Save the Children

Methodological Note[1]

Still left behind? – Tracking children's progress against the pledge to Leave No One Behind

1. Introduction

To fulfil the ground-breaking pledge to Leave No One Behind, it is necessary to know who the furthest behind are, monitor their progress, and build public and political understanding about the importance of reaching them. To do so, we have compiled a dataset including selected children's wellbeing indicators disaggregated by sex, location and wealth over time. We estimate trends of health, education and child protection indicators at the national level and for all subgroups and project those to 2030. This enables us to classify countries in terms of their performance and level of convergence between unequal groups. Furthermore, we aggregate our results at the global level, showing global progress of children from the poorest 20% of households in the world. The final output of this project in presented in the Save the Children report 'Still left behind?' (Save the Children 2018).

The paper is structured as follows: Part 2 explains the data sources used in that project. Part 3 explains the methodology to estimate trends and project indicators until 2030, both for the national average as well as for individual groups. Part 4 illustrates the level of inequality on a global level, aggregating country-level data, which includes the calculation of the poorest 20% of households.

2. Data

The following analysis is based on Save the Children's Group-based Inequality Database (GRID), with disaggregated data for selected indicators, including health (under-five mortality, stunting, wasting, vaccination), education (completion of primary, lower and upper secondary), and child protection (child marriage, birth registration). The data set includes currently 299 Demographic and Health Surveys (DHS) between 1985 and 2016, as well as 107 Multiple Indicator Cluster Surveys (MICS) of rounds 3 to 5 between 2005 and 2016. The total data set covers 114 mostly low- and middle-income countries and 72% of all children in 2018. All education data comes from the World Inequality Database on Education (WIDE) by UNESCO, which includes 339 surveys in 110 countries between 1990 and 2016. All surveys in this data set allow disaggregation at least by gender, location (urban/rural), and economic group (quintiles). In preparation to calculate the global poorest 20% at a later stage, we also computed all indicators by income deciles directly from the microdata, where possible.[2]

In addition to this group-specific data, we use national trend data where appropriate to calculate national trends for selected indicators. Trends in under-five mortality until 2016 are estimated by the UN Inter-agency Group for Child Mortality Estimation (2017). National data points for the prevalence of stunting are taken from

---

[1] For questions on the methodology, please contact Alexis Le Nestour (a.lenestour@savethechildren.org.uk) or Oliver Fiala (o.fiala@savethechildren.org.uk).

[2] If no access to the microdata is possible, we compute the indicator by deciles based on published wealth quintiles, assuming a linear interpolation between the mid-points of quintiles.

the Joint child malnutrition estimates (UNICEF/WHO/The World Bank Group 2018). Data for trends in child marriage are taken from our own research project, where we conducted a cohort analysis to recreate long-term trends in child marriage as well as projections until 2030 on country level (Save the Children and World Bank, forthcoming). Additional data on birth registration comes from World Bank (2018) and UNICEF Data (2018).[3] When using national rates or trends from other sources, we align group inequalities estimated in GRID accordingly, by applying the ratio between estimated national averages and national rates from other sources to the group inequalities, keeping relative inequalities constant.

3. Trends & projections

National average

To show trends in children's well-being indicators over time, including their possible development until 2030, we need to calculate trends and projections of the national average at the country level. National trends are computed by using all available information at the national level, and missing points between two years are linearly interpolated. Under-five mortality until 2016 (where yearly estimates by the UN Inter-agency Group for Child Mortality Estimation exist) and child marriage (where yearly data from 1990 have been calculated by Save the Children and World Bank) are two exceptions where no interpolation is needed to reconstruct national trends.

In all cases, forward projections are required from the last available data point to 2030 and in some cases backward projections if there is no information for the first years from 2000. A broad range of methods is possible to project national trends, however for this particular project some requirements are necessary. First, we want to keep the methodology cohesive and comparable across various outcome indicators. Secondly, we assume that countries are likely to follow a similar path of progress. Under this assumption, by observing trends in the sample, we can apply estimates of transition speeds (i.e., time trends) towards developments in countries for which recent data may be scarce. The technique can also be used for backwards projections. These methods have been used among others by Clemens (2004) and Lange (2014) to project likely educational attainment in the future. In its report on malnutrition, Save the Children (2016) also uses this method to project rates of stunting, wasting and overweight.

When implementing projections, one extreme alternative would be to simply assume that all countries follow a similar path, estimate a global transition speed, and apply it to all countries.

The alternative at the other extreme would be to consider that dynamics are specific to each country. While this might be possible for some indicators with available trend data (e.g. under-five mortality), for many countries and indicators, the available trend data is not sufficient to estimate country-specific transition speeds with precision.[4]

[3] Furthermore, we add a small number of additional estimates for selected group inequalities in stunting, child mortality and birth registration to add countries missing in the GRID data set. Data comes from Save the Children reports, UNICEF and national statistical offices.

[4] While the first method of a global transition speed might reflect more changes due to global developments (e.g. cultural or technological changes, changes in norm), a country-specific transition speed might be more connected to changes in national policies and systems within countries.

The option chosen for this paper consists in estimating transition speeds at both the regional and country levels and taking the weighted average of both. The weights for regional and country estimates are the inverse of the variance of the estimates, so that more precise estimates carry more weights. This follows the approach developed by Lange (2014).

Projections are based on a sigmoid function whereby we perform a logit transformation of the selected indicator to estimate regression models. We estimate models at both the regional and country level, including country fixed effects for the latter[5]. The transition speed used for a specific country is computed by estimating the weighted average of the regional and country transition speeds as follows:

$$\widehat{\beta_\iota} = \frac{\widehat{\sigma_\iota^2}}{\widehat{\sigma_\iota^2} + \widehat{\sigma_r^2}}\hat{\beta}_r + \frac{\widehat{\sigma_r^2}}{\widehat{\sigma_\iota^2} + \widehat{\sigma_r^2}}\hat{\beta}_i \tag{1}$$

where $\widehat{\sigma_\iota^2}$ and $\widehat{\sigma_r^2}$ are the estimated variances of the parameters obtained from the country specific and global regressions respectively.

Inequalities

Different methods for projections for all groups

Projecting values up to 2030 for all different groups (sex, urban/rural and wealth quintiles) can be challenging as there are usually few point estimates and relatively small sample sizes. Moreover, we assume projections at the group level need to satisfy some criteria to be credible. (1) Similarly to national projections, rates are naturally bounded between 0 and 1 and projections should all be in this interval. (2) Weighted averages of groups should always be equal to the national average. (3) Groups should not swap order as the result of projections. It is in theory possible that groups left behind catch up and overtake the most advantaged groups but usually the order of groups stay the same and what changes is their relative distance to the average.[6] (4) Projections at the group level should take into account the dynamic in the change in inequalities.

One simple method to project rates at the group level is to continue the observed trend that the group is following. Rates can be transformed into a logit to make sure that estimates will not exit the 0-1 interval. This method would satisfy criteria (1) and (4) by taking into account the dynamic of the group. If the national average is reconstructed from group average, criteria (2) can be satisfied but it may lead to a situation where national projections obtained from a type of group (say wealth quintiles) differ from the national projections obtained from another type of group (say girls and boys). One issue is that group dynamics can lead to cases where marginalized groups overtake best performing groups. This method also uses little information as usually data at the group level is derived from relatively smaller sample sizes and may vary greatly from one survey to another because of sampling error.

To avoid these problems, Lange (2014) uses all information available at the country and global levels to estimate a country specific transition speed that is applied to all groups in the country. The rationale is that all groups will tend to follow the same path of progress and that the most likely future outcome for a group can be deduced by looking at progress of all groups in the country. Moreover, using information from other

---

[5] The definition of world regions follows the World Bank classification, merging Europe & Central Asia with North America for the projection.

[6] This does not restrict the possibility that trends between existing data points are swapping order, but limits the projections after the last observation to reflect the last measured group order.

countries makes possible to estimate transition speeds with more precision and avoid extreme estimates coming from small sample sizes. This method satisfies our desired criteria (1) to (3) but fails to take into consideration the country dynamic in changes in inequalities. Actually, by applying a uniform transition speed for all groups in the country, relative inequalities from projected values will remain constant.

A third alternative is discussed by Guerreiro Osório (2008) in a technical paper on projections of education indicators by groups in Latin American countries. His proposed solution is to estimate the rate of progress of all groups relative the highest possible value of the indicator. One assumption is that the rate of progress tends to decrease when the value is close to 100% (for indicators when 100% is the maximum desirable value). To estimate how the rate of progress slows down close to 100%, he suggests using all available information from all groups over time and estimates the rate of progress for all observations over two periods. The expected rate of progress is obtained by estimating a linear model where the dependant variable is the rate of progress and the explanatory variable is the distance to the maximum value. The linear model is constrained to be 0 when the maximum value is achieved such as there is no more progress at this stage. His method satisfies criteria (1) and (3) but it is not clear how group averages are related to national averages. Dynamics in inequalities are captured in the sense that countries where groups further behind have experienced more rapid progress will tend to converge faster.

Method chosen for projection: estimation of a convergence index

Our method chosen for this project builds on these different methods and should satisfy all the desired criteria (1) to (4). We assume that group progress is relative to national progress, that is group progress can be either faster or slower than national progress. Thus, in cases where left behind groups tend to progress faster than the national average, they will eventually converge with the national average.[7]

To estimate the rate of convergence of groups left behind we observe group rates of progress relative to the national rates of progress for any two consecutive years of observations (thus, in a country with three available data points, we observe two points) by country and types of group. In countries exhibiting convergence, the rate of progress will be faster for groups left behind whereas, in the case of divergence, rates of progress will be slower for groups left behind.

A linear model is fitted to have an estimate of the rate of progress relative to the national average.[8] The explanatory variable is the difference in logit points between the group average and the national average and the dependent variable is the difference in rate of progress between the group and the national average, measured as the difference in logit points between the two periods. The estimated coefficient, the convergence index, is estimated separately for all indicators, country and types of groups. This parameter captures the evolution in inequalities in the country and makes our projections satisfy criterion (4).

The model is constrained to pass by 0, such as the group rate of progress equals the national rate of progress when a group average equals the national average. This ensures that group averages cannot cross the national

[7] Left behind groups can never exceed the national average in our method as this would violate either assumption (2) that the average of groups is equal to the national average, or assumption (3) that groups do not swap order.

[8] Linear model is weighted by the size of the sample used to compute the group estimates, such as larger sample sizes carry more weight, and by the length of the period between two points, such as convergence observed across a longer period carries more weight.

average and that group position relative to the average is unchanged in the projections, thus respecting our desired criterion (3).

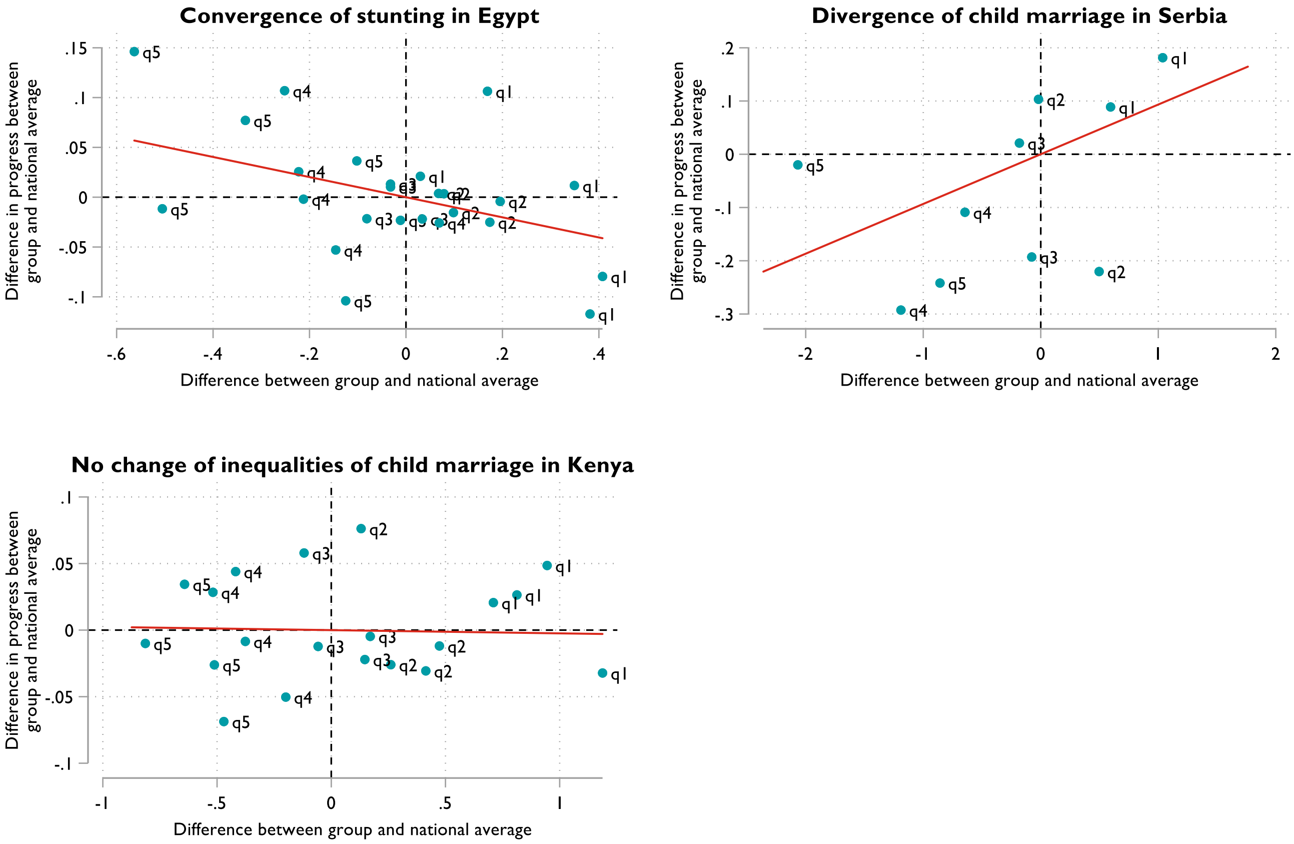

Figure 1: Convergence, divergence and no change over time, between groups (wealth) and national average

Error! Reference source not found. shows cases of convergence, divergence and no change over time. For instance, for stunting in Egypt, groups left behind (that is at the right of the national average rescaled at 0) have on average decreased faster than the national average (that is below the rate of progress of the national average rescaled at 0). The red line depicts the line of best fist across all points and represents the convergence index used at different values relative to the national average. The second case shows divergence in child marriage in Serbia where groups left behind have on average experienced a smaller decrease than the national average. Finally, the last example shows a case where group rate of progress is independent of the position of the group relative to the national average. In the case of child marriage in Kenya, we do not observe faster or slower progress of groups left behind and in this case the convergence index is 0 and indicates that relative inequalities have remained constant.

Based on the estimated national average until 2030, we estimate (starting from the latest available point) in an iterative process the likely rate of change of the group, considering the position of the group relative to the national average and the convergence index for this country and type of group. To ensure that characteristic (2) is satisfied group averages are rescaled at each iteration such as they equal the national average. This matters in the case of divergence when the best performing groups reach 0% or 100% (dependent on the indicator) and left behind groups need to progress to reflect national progress.

4. Global perspective

To draw a global picture of differences in development for different groups in comparison to the average, we are aggregating group-based trends and projections from the national to a global level. However, there are significant difficulties when comparing different groups in a global perspective, particularly economic groups. Household surveys usually classify one household's economic status relative to other households in the sample, grouping those in quintiles from the poorest 20% of households to the richest 20%. However, when comparing household quintiles between countries, we match economic groups which are defined on a national level and which might have almost no overlap using an absolute poverty measurement (e.g. comparing the poorest 20% in Democratic Republic of Congo with the poorest 20% in Mexico).

To address this issue and to measure progress for the poorest 20% globally, we follow the Development Initiatives 'P20' approach, which allows for the identification and comparison of poor households across countries (Development Initiatives 2018). The focus on the global P20 complements our focus on the poorest 20% for our national level analysis.

According to our calculations, the world's poorest 20% of households are living on less than $3 a day. Almost a third of the poorest 20% live in India, followed by 8% in Nigeria and 5% in the Democratic Republic Congo (DRC). At 85%, the DRC is the country with the highest share of its population in the world's poorest 20%. Table 1 shows the countries with the highest contribution to the P20 in 2018.

In order to aggregate globally the poorest 20%, the national headcount of the P20 (percentage of population in a country which is part of the poorest global 20%) must be calculated over time, as the composition is changing due to relative changes in economic growth. The following section explains our calculation of the P20 between 2000 and 2030 as well as the disaggregation of global trends for selected indicators using this concept.

Country	Contribution to poorest 20% (%)	Absolute population in poorest 20% (millions)	Country's population in the world's poorest 20% (%)
India	32	486.0	36
Nigeria	8	127.3	65
DRC	5	71.4	85
Indonesia	5	68.4	26
China	4	62.7	5
Bangladesh	3	49.9	30
Pakistan	3	46.2	23
Ethiopia	3	45.2	42
Tanzania	3	38.4	65
Philippines	2	30.9	29

Table 1: Countries with highest contribution to P20 in 2018

Calculation of P20

There are several pieces of information we need to calculate the national P20 headcount as well as changes in the composition of the P20. Firstly, we need to know the income distribution within countries as well as the mean income. Secondly, we need to know how national incomes have changed or are predicted to change over time, using changes in GDP. Thirdly, we calculate the percentage in every country which is part of the global poorest 20% and compute the contribution of a country towards the P20.

We use information on the country-specific income distribution by decile in 2013 (latest available data) from World Bank PovcalNet (2018). For countries with missing data, we use data on income distribution from the UNU-WIDER World Income Inequality Database (UNU-WIDER 2018).[9] We transform deciles into centiles, using a linear interpolation between mid-points of deciles.

To compute changes in the composition of the P20, we apply growth estimations and forecasts by IMF (2018) until 2022 (latest available year) and OECD (2018) (long term trends after 2022). We apply those growth rates to the mean income for every country in our sample, therefore projecting the mean income for the whole period between 2000 and 2030.[10]

Finally, we calculate the P20 headcount for each country and year in the sample, using the mean income per centile in every country across 30 years. We sort all centiles from the lowest to the highest mean income and identify the percentage of population in each country which falls into the lowest 20%.[11]

Global trends of selected indicators

We use the national trends and projections we developed earlier, both for the national averages as well as disaggregated by groups, to produce global estimates.[12] We compute the trends for the poorest 20% globally, by using the headcount calculated above and applying the trends by decile proportionally. For instance, if a

[9] When no deciles are available from either source, we impute those using information on quintiles or the general average of other countries in the same income group. Similarly, when the mean income is missing, we impute this by using the GDP per capita. All deciles will be converted to centiles when calculating the P20 headcount further below.
[10] While the mean income is changing over time (with effects on the composition of the P20), we are keeping the income distribution of 2013 within countries constant over time. We may include a changing income distribution at a later version of this analysis.
[11] In the 2013, the P20 headcounts calculated with the method are very close to those of Development Initiatives (2018).
[12] As trends for selected indicators in general and/or group-specific trends cannot be directly estimated based on the used data, we impute missing values based on a linear projection using GDP per capita (log) and regions as predictors.

country has a P20 headcount of 37% (meaning that 37% of this country's population is part of the global poorest 20%), we compute the trend for a specific indicator by calculating the average of the three lowest deciles plus 70% of the fourth decile.[13] We calculate the global trends, both for the average as well as for the P20, by computing the weighted average of the trends (using the number of children in the relevant age group for each indicator).[14] The global averages for stunting and under-five mortality (until 2017 and 2016, respectively) are taken from publicly available estimates (UN Inter-agency Group for Child Mortality Estimation 2017; UNICEF/WHO/The World Bank Group 2018). We aggregate country-level estimates (produced with above explained method) and use the rate of change to continue the published trends after the last year for which data is available until 2030. Global rates for child marriage come from our research project on child marriage trends (Save the Children and World Bank, forthcoming).

Figure 2 presents the trends for both the global average and the P20 for under-five mortality, prevalence of stunting, child marriage and primary completion.[15] The graphs also include the path to the SDG targets, illustrating the progress needed to reach the SDG targets by 2030.

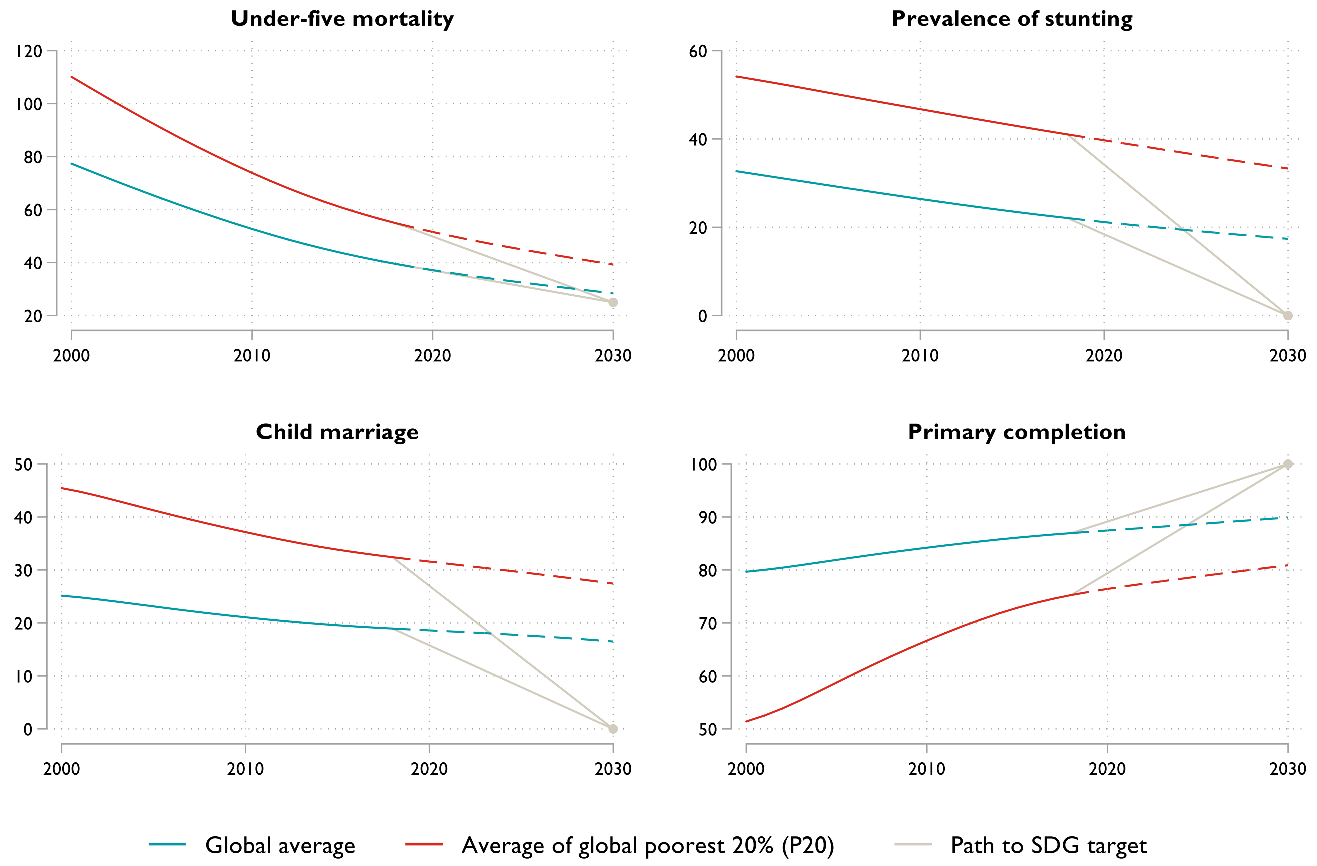

Figure 2: Trends in selected indicators, global average and P20

[13] If the P20 headcount is below 10%, we apply the average of the first decile for the selected indicator.
[14] Similarly, the global trends can be computed for girls and boys, as well as children in urban and rural areas.
[15] We smooth the estimated trends using a weighted running line.

References

Clemens M (2004) The long walk to school: International education glas in historical perspective. Cent Glob Dev Work Pap No 37

Development Initiatives (2018) The P20 Initiative: Data to leave no one behind. http://devinit.org/post/projects/p20-initiative-data-leave-no-one-behind/. Accessed 15 Mar 2018

IMF (2018) IMF Data Mapper: Real GDP growth. http://www.imf.org/external/datamapper/NGDP_RPCH@WEO/OEMDC/ADVEC/WEOWORLD/NGA. Accessed 15 Mar 2018

Lange S (2014) Projections to Zero

OECD (2018) OECD Data: GDP long-term forecast. https://data.oecd.org/gdp/gdp-long-term-forecast.htm. Accessed 15 Mar 2018

Save the Children (2018) Still Left Behind? – Tracking children's progress against the pledge to Leave No One Behind

Save the Children (2016) Unequal Portions

Save the Children, World Bank (forthcoming) Global and Regional Trends in Child Marriage: Estimates from 1990 to 2017

UN Inter-agency Group for Child Mortality Estimation (2017) Levels & Trends in Child Mortality: Report 2017, Estimates Developed by the UN Inter-agency Group for Child Mortality Estimation. New York

UNICEF/WHO/The World Bank Group (2018) Joint child malnutrition estimates - Levels and trends (2018 edition). http://www.who.int/nutgrowthdb/estimates2017/en/. Accessed 15 Mar 2018

UNICEF Data (2018) Birth registration. https://data.unicef.org/topic/child-protection/birth-registration/. Accessed 15 Apr 2018

UNU-WIDER (2018) WIID – World Income Inequality Database. https://www.wider.unu.edu/project/wiid-world-income-inequality-database. Accessed 15 Mar 2018

World Bank (2018) World Bank Open Data. https://data.worldbank.org/. Accessed 15 Apr 2018

World Bank PovcalNet (2018) PovcalNet: an online analysis tool for global poverty monitoring. http://iresearch.worldbank.org/PovcalNet/home.aspx. Accessed 15 Mar 2018