Leaving No One Behind:

Disaggregating Indicators for the SDGs[i]

October 26, 2015

Our poor ability to understand how people of different ages, capabilities, or income levels fared under the MDGs hampered the design and implementation of strategies to tackle discrimination and ensure achievement of the goals for all.[ii] A number of studies have demonstrated that progress has often been made amongst those groups that are easiest to reach or whose situations are the easiest to ameliorate, leaving many of the poorest and most vulnerable behind.[iii] Others have pinpointed cases where only the poorest benefitted.[iv] For this reason, it is very important that the indicators for Sustainable Development Goals can be disaggregated. This point is emphatically stressed in the many comments submitted to the Inter-Agency and Expert Group on SDG indicators (IAEG-SDGs) – the technical taskforce responsible for deciding the SDG indicator framework by March 2016.

Agenda 2030, the outcome document of the September 2015 UN Summit, calls for "no one to be left behind." As clearly stipulated in the UN Secretary General's Synthesis Report, *The Road to Dignity by 2030,* this means that targets can only be considered achieved if they have been met for all relevant income and social groups. The principle has since been widely accepted and reiterated in numerous other global reports, albeit often using slightly different terminology.[v] To ensure countries fulfill their commitments to leave no one behind, the IAEG-SDGs will need to identify the following:

i.   clear levels of disaggregation for relevant SDG indicators,
ii. a set of indicators that specifically reflect inequalities that are not captured by disaggregation of other indicators,
iii. the capacity gaps and investment requirements of national statistical systems, if countries are to be able to disaggregate indicators with the requisite level of detail, and bring these to the attention of the international community.

The identification of stratification variables can pose major analytical and operational challenges. Data collected through survey instruments or other tools must collect all stratification variables for each household. In particular, problems arise when detailed responses are required for each member of a household, which causes the number questions to increase by a multiple of the number of respondents. The length of each survey and the need to maintain confidentiality for the collection of sensitive data (e.g. on ethnicity) may constrain opportunities for stratifying socioeconomic and other data. In addition, to maintain statistical accuracy, larger sample sizes are necessary for small populations, for example when we want to disaggregate results to subnational areas, or for smaller groups like migrants, income groups, or people with disabilities. The unavoidable effect is a larger, more expensive survey. Similar constraints may apply on the monitoring side due to the limited capacities of many national statistical offices.

Given the importance of disaggregated data, the SDSN recommends that relevant SDG indicators be disaggregated according the following specific dimensions: [vi]

- Sex and gender,[vii]
- Age,[viii]
- Income quintiles/deciles,
- Location or spatial disaggregation (e.g. by metropolitan areas, urban/rural, or districts),
- Disability,
- Ethnicity and indigenous status, to the extent possible,
- Economic activity,[ix]
- Migrant status, to the extent possible.

It is important to note that disaggregation according to these dimensions would not be relevant for all indicators. Indeed, of SDSN's 100 proposed indicators, only 40% are can be stratified by these criteria (see *Indicators and a Monitoring Framework: Launching a data revolution for the SDGs*, Annex 3). For example, SDSN proposes a measure of 'Total Fertility Rate' to track sexual and reproductive health and population growth. This is a measure of the average number of children born to a woman over her lifetime, so disaggregation by sex is obviously unnecessary. Similarly, many of the indicators under Goals 3 and 5 specifically relate to women and children. However, the sample size problem remains, as a sufficient number of women in a survey is required to yield reliable results.

In general terms, data on health, education, and select aspects of wellbeing can already be disaggregated by sex, age, location, and income (by quintile/decile) in most countries using international household surveys such as the Demographic Health Surveys (DHS), Multi-Indicator Cluster surveys (MICS), and Living Standards Measurement Study (LSMS). Information can also be gleaned from national census and vital registration information. However, data collection is patchy (DHS is only collected every 5.88 years[x]), coverage is limited to developing countries, and often data produced by these different surveys is non-comparable.

Substantive investments in national statistical capacity will therefore be required to ensure standardized collection of data relating to all of the above-defined dimensions, including investments in geospatial data infrastructures to enable spatial disaggregation. Big data, remote sensing and new methodological approaches should also be considered to provide an extra overlay of information and/or to reduce the costs of current methods. As highlighted in *Data for Development: A Needs Assessment for SDG Monitoring and Statistical Capacity Development,* the scale of these investments is relatively modest - $1 billion per annum - and yet is critical to us fulfilling our commitments to the poorest and most vulnerable.

In the interim, the IAEG-SDGs should prioritize indicators that can already be disaggregated to the greatest extent possible, including those from internationally compiled household surveys. These surveys will also need to bolster their collection of data relating to disability and ethnicity, and improve the quality and comparability of spatially disaggregated data, but they give us sufficient detail to kick start the process, while we work to strengthen statistical systems and refine additional or alternative indicators for the future.[xi]

Regarding additional indicators focused on acute inequalities, the SDSN proposes to include indicators on relative poverty as well as the income share of the top decile (or a ratio of the top decile to the bottom 4 deciles) to measure income inequalities within countries. Similarly, a number of dedicated indicators have been proposed to capture gender inequality and other inequalities under Goals 5 and 10.

[i] Prepared by Jessica Espey (SDSN) with Eve de la Mothe Karoubi (SDSN), and with thanks to Eric Swanson (Open Data Watch).

[ii] See i) Melamed, C and Samman, E (2013), *Equity, inequality and human development in a Post-2015 Framework*. UNDP HDR Office: New York. And ii) Watkins, K (2013). *Leaving no one behind: an equity agenda for the post-2015 goals*. ODI: London.

[iii] See i) Save the Children (2010). *A Fair Chance At Life: Why Equity Matters for Child Mortality*. Save the Children: London, UK. ii) Wirth, ME et al (2006). "Setting the stage for equity-sensitive monitoring of the maternal and child health MDGs." *Bulletin of the World Health Organization* 84 (7), p 519–27. And iii) Borooah, VK (2004). Gender bias among children in India in their diet and immunization against disease." *Social Science & Medicine* 58:9, p 1719–31.

[iv] In an OPHI study, in nine out of 34 countries, the poorest region reduced the Multidimensional Poverty Index the fastest; in eight countries all subnational regions reduced poverty, and in Kenya the poorest ethnic group reduced multidimensional poverty the fastest.

[v] See i) High Level Panel of Eminent Persons on the Post-2015 Development Agenda (2013). *A New Global Partnership: Eradicate Poverty and Transform Economies*. ii) SDSN (2013). *Action Agenda for Sustainable Development*. And iii) UN Secretary General (2013). *A life of dignity for all: accelerating progress towards the Millennium Development Goals and advancing the United Nations development agenda beyond 2015*.

[vi] These dimensions are based on the key income and social groupings identified in the report of the High Level Panel of Eminent Persons on the Post-2015 Development Agenda. Key vulnerable groupings, discussed in Annex IV, are captured under the aggregated variables proposed above. HLP (2013) *A New Global Partnership: Eradicate Poverty and Transform Economies.*

[vii] For a internationally accepted definition of the distinction between sex and gender, see www.who.int/gender/whatisgender/en/

[viii] We recommend that the disaggregation by age should at a minimum be by the following set of groups: 0-2 years (infants), 2-5 years (pre-school age), 5-14 years (school age), 15-49 years (childbearing age), 15-64 years (working ages) and 65 years and older (elderly persons).

[ix] For example, water use should be accounted for by economic activity using ISIC Rev 4.

[x] According to Alkire (2014), "DHS have been updated every 5.88 years across all countries that have ever updated them (across a total of 155 'gaps' between DHS surveys). Dropping all incidents where 10 or more years have passed between DHS surveys, that average falls only to 5.31 years."

[xi] UNSD advises that the "required disaggregation of statistical indicators by age, gender, geography, income, disability etc. is currently not available for many statistical areas. However, in many administrative data sources, such as vital registration, some of the parameters such as age and gender are part of the original microdata sets. Also location information may frequently be either part of the dataset or its metadata. On the other hand, such parameters can be easily included in surveys, although representativeness in respect to them will require increased sample sizes (thereby significantly increasing the costs). In particular the data collection for countries in special situations and countries affected by conflict will require strong efforts as the abovementioned data sources are frequently not available." See UNSD, (2014), footnote 3.