# **Dataset construction protocol**

#### Countries:

Data collection for this dataset was carried out in all low and middle income countries with population above 1 million as defined in the World Bank 2016 Country Classification table (see historical classification for income file available from <a href="https://datahelpdesk.worldbank.org/knowledgebase/articles/906519">https://datahelpdesk.worldbank.org/knowledgebase/articles/906519</a>).

The table identifies 30 low income countries; 50 lower middle income countries and 51 upper middle income countries in 2016. In total 131 countries. The SALMIC dataset contains information for social protection programmes for 109 countries. Chile and Uruguay became high income countries in 2014 but were included. Some low and middle in countries with populations less than 1 million (mainly Pacific Islands) were included where information was available.

Countries with no social assistance programme in place by 2015 are not included in the dataset.

#### Programmes:

Social assistance programmes are the main focus of the dataset. Social assistance programmes are budget-financed and rules based programmes supporting individuals and households vulnerable to poverty through transfers in cash and/or kind.

In low and lower middle income countries, emergency and humanitarian assistance is an important component of development assistance and the work of international organisations. Multilateral and bilateral international organizations disbursing international aid often use the term safety nets to describe combined social assistance and emergency and humanitarian assistance. National governments may work with narrower or broader definitions of social security, safety nets, social assistance.

From an analytical perspective, social assistance is distinct from emergency or humanitarian assistance. Emergency or humanitarian assistance figures in all forms of social organisation as a response to misfortune. Social assistance, on the other hand addresses poverty, vulnerability, and disadvantage as outcomes of economic structures.

The dataset focuses exclusively on social assistance. An inventory of potential social assistance programmes was developed for each country. The definition above was then applied to identify social assistance programmes.

For some countries with a large number of very small or localised programmes, the data collection focused on nationwide, large-scale, and/or leading programmes. It also focused on public programmes. In some countries, especially low income countries, programmes might be partially funded by multilateral or bilateral agencies and/or delivered by not-for-profits or private providers. They were included in the dataset where programme governance structures involved public agencies as main partners.

## Data sources:

The collection of the data was organised around a questionnaire or codebook, describing each of the variables and the specific coding of the information. The codebook was constructed after extensive consultation with specialist researchers. The codebook is available from this website.

The data were collected from a variety of sources: global and regional datasets (ASPIRE, ODI, CEPAL, ADB's SPI, IPC-PG); national government websites; programme agency reports; research papers; evaluation reports; policy documents; IFIs project documentation and reports; and personal communication with programme agencies. The data were checked against existing international databases and national agency reports where available.

## Data harmonisation:

As the main purpose of the dataset is to support comparative research on social assistance in low and middle income countries, data harmonization was a key stage of the dataset construction.

The questionnaire or codebook defines a common structure to the data collection.

Post data collection, further steps were taken to enable comparative analysis.

Reach – Reach is the number of programme beneficiaries. The dataset reports on the reach of programmes in terms of individuals and households. Where the raw data reported the number of individual recipients, it was multiplied by the average household size for the specific country to calculate the full reach of the programme in terms of individuals. Where the raw data reported the number of households, it was multiplied by the average household size for the specific country to calculate the full reach of the programme in terms of individuals. Where the raw data reported the number of households, it was multiplied by the average household size for the specific country to calculate the full reach of the programme in terms of individuals.

Information on average household size was taken from: UNDESA. (2017). *Household Size and composition around the world 2017 - Data Booklet* (Data Booklet No. ST/ESA/SER.A/405). New York: United National Department of Economic and Social Affairs, Population Division. Available from https://www.un.org/en/development/desa/population/publications/pdf/ageing/household

https://www.un.org/en/development/desa/population/publications/pdf/ageing/household size and composition around the world 2017 data booklet.pdf

Budget – All financial data, transfer values and programme budgets for example, are reported in domestic currency and in purchasing power parity (PPP).

PPP conversion factor is the number of units of a country's currency required to buy the same amounts of goods and services in the domestic market as U.S. dollar would buy in the United States. PPP enable consistent comparison across countries. PPP conversion factors were taken from World Development Indicators. Available from <a href="https://data.worldbank.org/indicator/PA.NUS.PPP">https://data.worldbank.org/indicator/PA.NUS.PPP</a> (accessed 21/11/2017).

Reporting on the financial resources associated with social assistance programmes shows significant variation across countries. The codebook included three data fields: (i) budgeted resources (*bugt*); (ii) programme expenditure (*cost*); and (iii) donor/government financing expenditure (*dfinex/govfinex*). One or more of these was reported for the majority of programmes. Where the amount reported referred to more than one year, it was allocated proportionally to each year.

A derived variable *finres* consolidates the information for each country/year by reporting only one data format where more than one was available. This was done by ordering the data (i), (ii), (iii) and selecting the first one available. The variable *finresour* indicates the data format selected.

AB.19.06.19