revealed a severe problem with the uptake and retention of a recently introduced Guaranteed Minimum Income benefit (Ajutorul Social).[14] Diagnostic analysis was used to compare the characteristics of beneficiary households and eligible households, using the benefit's administrative database, MSAS. To further interrogate initial insights, an additional question about the reason for non-application was added to the Household Budget Survey (HBS), revealing severe awareness issues.[15] Data from the MSAS and HBS were subsequently used to purposely sample households for qualitative assessments, to dig further into reasons for low uptake and retention. Insights from the research led to the development of a tailored and multi-pronged communication campaign. The results were immediately visible in the MSAS administrative data: the number of beneficiaries swiftly increased from 20,000 to over 37,000 and the retention problem was almost entirely resolved (**Barca & Carraro, 2013**).

14 Of the 71,000 households estimated as eligible, only some 21,000 were receiving the benefit in December 2009. Moreover, 10,500 households had stopped receiving the benefit because they did not re-apply.

15 In Quarter 1 of 2009 HBS data revealed 47.3% of eligible households were not aware of the benefit and almost 39% believed they were not eligible.

- In **Switzerland**, research into the marginalisation of persons with disabilities in applications for a disability insurance benefit (DB) used administrative data on applications, linked to a social protection and labour market cross-sectional survey (SESAM). The analysis looked at how long-term health-related activity limitation, region of residence, and demographic and socioeconomic characteristics affected the probability of applying for the benefit. Programme data allowed the measurement of actual outcomes, in this case DB applications, which are not available in surveys (**Altwicker-Hámori, 2021**).

## 2.3 TRANSPARENCY, ACCOUNTABILITY AND ADMINISTRATIVE DATA AS A PUBLIC GOOD

**Administrative data can serve internal analytical purposes, but can also be shared externally to enhance transparency and accountability, while enabling the use of the data** (e.g. by academic researchers) **as a public good. The 'Open Data' movement, which has been gaining traction globally, also applies to the 'opening' of social protection administrative data** as an important mechanism for transparency, external accountability and civic participation (**Rohwerder, 2016**; **Baldeon & Arribas-Baños, 2018**; **ISSA, 2019**). Providing public access to administrative data provides citizens and civil society organisations with opportunities to hold duty bearers accountable. It also provides an opportunity for research institutions such as universities to feed deeper/richer research and insights to time-strapped government policy and decision makers. Some further argue that "by encouraging the use, reuse and free distribution of datasets, governments promote business creation and innovative, citizen-centric services" by fostering "higher collaboration within governments, as well as between government agencies and the wider society" (**OECD, 2013**).

**'Opening' social protection administrative data can be done in two main ways,** both of which require specific efforts to protect the privacy of those whose data is being shared:

- **Public-facing dashboards,** sharing aggregated data and enabling varying levels of investigation (depending on how these have been set up). Examples abound in the sector, but most are programme specific, rather than offering an overview across programmes.
- **Sharing of:**
  - **'Restricted-use' microdata files,** which can be only be accessed under very specific conditions (e.g. only by research institutions, within research data centres[16] or via restricted and secure remote access, upon submission of a specific data access request and research proposal)
  - **'Public use' data files**, that have been processed, prepared (e.g. anonymised) and disseminated for research purposes

16 See for example the US Federal Statistical Research Data Centers, which are partnerships between US federal government statistical agencies and leading research institutions to provide secure facilities located throughout the United States that provide access to restricted-use microdata for statistical purposes to authorized individuals.

**Table 6** Public-facing dashboards and sharing of data files, key considerations and examples

| | Key considerations | Some examples |
|---|---|---|
| Public-facing dashboards | Cannot be truly classified as 'open data', as they typically provide an aggregated view and do not enable further analysis beyond what the dashboard has been set up for. However, such dashboards can enable simple and visual interrogation by non-experts, making them a useful tool for accountability, if well-designed. | • **India MGNREGA**<br>• **Philippines 4Ps** and **Social Amelioration Program**<br>• **Bangladesh social protection budget portal**<br>• **Indonesia PKH**<br>• **Kenya HSNP**<br>• **Cambodia ID-poor registry**<br>• **Chile digital analysis of social information** |
| Sharing of: 'restricted-use' microdata files and 'public use' data files | Where these are shared according to the Open Data Principles (Box 9), they can be an extremely valuable resource. | • **Ireland**<br>• **Uruguay**<br>• **Moldova social assistance and social insurance**<br>• **USA** |

**Of course, 'opening' social protection data should not be an end in itself, but a means to an end, which is better and more strategic use of data to inform strategic policy and operational decision-making over time,** ultimately better serving those who the social protection sector is meant to serve. In order for this to be achieved, it is not sufficient to just 'share' the data, but to do this in a way that makes it as useful as possible to potential users. Open data campaigners have been making very specific requests for this to be achieved, as summarised in Box 9. Box 10 contextualises this by providing some country examples of experiences to date sharing social protection data with the public.

## BOX 9 'Opening data' is not enough: principles for making open data useful

'Open Data' is only useful to potential users when it is:

- **Published in electronic formats that are machine readable and non-proprietary** (open licence): So that anyone can access and use the data using common, freely available software tools (e.g. this means using formats such as XML and XSLT).
- **Clearly documented:** Includes documentation about the format and meaning of data, including standardised tagging and meta-data.
- **Complete:** All public data (that is not subject to valid privacy, security or privilege limitations) are made available.
- **Primary:** Collected at the source, with the highest possible level of granularity, not in aggregate or modified forms.
- **Timely:** Made available as quickly as necessary to preserve the value of the data.
- **Re-useable:** Including terms that permit reuse and redistribution, including intermixing with other datasets.

- **Designed with public input:** Clearly driven by the public, who want and use the data, including what data are released when and in what form. Open data policies that release as much data as possible and leave the rest to market dynamics are inadequate. A stronger approach curates data to ameliorate specific problems of people and communities, ensuring that data can be used in line with their needs, capacities and vulnerabilities.

**Source:** Consolidated from the 8 Principles of Open Government Data, Sunlight Foundation's Open Data Policy Guidelines, World Bank Open Government Data Toolkit and FAIR Data principles. Note that these sources include many other relevant principles and insights.

## BOX 10 'Opening' social protection data to the public: country examples

- In **India**, the National Rural Employment Guarantee Scheme (NREGS), the country's legally backed public works programme, has published transaction-level data **online** since 2010. The Act underpinning the programme stipulates that all information related to the programme be proactively published in the public domain in order to uphold the transparency and accountability measures enshrined in the Act (**Hanbal & Prakash, 2019**). However, transparency does not automatically lead to accountability, especially when low literacy target populations have to contend with dense layers of administrative data (**Van Zyl, 2013**). Civil society organisations play an important role in bridging this gap. For instance, civil society organisations in India have developed trackers to continuously monitor and shed light on the payment delays associated with NREGS. An analysis of nearly 1.8 million transactions between April 2021 and September 2021 showed that 71% of payments were not made within the mandated 7 days (**LibTech India, 2021**).

- As of 2012, **Brazil**'s Secretariat for Evaluation and Information Management (SAGI) created a Cadastro Unico data visualiser, called the **Information Consultation, Selection and Extraction Tool (CECAD)**. The creation of the tool was an iterative and participatory process, focused on the needs of different intended users. Interestingly, in the course of the tool-building process, the Cadastro Unico managers were also able to identify previously undetected problems resulting from unsystematic use of the database. Since its launch, the CECAD has been upgraded several times. It currently receives around 500,000 visits and over 2 million queries a month by, among others, researchers, federal, state and municipal managers, electricity utilities, and the National Social Security Institute (**WWP, 2017**). Brazil's social welfare ministry, MDS, also released anonymised administrative datasets of a sample of Cadastro Único in 2017 (**World Bank, 2019a**)

- In **France**, the Family Allowances Fund (CAF) opened its **data** in 2011, as part of the broader Etalab mission by the then prime minister to facilitate the widest reuse of public data. To increase public awareness and use of open data, the CAF organised a 36-hour non-stop hackathon to identify delivery gaps and co-develop solutions (**National Family Allowances Fund, 2015**).

- In **Belgium**, the National Employment Office (ONEM) developed an **open data platform** as a strategic decision to innovate the institution's statistical strategy. The initiative overcame the issue of data scattered across a multitude of databases by capturing the data landscape in detailed and highly useful monthly tables.
- In **Canada**, Statistics Canada had four different modes of providing researchers with access to microdata, with varying degrees of flexibility and each serving different user communities and, in some cases, providing access to different datasets. The most flexible mode of access allows the public use of anonymised microdata. A more restrictive mode is direct access to anonymised microdata, including administrative data from social and health programmes, in secured research data centres. These are accessible only to researchers with approved projects who have been sworn in under the Statistics Act as 'deemed employees' of Statistics Canada (**OECD, 2020**).

## 2.4 EVALUATING AND TESTING WHAT WORKS

**The use of administrative data for *monitoring* purposes is widely understood, although often underutilised, as argued in sections above. The use of administrative data for *evaluation* purposes, on the other hand, is often entirely ignored.** This is because it has only been in recent years that the digitisation and higher quality of administrative records has made it possible for researchers to leverage administrative data sources for evaluation purposes, especially in middle and high-income countries. The potential for this trend to increase is huge, with major implications for the social protection sector.

**In short, where the limitations of administrative data are thoroughly understood and mitigated, and the benefits explicitly pursued** (see Table 1 for the full list)**, leveraging administrative data can improve data collection and analysis in several ways:** "(i) cheaply and swiftly increasing effective sample sizes to ensure sufficient statistical power to measure policy-relevant changes in outcomes, (ii) allowing for national samples that improve external validity, (iii) providing real outcomes rather than self-reported data" (**Alatas *et al.*, 2021**), (iv) providing multiple observations over time, which makes it possible to analyse duration of participation in social protection schemes, and (v) acting as a full 'roster' from which purposeful samples can be extracted for research purposes, both simple assessments as well as complex evaluations.

**Linking administrative and survey data can also help mitigate the risks and challenges that are inherent to survey data.** Lack of relevant variables, measurement error, nonresponse, and imputed responses can seriously bias official poverty statistics and detected programme effects, and administrative data can help address this (see Box 11). This linkage can be established by including unique identifiers (such as national IDs or social security cards) in household surveys, allowing individuals and households to be matched to those in administrative data records. Alternatively, households can be matched using data such as names and addresses. Whatever the method, data files should be designed in a way that prevents the disclosure of individual identities (**McNabb *et al.*, 2009**).

**BOX 11 Examples of challenges using survey data alone and improvements via administrative data use**

- In the **USA**, **Meyer and Mittag (2019)** show how the use of survey data in assessing the impact of transfer programmes understates the income of the poorest families and the value of transfers received. Revising the impact evaluation using linked administrative data, the estimated programme effects are nearly double.[17]
- Also in the **USA**, a randomised evaluation of the impact on emergency room use of covering uninsured low-income adults with Medicaid health insurance in Oregon found no statistically significant impact when measured using survey data, but a statistically significant increase in emergency room use of 40% using administrative data. Part of this difference was due to greater accuracy in the administrative data than the survey reports for the same time periods and the same set of individuals (**Finkelstein & Taubman, 2015**).
- In **Chile**, **Hoces de la Guardia *et al.* (2011)** discuss how they were unable to use survey data for the evaluation of the Chile Solidario cash transfer programme. As it had no baseline, sample selection was inadequate, the data did not account for attrition, and the unit of observation was not formally established.[18]

**Two key trends of evaluative uses of administrative data emerging from recent global experiences are discussed in the following subsections.**

## 2.4.1 EVALUATING PROGRAMME IMPACTS

**The first key trend is the use of administrative data, alongside more traditional data sources, to evaluate programme outcomes, e.g. enabling quasi experimental methods.** Many social protection programmes, especially if donor-funded, have evaluation requirements, in order to test whether or not intended outcomes are being achieved, as per the programme's theory of change. While traditionally this is done via an ad-hoc before-and-after impact evaluation, leveraging administrative data offers a lower cost alternative, with larger sample sizes, and longitudinal data unaffected by non-response and recall biases.

**Administrative data can help overcome the lack of a baseline for programmes that have already been running for some time.** This has been achieved via Propensity Score Matching and Regression Discontinuity Design (RDD) approaches. For example, in Lesotho, no baseline data was collected for the most recent evaluation of the Child Grant and SPRINGS Programmes. Instead, researchers used administrative records from the NISSA registry, which has national coverage, to compare outcomes among beneficiaries and non-beneficiaries matched on socio-economic characteristics (**Pace *et al.*, 2021**). Similarly, Moldova, Mongolia and the Philippines have successfully used RDD, drawing extensively on administrative data.

**Going one step further, using integrated data infrastructures that routinely link administrative data across sectors allows analysis of the impact of social protection programmes on a wider range of outcomes, with a potentially shorter time-lag.** This expands the scope for research to

17 The transfer programmes are: Supplemental Nutrition Assistance Program (SNAP), Temporary Assistance for Needy Families (TANF), and General Assistance. The survey data comes from the Current Population Survey (CPS). Administrative records on monthly payments for all individuals in New York State were compared to figures in the New York State sample of the CPS. This revealed significant measurement errors in the survey. Receipt of the transfer was not recorded for over one-third of housing assistance recipients, 40% of SNAP recipients, and 60% of TANF recipients.

18 The Chile Solidario Panel Survey (PCHS) was an instrument originally designed for an impact evaluation of Chile Solidario. Instead of relying on the low-quality PCHS for evaluation, the authors used administrative data from the Puente Program and CAS to construct a new panel with seven rounds and over 1,000,000 records for each round. Administrative records ensured a larger sample size in a longitudinal format that included the baseline year as well as detailed socio-economic information.

span policy areas that are traditionally seen as siloes, while shortening the policy cycle (**Langendijk *et al.*, 2018**). As an example, New Zealand has pioneered a data repository – the Integrated Data Infrastructure (IDI) – which brings together administrative records on social protection, education, health, housing, labour market, justice, and population using a unique identifier number. The IDI is made available to researchers and government for impact evaluation purposes, allowing the tracking of individuals across their life cycle and the analysis of how social protection programmes affect outcomes beyond standard indicators.

**An innovative approach is combining administrative data with new sources of 'big data' for impact evaluations.** This was recently seen in Colombia, where records from the country's social registry were matched with geo-located cell phone data to assess the impact of COVID-19 cash transfers on social distancing. In addition to achieving national coverage of beneficiaries and non-beneficiaries, researchers had access to high-frequency data on mobility – going well beyond the 'standard' outcomes that are typically evaluated (**Alvarez *et al.*, 2022**).

## BOX 12 Quasi-experimental evaluation approaches building on administrative data

- In **Lesotho**, the Child Grant Programme (CGP) impact evaluation was based on administrative data from the social registry (NISSA), given the lack of baseline survey data (**Pace *et al.*, 2021**). Beneficiary and non-beneficiary households with similar socio-demographic characteristics were matched and compared using Propensity Score Matching (PSM) to determine programme impacts. A key challenge was that administrative records did not contain monetary measures of income, limiting researchers to self-reported indicators.
- In **Moldova**, a Regression Discontinuity Design (RDD) evaluation[19] (at the proxy means test [PMT] cut-off) was made possible via a combination of administrative data from the Social Assistance Automated Information System (SAAIS) and an ad-hoc once-off survey. While the former was the best source of information on the characteristics of beneficiaries, the amount of money they received and for how long they received it, the latter was fundamental for determining and assessing impacts. A nationally representative Household Budget Survey, to which questions on the receipt of social protection transfers had been added, further helped assess targeting effectiveness. Together, these approaches helped address the lack of a pre-programme baseline survey (Carraro & Ministry of Labour, Social Protection and Family, 2017).
- Similarly, the **Philippines** used administrative data in the RDD evaluation of the Pantawid conditional cash transfers programme (**DSWD, 2020**) as did **Mongolia** in the evaluation of the Food Stamps and Medicard (**OPM, 2014**).
- **New Zealand** routinely makes available linked administrative data from its Integrated Data Infrastructure (IDI).[20] **Wilson and McLeod, (2021)** used the IDI in evaluating the impact of the 2018 Families Package, citing the benefits of: a longitudinal data source unaffected by non-response bias, with national coverage, and that contains information on important sub-groups: newborns, parents, and ethnic minorities.[21] In future studies, the IDI will allow the examination of parental employment and incomes, hospitalisations, school attendance, and education.

19 This is particularly useful for programmes that have already started, as RDD does not require the comparison of a baseline and follow-up survey and instead ensures the identification of a sample of recipients and a credible control group in a once-off data collection exercise.

20 The microdata brings together information about individuals and households in the areas of education, income, benefits, migration, justice, and health among other areas.

21 Sub-group analysis included ethnic groups: Māori, Pacific, or non-Māori, non-Pacific, mothers/first parents supported by benefits, mothers/first parents eligible for paid parental leave, mothers/first parents not eligible for paid parental leave.

- In **Colombia**, **Alvarez *et al.* (2022)** combined administrative data from a social registry[22] with geolocated cell phone data to evaluate the impact of an unconditional COVID-19 cash transfer programme on social distancing. This allowed for national coverage and the identification of de jure beneficiaries, who could also be geo-located in census tracts – allowing for the timely tracking of changes in mobility.
- In **Turkey**, an evaluation of conditional cash transfers for education for Syrian refugees (CCT-E) used a non-experimental approach that leveraged data from the programme's child protection database, as well as payment and attendance data, to evaluate impacts on indicators of interest (in a non-causal way) (**Ring *et al.*, 2020**).

**Another notable approach to evaluating programme impacts is to track changes in people's circumstances and behaviour over time.** Such analysis can help to identify how individuals and households move in and out of social protection programmes and, ideally, the key factors that influence these movements. The longitudinal nature of administrative data makes it a particularly valuable source of information in this respect. Insights can be further deepened when social protection administrative data is linked to the databases of other sectors and/or to survey data. Examples of this approach include various pieces of research in high income countries on the duration of social assistance receipt (and determinants) and analysis of worker participation in social insurance (e.g. in Thailand) (see Box 13). Such studies are commonly undertaken by researchers independent of governments, drawing on publicly available databases (discussed in section 2.3).

## BOX 13 Analysis of changes in circumstances and behaviour over time

- A 2011 study in **Denmark** analysed the extent to which active labour market policies influenced the exit rate from employment of immigrant recipients of social assistance. The study applied a timing-of-events duration model to several administrative datasets collected and merged by Statistics Denmark. The research found large positive effects of labour market programmes on gaining employment (**Heinesen *et al.*, 2011**)
- A similar methodology was used in a 2014 study in **Belgium** exploring determinants of the duration of receipt of social assistance. The study found that shorter durations are influenced by factors including the presence of labour market programmes, living in a larger municipality, and in places where there is provision of more generous supplementary support by a municipality (Carpentier *et al.*, 2014).
- A 2020 study in **Thailand** used administrative data from the national Social Security Fund to explore labour market dynamics. This study applied machine learning techniques to the analysis of administrative data from the fund. One notable finding of the study is that informal and formal employment are quite connected, with many workers moving between different forms of employment over time (Wasi *et al.*, 2020).

22 Administrative data came from the System for the Identification of Potential Beneficiaries, a master household administrative database.

## 2.4.2 INNOVATING AND TESTING TO UNDERSTAND WHAT WORKS

**The second key trend is the piloting and testing of design and implementation options, to understand how these mediate impacts and to inform future design choices.** Indonesia, for example, has been a pioneer in using administrative data to implement experiments aimed at iteratively improving programme design (see Box 14).

**Overall these apply experimental or quasi-experimental research designs, at a lower cost.** Administrative data is used to determine who is in the programme (i.e. define the treatment group that will later receive the 'innovation' in programme design).

**Detailed socio-economic information on beneficiaries available in some (but not all) administrative datasets can also be key to implementing the experimental 'treatment' and in measuring outcomes.** In the Indonesian rice subsidy programme, Raskin, researchers mailed Social Protection Cards with eligibility information to beneficiaries using addresses and names from administrative records. Admin data also helped to measure outcomes more precisely, especially in the context of sensitive data such as health insurance claims and enrolment, which beneficiaries might be reluctant to disclose in a survey.

**Linking admin data from social protection programmes to other administrative databases (i.e. from banks) allows the analysis of a wider range of programme features.** Mode of payment, digitalisation, and efficient spending are topics at the fore-front of discussions on how to improve the functioning of social protection programmes. Access to information on the timeliness of payments, number and qualifications of bank agents by region, and programme spending by category allows researchers to consider the quality of services as outcomes in experimental research designs, and ultimately improve delivery.

### BOX 14 Examples of innovating and testing, while leveraging administrative data

- Administrative records from the **Indonesia** Unified Database (UDB) made possible the design of a randomised control trial to test if the uptake of the rice subsidy programme Raskin could be improved (**Alatas *et al.*, 2021**). Using detailed data from the UDB (names, addresses, assets, eligibility status), the government mailed Social Protection Cards clearly stating eligibility to treatment villages, which increased the amount of subsidy received by 26% (**Banerjee *et al.*, 2018**).
- In another experiment from **Indonesia**'s Raskin programme, administrative records were used in a pilot covering 216 districts where digital payments were introduced instead of in-kind transfers. Surveying all districts would have been prohibitively expensive, so researchers used detailed administrative data from banks to see how the number of agents in each village and their qualifications affected the quality of service (**Alatas *et al.*, 2021**).
- In the evaluation of **Indonesia**'s national health insurance scheme (JKN), researchers wanted to see if a subsidy would boost enrolment among non-poor informal workers and healthier individuals (**Alatas *et al.*, 2021**). Using actual administrative records on enrolments and claims helped to minimise response as well as recall bias. One challenge was that no records were kept on workers who left the programme. However, this was overcome by using a unique national ID number that can be linked to the Social Security Agency administrative database.

- Evaluating the impact of e-governance[23] on corruption in **India**'s public works scheme (MGNREGS) in the state of Bihar, **Banerjee *et al.* (2015)** ran a randomised control trial using actual programme data on spending to measure fund leakage. Data from MGNREGS allowed researchers to track bank balances, credit and debit programme transactions, as well as detailed information on work spells and payments for all workers, none of which would have been available in a survey.
- Interesting approaches have also been used in the study of non-uptake. Examples include an evaluation of policies to increase applications for food stamps among likely eligible individuals in the **USA** (**Finkelstein & Notowidigdo, 2019**) and the TAKE project, which applied different research methodologies including microsimulation models, a field experiment and microeconometric analyses, based on survey and administrative data, to study the non-uptake of various benefits in **Belgium** (**TAKE, n.d.**).

## 2.5 SIMULATING BETTER SOCIAL PROTECTION POLICIES

**The longitudinal nature of social protection data means that it can potentially provide a rich source of information for simulating policy reforms.** This can add significant value, compared to data that only considers individuals and households at a given moment in time. In theory, information within administrative datasets can also shed light on behavioural responses to reforms.

**Arguably the best established model for policy simulation that draws on administrative data is actuarial modelling, which is used for social insurance.** Actuarial models draw on an array of economic, demographic and labour force data, however, of critical importance is information on existing members of a social insurance scheme. In this respect, actuarial models include data on demographics, contribution density, wage levels, entry and exit from a scheme, and benefit payments (**ILO, 2001**). Once these models are built, an important asset of actuarial models is the capacity to simulate how parametric adjustments to social insurance schemes affect their sustainability, coverage and adequacy.

**The analytical potential for simulation tools becomes greater where administrative data is linked to survey data.**[24] The USA's Social Security Administration has made significant efforts to link its various administrative datasets to surveys routinely implemented by the Census Bureau. As well as serving a broader purpose of improving the accuracy of survey data, these linked data sources provide a source for simulation models. The SSA has developed a variety of microsimulation models that can be used to analyse the scope and impact of its programmes in the future and the effect of proposed changes to the Social Security system. The outputs of these models "describe the impact of SSA programs on our economy, society, and beneficiary populations, and provide detailed demographic and economic information on beneficiaries and covered workers" (**McNabb *et al.*, 2009**).

23 In this context, e-governance is defined as the application of information and communications technology for delivering public services, which makes it possible to design mechanisms with fewer agents intermediating the delivery process.

24 Note these are different from tax-benefit simulation models that are based only on survey data. These also play a useful role, and their use has been expanding. See for example the United Nations University World Institute for Development Economics Research (UNU-WIDER) coordinated SOUTHMOD and associated country simulations.

# 3 CONCLUSIONS

This paper has only scratched the surface of the possibilities offered by more strategic uses of social protection administrative data, especially when combined with other data sources. As noted at the outset of this paper, these applications should be viewed as a continuum, with policy priorities, information systems maturity, implementation capacity, and financing influencing how much a country is able to harness them. For more generic considerations also see the Eurostat report *Good Practices in Accessing, Using and Contributing to the Management of Administrative Data* (Eurostat, 2018).

## 3.1 'BIG PICTURE' CONSIDERATIONS

**USE THE DATA.** The first key message of this paper is clear: do not let your administrative data go to waste, use it in whatever ways work best for you and your institution to reveal trends and opportunities, and to guide decision-making and strategic action. Public sector institutions should be no different from private sector ones in leveraging the data they have to do better, transforming data into knowledge and knowledge into action.

**THINK OUTSIDE THE BOX WHEN USING DATA.** Even when administrative data is used, it is rarely used to its full potential, to truly serve the sectoral outcomes we are aiming to achieve (e.g. better protecting people over the course of their lifecycle). This paper puts forward some initial ideas and 'nudges' in terms of what *more* could be done with the data you have, alongside country examples for inspiration. Some of these may feel more complex and unachievable for your country in the short term (e.g. leveraging data for evaluation purposes), but others are simple low-hanging fruit that you can start working on today.

**ADMINISTRATIVE DATA IS INVALUABLE FOR MONITORING AND EVALUATION.** This paper has provided many examples, but the overall message is that administrative data can lower the costs and increase the effectiveness of M&E, serving the social protection sector and beyond. This also requires some shifts in standard practices. For example, old observations are often discarded or archived because they do not have 'administrative value'. However, for evaluation and research purposes, analysing longitudinal data to understand variations over time can be fundamental.

25 The guidance stresses "the two data sets should only meet again if necessary to adjust the data matching technique" (Feeney *et al.*, 2015).

**TAKE DATA PROTECTION SERIOUSLY.** Data protection should be considered at all times, starting from informed consent all the way down to ensuring the full anonymisation, as and when necessary. As a reference, see the SPIAC-B commissioned ***Implementation Guide on Data Protection and Privacy in Social Protection Systems***. In addition, the guide on ***Using Administrative Data for Randomised Evaluations*** also has many useful suggestions on de-identifying data for research purposes, including e.g. separating the 'identifiers' data set from the 'analysis' data set and encrypting the former.[25]

**IF THERE IS NO GOOD REASON NOT TO OPEN THE DATA, OPEN IT (while ensuring strong data protection guarantees, of course).** Your institution may not have the capacity to leverage your data to its full potential, but opening the data for use by other actors (universities, civil society, even the private sector) could have many positive repercussions, as discussed in section 2.3.

**FULLY LEVERAGING ADMINISTRATIVE DATA REQUIRES OTHER INFRASTRUCTURAL INVESTMENTS, especially in foundational IDs and civil registration.** Having a unique identifier that enables the identification and authentication of all individuals in a country can make a huge difference in terms of linking different datasets and broadening the scope of analysis. While some countries have used functional IDs or fuzzy matching for such analysis, this does not offer the same potential. Special attention also needs to be paid to uniquely identifying all individuals within administrative data, even when the unit of observation is the household, not merely, for example, the household head or main recipient. As one 'advanced' example, statistics laws in many Nordic economies give their national statistics offices access to administrative data at unit level (with identifiers) and allow them to link these data with other administrative data sources for statistical purposes.

## 3.2 SUPPLY-SIDE: OPERATIONAL AND DESIGN CONSIDERATIONS

**BUILD YOUR INFORMATION SYSTEMS WITH DATA USE IN MIND.** When designing programme MISs or information systems serving the sector as a whole (e.g. social registries), ensure that the data is being collected, stored and managed to maximise its usefulness. This may include, for example: ensuring that the data complies with relevant data standards and is clearly described in a data dictionary and via metadata, and ensuring that key datapoints are collected (such as timestamps).

**AUTOMATE ANALYSIS WHERE POSSIBLE IN REALTIME.** Especially for routine analysis (section 2.1), the full benefits discussed in this document can only be reaped via full automation of the data extraction and analysis process, in real-time, via dedicated dashboards.

**RETAIN FLEXIBILITY IN YOUR AUTOMATED REPORTING.** Ensuring flexibility in disaggregation and cross-tabulations (e.g. in query and reporting functions) helps different users navigate the data according to their needs. For example, this could mean enabling different cross-tabulations, disaggregation and comparisons against relevant benchmarks.

**DENOMINATORS MATTER.** Presenting data in relative, rather than absolute, terms provides much more analytical value, allowing the creation of indicators that can be measured against desirable outcomes (coverage, adequacy, etc), by using policy-relevant denominators as benchmarks. The more these denominators are pulled into our reporting systems automatically and regularly, so they can be fully embedded in automatised routine reporting, the better (see section 2.1 for examples).

**BROADEN USE OF DISAGGREGATION AND BENCHMARKS.** As stated above, these enable much richer and meaningful interrogation of the data.

**VISUALS, GRAPHS AND MAPS REALLY HELP.** Even the most data-savvy among us can struggle to see patterns in data presented in numeric and table formats. Visuals, graphs and maps make the data speak on its own, while also making it easier to interpret, communicate and action. Building these into your dashboards will make a difference.

**STANDARDISE AND 'EXPLAIN' YOUR DATA.** This is fundamental for two key reasons. First, the wider the variety of actors using your data, the more it is essential that every single data point is understood for exactly what it represents. Second, standardisation will enable comparison and linking with a wide variety of other data sources, multiplying the analytical potential of your data.

**ENSURE ADEQUATE HUMAN RESOURCES FOR DATA PROCESSING AND ANALYSIS.** Key factors include ensuring adequate staffing for these functions and providing appropriate incentives for the processing and use of administrative data. These human resource requirements can be minimised with greater automation.

## 3.3 DEMAND-SIDE: CAPACITY AND ATTITUDE CONSIDERATIONS

**FOCUS ON YOUR USERS (beyond usual suspects and beyond upward accountability).** All too often the only reason we use administrative data is to aggregate insights for upward accountability purposes, for a handful of central level decision makers, M&E officers and 'wallet-holders'. This is just the tip of the iceberg of the potential users of our data, who include: managers and decision makers at all levels of implementation, civil society organisations and research think-tanks supporting us to do better, academics looking at medium-long-term trends, and international organisations benchmarking data globally,[26] to name a few. Our data, and the insights emerging from it, need to serve their (multiple, diverse and flexible) needs. Thus, we need to involve them from the start and iteratively improve based on their feedback.

**TRAIN AND SUPPORT YOUR USERS.** Even where reporting is automated or semi-automated via dashboards, users still require training and support to navigate more complex querying of the data and interpretation of trends. They will also require hand-holding in terms of understanding how further/deeper analysis of the data can make their work easier (usefulness to them specifically). This requires dedicated central-level staff whose function is not to write reports that no-one will ever read, but to support lower levels of implementation to use and interpret the data they need to perform their functions.

**WORK ALONGSIDE YOUR NATIONAL STATISTICS BUREAU.** They can be powerful allies from many perspectives: ensuring national statistics include a focus on social protection (e.g. even only one question on social protection added to a regular household survey can make a big difference[27]), supporting more in depth data analysis, supporting protected sharing of micro-data, and helping to harness data demand.

26 E.g. World Bank ASPIRE and the ILO Social Protection Monitor

27 For more on this see references in this document. A key resource, once it is published, will be Leite, Alas and Reboul (forthcoming, 2023).

# RESOURCES AND REFERENCES

## BIBLIOGRAPHY

ADB, *Part IV: Harnessing administrative data for a more resilient data and statistical system*, Asian Development Bank, 2022, **https://www.adb.org/sites/default/files/publication/812946/part4-nsss.pdf**

ADR UK, *What is administrative data?* [Online], n.d., **https://www.adruk.org/our-mission/administrative-data/**

Alatas, V., Amalia, F., Banerjee, A., Olken, B. A., Hanna, R., Sumarto, S. and Widyasari, P. P., 'Using administrative data to improve social protection in Indonesia', In: Cole, S., Dhaliwal, I., Sautmann, A. and Vilhuber, L. (eds),
*Handbook on Using Administrative Data for Research and Evidence-based Policy*, 2021, **https://admindatahandbook.mit.edu/book/latest/indonesia**

Altwicker-Hámori, S., 'Disability insurance benefit application in Switzerland: An analysis of linked administrative and survey data', *BMC Health Services Research*, 21(1), 982, 2021, **https://doi.org/10.1186/s12913-021-06992-2**

Álvarez, C., Devoto, F. and Winters, P., 'Why do beneficiaries leave the safety net in Mexico? A study of the effects of conditionality on dropouts', *World Development*, 36(4), 641–658, 2008, **https://doi.org/10.1016/j.worlddev.2007.04.014**

Alvarez, E., Gallego, J. and Pecha, C., *Can cash transfers increase social distancing during COVID-19? Evidence from Colombia*, Working Paper Series 34, United Nations Development Programme, 2022, **https://www.undp.org/latin-america/publications/can-cash-transfers-increase-social-distancing-during-covid-19-evidence-colombia**

Baldeon, C. and Arribas-Baños, M. D., *Management information systems in social safety net programs: A look at accountability and control mechanisms*, Social Protection Discussion Paper No. SP 0819, World Bank Group, Washington, DC, 2008, **http://documents.worldbank.org/curated/en/225171468150316147/Management-information-systems-in-social-safety-net-programs-a-look-at-accountability-and-control-mechanisms**

Banerjee, A., Duflo, E., Imbert, C., Mathew, S. and Pande, R., *Can e-governance reduce capture of public programmes? Experimental evidence from India's employment guarantee scheme in Bihar*, 3ie Impact Evaluation Report 31, 2015, **https://www.3ieimpact.org/evidence-hub/publications/impact-evaluations/can-e-governance-reduce-capture-public-programmes**

Banerjee, A., Hanna, R., Kyle, J., Olken, B. A. and Sumarto, S., 'Tangible information and citizen empowerment: Identification cards and food subsidy programs in Indonesia', *Journal of Political Economy*, 126(2), 451–491, 2018, **https://doi.org/10.1086/696226**.

Barca V., *Monitoring mechanisms of the social safety net programmes in the COMCEC Region*, Unpublished, 2014.

Barca, V. and Carraro, L., *Monitoring implementation and evaluating performance: Experiences from cash social assistance in Moldova*, Oxford Policy Management Working Paper, 2013, pp. 2042–1257

Carpentier, S., Neels, K. and Van den Bosch, K., 'How do exit rates from social assistance benefit in Belgium vary with individual and local agency characteristics?', *Research in Labor Economics*, 39, 151–187, 2014, **https://repository.uantwerpen.be/docman/irua/8fcd27/b251eb8a.pdf**

Carraro, L. and Ministry of Labour, Social Protection and Family, *Strengthening the effectiveness of the social safety net project: Consultancy for Ajutorul social evaluation and design*, 2017

Chirchir, R. and Barca, V., *Building an integrated and digital social protection information system*. GIZ, 2020

Cole, S., Dhaliwal, I., Sautmann, A. and Vilhuber, L., 'Using administrative data for research and evidence-based policy: An introduction', In: Cole, S., Dhaliwal, I., Sautmann, A., and Vilhuber, L. (eds), *Handbook on Using Administrative Data for Research and Evidence-based Policy*, 2020, **https://admindatahandbook.mit.edu/book/v1.1/intro.html**

Connelly, R., Playford, C. J., Gayle, V. and Dibben, C., 'The role of administrative data in the big data revolution in social science research', *Social Science Research*, 59, 1–12, 2016, **https://doi.org/10.1016/j.ssresearch.2016.04.015**

Development Pathways, *Launch of new Disability Benefits Database*, Development Pathways, 2019, **https://www.developmentpathways.co.uk/news/launch-of-new-disability-benefits-database/**

DSWD, *Pantawid Pamilyang Pilipino Program Third Wave Impact Evaluation (IE Wave 3) regression discontinuity report*, 2020, **https://pantawid.dswd.gov.ph/wp-content/uploads/2020/11/4Ps-Impact-Evaluation-Wave-3-RDD-Report.pdf**

Dubois, H., & Ludwinek, A. (2014). *Access to benefits*, Eurofound, Dublin, 2014

Eurostat, *Final report: Good practices in accessing, using and contributing to the management of administrative data*, Eurostat, 2018, **https://cros-legacy.ec.europa.eu/system/files/admin-wp1.2_good_practices_final.pdf**

Feeney, L., Bauman, J., Chabrier, J., Mehra, G. and Woodford, M., *Using administrative data for randomized evaluations*, J-PAL North America, 2015, **https://www.povertyactionlab.org/sites/default/files/Admin_Data_Guide.pdf**

Finkelstein, A. and Notowidigdo, M. J., 'Take-up and targeting: Experimental evidence from SNAP', *The Quarterly Journal of Economics*, 134(3), 1505–1556, 2019, **https://doi.org/10.1093/qje/qjz013**

Finkelstein, A. and Taubman, S., *Using randomized evaluations to improve the efficiency of US healthcare deliver*, 2015, **https://www.povertyactionlab.org/review-paper/using-randomized-evaluations-improve-efficiency-us-healthcare-deliver**

Grosh, M., Del Ninno, C., Tesliuc, E. and Ouerghi, A., *For protection and promotion: The design and implementation of effective safety nets*, World Bank Publications, 2008

Hagemejer, K. and Scholz, W., Social budgeting, In: *Handbook on Social Protection Systems,* Edward Elgar Publishing, 2021

Haines, D.E. and Greenberg, B., *Statistical uses of social security administrative data*, American Statistical Association Joint Statistical Meetings, Minneapolis, MN, 2005, **https://www.oecd.org/sdd/36237686.pdf**

Hanbal, R. D. and Prakash, A., 'A rights-based approach to open government data', In: *Proceedings of the Tenth International Conference on Information and Communication Technologies and Development*, New York, NY, USA, Association for Computing Machinery (ICTD '19), 2019, pp. 1–4, **https://doi.org/10.1145/3287098.3287148**

Hatry, H., *Performance measurement: Getting results*, Urban Institute Press, Washington, DC, 1999 (updated in 2006)

Heinesen, E., Husted, L. and Rosholm, M., *The effects of active labour market policies for immigrants receiving social assistance in Denmark*, 2011, **https://www.iza.org/publications/dp/5632/the-effects-of-active-labour-market-policies-for-immigrants-receiving-social-assistance-in-denmark**

HelpAge International, *Work, family and social protection*, 2017, **https://socialprotection.org/discover/publications/work-family-and-social-protection**

HelpAge International, *Visualisation: Adequacy of social pensions. Social pensions database*, [Online], Pension Watch, 2018, **http://www.pension-watch.net/pensions/social-pensions-database/visualisation-adequacy-of-social-pensions/**

Hoces de la Guardia, F., Hojman, A. and Larrañaga, O., 'Evaluating the Chile Solidario program: Results using the Chile Solidario panel and the administrative databases | socialprotection.org', *Estudios de Economía*, 38(1),129–168, 2011

Hotz, V. J., Goerge, R., Balzekas, J. and Margolin, F., *Administrative data for policy-relevant research: Assessment of current utility and recommendations for development*, Report of the Advisory Panel on Research Uses of Administrative Data of the Northwestern University/University of Chicago Joint Center for Poverty Research, 1999.

IBM, *Applying business intelligence to social services and social security*, International Business Machines (IBM), 2001

ILO, *Resolution concerning the development of social security statistics*, adopted by the Ninth International Conference of Labour Statisticians, International Labour Organization (ILO),1957

ILO, *Social security data required for the valuation of a national social security system*, International Labour Organization (ILO), Geneva, 2001, **https://www.ilo.org/wcmsp5/groups/public/---ed_protect/---soc_sec/documents/genericdocument/wcms_205313.pdf**

ILO, *Recommendation concerning national floors of social protection*, R202, 14 June 2012, International Labour Organization (ILO), 2012, **https://www.refworld.org/pdfid/5c77a49f7.pdf**

ILO, *World social protection report 2020–22: Social protection at the crossroads – in pursuit of a better future*, International Labour Organization (ILO), Geneva, 2021, **https://www.ilo.org/wcmsp5/groups/public/@ed_protect/@soc_sec/documents/publication/wcms_817572.pdf**

ILO, *World social protection data dashboards*, [Online], International Labour Organization (ILO), 2022, **https://www.social-protection.org/gimi/WSPDB.action?id=32**

ISSA, *ISSA guidelines: Information and communication technology*, International Social Security Association (ISSA), 2019, **https://ww1.issa.int/guidelines/ict/174551**

ISSA, *The use of analytical technology in social security systems during the pandemic*, International Social Security Association (ISSA), 2021, **https://ww1.issa.int/analysis/use-analytical-technology-social-security-systems-during-pandemic**

ISSA, *16th ISSA International Conference on Information and Communication Technology in Social Security*, International Social Security Association (ISSA), 2022a

ISSA, *Data-driven innovation in social security: Good practices from Asia and the Pacific*, International Social Security Association (ISSA), 2022b, **https://ww1.issa.int/analysis/data-driven-innovation-social-security-good-practices-asia-and-pacific**

Jones P. and Elias P., *Administrative data as a research resource: A selected audit. UK National Data Strategy*, Economic and Social Research Council (ESRC), 2006

Koudou, S. and Dely, J., *Quick disability determination cases: Descriptive statistics for fiscal years 2015–2020*, Social Security Administration Research, Statistics, and Policy Analysis, 2022, **https://www.ssa.gov/policy/docs/rsnotes/rsn2022-02.html**

Langedijk, S., Vollbracht, I. and Paruolo, P., 'The potential of administrative microdata for better policy-making in Europe', In: Crato, N. and Paruolo, P. (eds), *Data-Driven Policy Impact Evaluation*, Springer, Cham, 2018, pp. 333–346, **https://doi.org/10.1007/978-3-319-78461-8_20**

Leite, P, Alas, C. R. and Reboul V., *Measuring performance of social protection programs: Guidance to improve household survey design to collect better data*, (forthcoming, 2023)

Leite, P., George, T., Sun, C., Jones, T. and Lindert, K., Social registries for social assistance and beyond, World Bank, 2017

LibTech India, *Heavy wait: Wage payment delays in NREGA by the central government across caste and payment type*, 2021, **http://libtech.in/wp-content/uploads/2021/11/Heavy-Wait_LibTech_NREGAPaymentDetailsCastePaymentType_AprilSep2021_FINAL.pdf**

Lindert, K., Karippacheril, T. G., Caillava, I. R. and Chávez, K. N. (eds), *Sourcebook on the foundations of social protection delivery systems*, World Bank Publications, 2020

McClanahan, S. and Gelders, B., *Assessing the potential for multi-tiered child benefits in Viet Nam*, International Labour Organization (ILO), 2019, **https://www.developmentpathways.co.uk/publications/assessing-the-potential-for-multi-tiered-child-benefits-in-viet-nam/**

McNabb, J., Timmons, D., Song, J. and Puckett, C., 'Uses of administrative data at the social security administration', *Social Security Bulletin*, 69(1), 2009, **https://www.ssa.gov/policy/docs/ssb/v69n1/v69n1p75.html**

Meyer, B. and Mittag, N., 'Using linked survey and administrative data to better measure income: Implications for poverty, program effectiveness, and holes in the safety net', *American Economic Journal: Applied Economics*, 11(2), 2019, **https://www.aeaweb.org/articles?id=10.1257/app.20170478**

National Family Allowances Fund, *Hackathon des Allocations Familiales*, International Social Security Association (ISSA), 2015, **https://ww1.issa.int/gp/162306**

OECD, *Open government data: Towards empirical analysis of open government data initiatives*, Organisation for Economic Co-operation and Development (OECD), 2013

OECD, *Data and diagnostics to leave no one behind*, Organisation for Economic Co-operation and Development (OECD), 2018, **https://www.oecd-ilibrary.org/sites/dcr-2018-10-en/index.html?itemId=/content/component/dcr-2018-10-en**

OECD, *Harnessing new social data for effective social policy and service delivery*, Organisation for Economic Co-operation and Development (OECD), 2019

OECD, *Impact evaluation of labour market policies through the use of linked administrative data*, VS/2019/0261, Joint OECD-EU analysis of labour market policies, Organisation for Economic Co-operation and Development (OECD), 2020, **https://www.oecd.org/els/emp/Impact_evaluation_of_LMP.pdf**

O'Hara, A., Shattuck, R. M. and Goerge, R. M., 'Linking federal surveys with administrative data to improve research on families', *The Annals of the American Academy of Political and Social Science*, 669(1), 63–74, 2017

OMB, *Using administrative and survey data to build evidence*, The White House Barack Obama Archives, 2016, **https://obamawhitehouse.archives.gov/sites/default/files/omb/mgmt-gpra/using_administrative_and_survey_data_to_build_evidence_0.pdf**

OPM, *Mongolia: Food and Nutrition Social Welfare Program and Project*, Project Number 42322, 2014, **https://www.adb.org/sites/default/files/project-document/153512/42322-012-tacr-01.pdf**

Pace, N., Daidone, S., Bhalla, G. and Prifti, E., *Evaluation of Lesotho's Child Grants Programme (CGP) and Sustainable Poverty Reduction through Income, Nutrition and Access to Government Services (SPRINGS) project*, Food and Agriculture Organization (FAO), 2021, **https://www.fao.org/3/cb4862en/cb4862en.pdf**

Paes-Sousa, R., Regalia, F. and Stampini, M., *Conditions for success in implementing CCT programs: Lessons for Asia from Latin America and the Caribbean*, Inter-American Development Bank, 2013, **https://publications.iadb.org/en/publication/16347/conditions-success-implementing-cct-programs-lessons-asia-latin-america-and**

Palacios, R. and Sluchynsky, O., *Social pensions part I: Their role in the overall pension system*, Discussion Paper 0601, World Bank, 2006, **https://documents1.worldbank.org/curated/en/116531468159898116/pdf/362370Social0P10rev0SP060101PUBLIC1.pdf**

Palisade, *The analytics pyramid*, [Online], 2023, **https://www.palisade.com/analytics-pyramid/**

Ring, H., Rothbard, V., Seidenfeld, D., Stuer, F. and Kamto, K., *Programme evaluation of the Conditional Cash Transfer for Education (CCTE) for Syrians and Other Refugees in Turkey*, American Institutes for Research (AIR), 2020, **https://www.air.org/sites/default/files/UNICEF-Turkey-CCTE-Evaluation-Report-September-2020.pdf**

Rohwerder, B., *CSOs supporting accountability in cash transfer programmes – GSDRC*, GSDRC Publications (Preprint), 2016, **https://gsdrc.org/publications/csos-supporting-accountability-cash-transfer-programmes/**

SASSA, *Fifth statistical report: Payment system*, Department of Monitoring and Evaluation, South African Social Security Agency, 2021, **https://www.sassa.gov.za/statistical-reports/Documents/August%202021%20-%20Report%20on%20Social%20Grants%20-%2016%2009%202021.pdf**

Symons, T., *Datavores of local government: Using data to make services more personalised, effective and efficient*, Nessa Discussion Paper, 2016, **https://media.nesta.org.uk/documents/local_datavores_discussion_paper-july-2016.pdf**

TAKE, *Methodology*, [Online], WordPress, n.d., **https://takeproject.wordpress.com/methodology/**

Tonder, J.-K., *The register-based statistical system preconditions and processes*, International Association for Official Statistics Conference Shanghai, 2008, **https://www.fao.org/3/I9360EN/i9360en.pdf**

Tran, A., *How effective are pension systems in ensuring a minimum income in old age? An index to measure global progress*, Development Pathways, 2021, **https://www.developmentpathways.co.uk/publications/how-effective-are-pension-systems-in-ensuring-a-minimum-income-in-old-age-an-index-to-measure-global-progress-2/**

UNDP, *Implementing comprehensive HIV and STI programmes with transgender people*, United Nations Development Programme (UNDP), 2016, **https://www.unfpa.org/sites/default/files/pub-pdf/TRANSIT_report_UNFPA.pdf**

UNPRPD, *Social protection responses to the COVID-19 crisis for persons with disabilities*, Synthesis Paper, social-protection.org, United Nations Partnership on the Rights of Persons with Disabilities (UNPRPD), 2021, **https://socialprotection.org/discover/publications/social-protection-responses-covid-19-crisis-persons-disabilities-synthesis**

Van Zyl, A., *Greasing the wheels of the accountability system: How civil society organizations close the gap between transparency and accountability*, Paper presented at 3rd Global Transparency Conference in Paris, France, 2013, **citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.646.8547&rep=rep1&type=pdf**

Wasi, N., Devahastin Na Ayudhya, C., Treeratpituk, P. and Nittayo, C., *Understanding a less developed labor market through the lens of social security data*, 2020, **https://www.pier.or.th/files/dp/pier_dp_147.pdf**

White, P., Hodges, A. and Greenslade, M., *Measuring and maximising value for money in social protection systems*, Department for International Development (DFID), London, 2015

Wilson, M. and McLeod, K., *Estimating the impact of the introduction of the 2018 Families Package early-years changes*, Technical report, Ministry of Social Development New Zealand, 2021, **https://www.msd.govt.nz/documents/about-msd-and-our-work/publications-resources/evaluation/family-packages/early-years-impacts-full-report.pdf**

World Bank, *Implementation completion and results report Second Bolsa Familia Project*, Report No. ICR00004576, Social Protection and Jobs Global Practice Latin America and the Caribbean Region, World Bank, 2019a, **https://documents1.worldbank.org/curated/en/578141568314136393/pdf/Brazil-Second-Bolsa-Familia-Project.pdf**

World Bank, *Project performance assessment report: Social Welfare and Development Reform Project in the Philippines*, Report No. 137065, World Bank, 2019b, **https://ieg.worldbankgroup.org/sites/default/files/Data/reports/ppar_philippinesocialwelfare.pdf**

World Bank, *World development report 2021: Data for better lives*, World Bank, Washington, DC, 2021, **doi:10.1596/978-1-4648-1600-0**

WWP, *Unified registry information consultation, selection and extraction tool (CECAD)*, World Without Poverty (WWP), 2017, **http://wwp.org.br/wp-content/uploads/2016/12/wwp_information_management_cecad_eng.pdf**

## ANNEX: THE ADVANTAGES OF THE IMF'S GOVERNMENT FINANCE STATISTICS MANUAL (GFSM) FOR CLASSIFICATION OF SOCIAL PROTECTION EXPENDITURE

**Given the significant issues in classification of social protection expenditure discussed in section 2.1.5, there is a strong case for greater use of classifications in the IMF's Government Finance Statistics Manual (GFSM).** The two main classifications described by the GFSM (see Box 15) have some important advantages, including:

- The GFSM is based on well-established and extensively tested systems of classification (developed by various organisations including the IMF, United Nations, European Union and OECD).
- National governments already routinely report expenditure according to the IMF GFSM (particularly the economic classification). While there are gaps and inconsistencies in reporting, efforts to classify social protection according to the GFSM can contribute to improving the

accuracy and consistency of reporting of government expenditure in general.

- The use of the economic and functional classification makes it possible to disaggregate social protection expenditure by two important dimensions (the type of scheme and the social risk/contingency they address), which are sometimes missing in other statistical classifications. The separation of public servant benefits ('employment-related social benefits') from social security and social assistance benefits is also analytically useful given that these benefits can constitute a large share of social protection expenditure in many countries, but are often aggregated with other 'social insurance' benefits.
- The use of the GFSM means that reported social protection expenditure data is compatible with a whole of government view of expenditure. In other words, social protection can be compared to other sectors (such as health, education etc.), which is not possible when using customised systems of classification that may overlap with other sectors (as in the case of the ADB's Social Protection Index and World Bank's ASPIRE).

## BOX 15 Better classifying social protection expenditure

The IMF GFSM 2014 describes two main approaches to classify government expenditure (with sub-components relating to social protection):

- **The economic classification of expense,** which "identifies the types of expense incurred according to the economic process involved": Under the economic classification, social protection programmes are primarily categorised under Social benefits (27) which may be provided in cash or in kind. This is further disaggregated into: Social security benefits (271); Social assistance benefits (272); Employment-related social benefits (272).
- The **Classification of Functions of Government (COFOG),** which is "a detailed classification of the functions, or socioeconomic objectives, that general government units aim to achieve through various kinds of expenditure": COFOG was originally developed by the OECD and published under the UN. Social Protection (710) is one of ten functions identified by COFOG. This function is disaggregated according to the main risk and vulnerabilities (sometimes called 'contingencies' or 'social risks') that social protection seeks to address, namely: Sickness and disability (7101); Old age (7102); Survivors (7103); Family and children (7104); Unemployment (7105); Housing (7106); Social exclusion not elsewhere classified (n.e.c.) (7107); Research and Development (R&D) on Social protection (7108); Social protection not elsewhere classified (n.e.c) (7109).

An important advantage of this framework is that it relates to both the main kind of social protection (under the economic function) and the main function in relation to different risks (COFOG). The GFSM proposes that cross-classification of economic and functional classifications is analytically useful, and this can be applied to social protection. For example, with adequate data collection and classification it would be possible to identify the different kinds of social benefits (e.g. social assistance) relating to a given function (e.g. family and children).[28]

28 This said, it should be noted that these two classifications (economic and functional) do not perfectly align, that is, it is possible for a social benefit to exist that is not social protection, and possible for a social protection expenditure to be in place that is not a social benefit.

## ACKNOWLEDGEMENTS

This paper has greatly benefited from the in-depth peer review and suggestions provided by Philippe Leite (World Bank), Ludovico Carraro, and Fred Merttens, to whom we are extremely grateful. It has also been enriched by wide ranging consultations during its design stage, with suggestions from Veronika Wodsak, Rodrigo Ortiz D'Avila Assumpcao and Christina Behrendt (ILO); Ernesto Brodersohn (ISSA); Foyzunnesa Khatun, Philip Stokoe and Mauricio Soto (IMF); Melis Guven (World Bank); Anita Mittal, Ralf Radermacher, Saurav Bhattarai and Dominique Leska-See (GIZ); Tomoo Okubo (UNICEF) and; Stephanie Brockerhoff and Judith Munster (OPM).

# IMPRINT

**Published by:**
Deutsche Gesellschaft für
Internationale Zusammenarbeit (GIZ) GmbH

**Registered offices:**
Bonn and Eschborn

**Address:**
Friedrich-Ebert-Allee 32 + 36
53113 Bonn / Germany
T +49 228 44 60 - 0
F +49 228 44 60 - 17 66

Dag-Hammerskjöld-Weg 1-5
65760 Eschborn / Germany
T +49 61 96 79 - 0
F +49 61 96 79 - 11 15

E info@giz.de
I www.giz.de

**Programme/project description:**
We Have the Data, Let's Use it Better – Pushing the Boundaries of Social Protection
Administrative Data Analysis and Use

**Authors:**


**Coordination:**
Dominique Leska-See

**Editing:**
Susan Sellars

**Layout / Illustrations:**
EYES-OPEN, Berlin



**Responsible for the content:**
Ralf Radermacher (GIZ)
Bonn, October 2023

Deutsche Gesellschaft für
Internationale Zusammenarbeit (GIZ) GmbH

Registered offices:
Bonn and Eschborn

Friedrich-Ebert-Allee 32+36
53113 Bonn / Germany
T +49 228 44 60-0
F +49 228 44 60-17 66

Dag-Hammarskjöld-Weg 1-5
65760 Eschborn / Germany
T +49 61 96 79-0
F +49 61 96 79-11 15

info@giz.de
www.giz.de

www.facebook.de/gizprofile
twitter.com/giz_gmbh

On behalf of

Federal Ministry
for Economic Cooperation
and Development