Proposal for a refined approach to monitoring and reporting indicator 4.1.4: Completion rate

GAML6/REF/5
TCG6/REF/13

Prepared by the Global Education Monitoring Report team

Proposal for a refined approach to monitoring and reporting indicator 4.1.4: Completion rate

Background

The completion rate relates directly to SDG target 4.1., which aims to "ensure that all girls and boys **complete** free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes". As such, it was adopted by the TCG as thematic indicator 4.1.4 in 2016. In light of the fact that the global indicator 4.1.1 neglects the often large populations of children who may not reach the grades where learning is assessed, the completion rate was recently proposed to the IAEG-SDG as a second, complementary global indicator for target 4.1.

As noted in said proposal, the completion rate indicator is calculated from household survey data and is therefore subject to time lag in the availability of data, potentially conflicting information from different sources, in addition to sampling and non-sampling errors in survey data. Additional challenges are enumerated below.

To address these challenges and to provide up-to-date and more robust data, the Technical Cooperation Group recommended making use of information from older cohorts who are outside of the age bracket specified in the definition of the indicator to obtain estimates for different years, and following an approach similar to that used for the estimation of child mortality rates.

The international health community faced a similar challenge in measuring indicators, such as under 5 mortality or maternal mortality rates, based on multiple sources. The UN Inter-agency Group for Child Mortality Estimation adopted a consensus model to generate annual estimates for under 5 (Alkema and New, 2012) and neo-natal mortality (Alexander and Alkema, 2018) in each member state. The Inter-Agency Group for Maternal Mortality Rates followed a similar process (Alkema et al., 2016).

Crucially, this approach goes beyond the imputation of gaps. Results of newly-released individual surveys are not treated as 'latest available observations' for purposes of SDG monitoring of mortality. Instead, together with older surveys, they are merely considered raw input data for the generation of model-based estimates that supersede the face-value results of individual surveys.

Here, an adaptation of this approach to estimating school completion rates is proposed. This model allows us to estimate and project completion rates and trends for countries and regional aggregates that are less sensitive to individual surveys and their vagaries, in which years they are conducted, and which one happens to be the latest available for a given country. This paper proposes that, similar to mortality, for purposes of monitoring and reporting school completion the consolidated model-based estimates are used instead of raw survey figures directly.

Rationale for methodological refinement

Current practice is to calculate the completion rate indicators for the cohorts currently at the relevant ages based on each survey, to treat all of these as equally-valid 'observed values' of the indicator, and to report the 'latest available' of these where current figures are presented.

This approach suffers a number of shortcomings:

1. Low periodicity of household surveys mean the 'latest available' data is often several years out of date.
2. Low periodicity also means time series for individual countries are very incomplete. As a consequence, consistent regional time series are unavailable, because the set of countries with available data is different at different points in time.
3. Using only information for individuals within the indicator age bracket at the time of survey throws away useful retrospective information on completion among older cohorts.
4. Carrying forward of the 'latest available' value ignores trends. This is exacerbated by long lags.
5. There is no way to estimate the extent to which changes between observations from different surveys represent real trends in completion or instead differences in bias between different sources.
6. Where several data sources *are* available for a single year, their different levels of sampling variation (and bias) are ignored in calculating the average.
7. While the indicator age bracket of 3 to 5 years above theoretical age for the final grade has a meaningful interpretation in terms of 'reasonably timely completion', no information is gained on the age pattern of eventual completion.

With respect to periodicity, for half of all countries in the UIS database, the latest available survey data on primary completion is for the year 2014 or before. Only 22 countries have data in each of the 5-year windows centred on 2000, 2005, 2010, and 2015, ruling out even an approximate global (not to speak of regional) trend based on a consistent set of countries.

An attempt could be made to resolve these two issues through a simple interpolation between surveys, and by 'carrying forward' the latest available value until a new survey becomes available, while holding on to 'observed values' for the years where a survey is available. *However, this would leave points 3 through 7 unresolved*.

Making use of completion among older cohorts as retrospective information to fill gaps in the country time series throws challenges 4 to 7 into stark relief.

For example, suppose the age bracket for the primary completion rate is 14 to 16 in a given country, and for the time being that at this age, the ultimate level of completion has already been reached. Then a survey in the year 2015 allows for the calculation of the 2015 Completion Rate based on the 14 to 16-year-olds in the sample. In addition, however, completion among 17- to 19-year-olds in the sample may be taken as a proxy for the Completion Rate among 14- to 16-year-olds three years prior in 2012. These retrospective estimates will be more uncertain due to the possibility of selection in terms of migration and mortality. In this way, a single survey contributes completion rate estimates for a series of years.

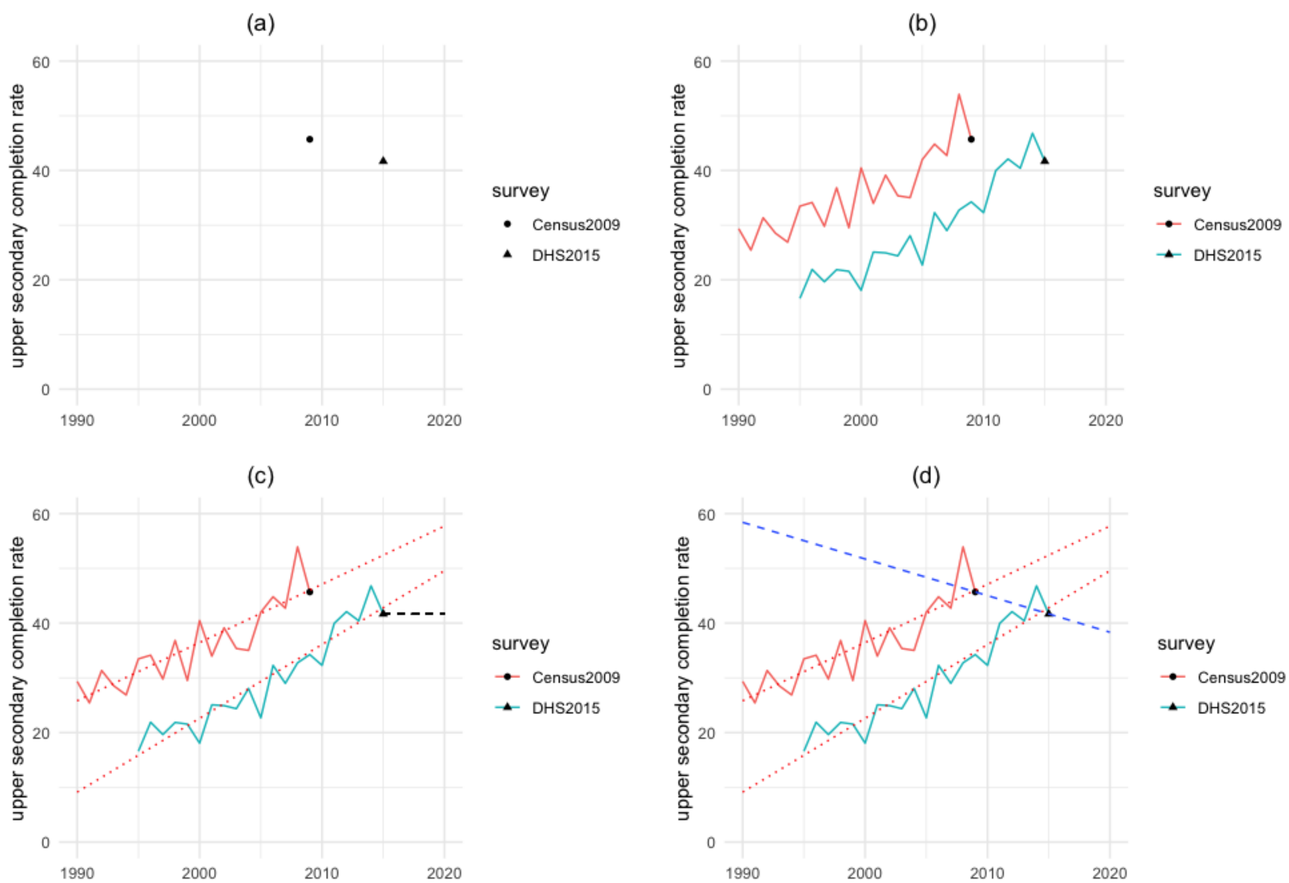

Figure 1: India, upper secondary completion rate. (a) two most recent household surveys. (b) reconstructed time series. (c) trend without (dashed) and with (dotted) reconstructed time series. (c) dashed line = 'latest available' extrapolation.

By way of illustration, consider for simplicity two recent data sources for India (Figure 1). When (retrospective data on) all cohorts are used (Figure 1.b), even countries with only 1 or 2 surveys can provide compelling evidence of a strong trend over time. In this example, it is also obvious that carrying forward the 'latest available' value to the year 2019 (Figure 1.c) is far from plausible given the clear trend, demonstrating challenge 4. With respect to challenge 5, Figure 1.d shows where extracting only the current cohorts' completion from each survey suggests a trend that goes in the opposite direction of the trend evident across cohorts. In other cases the trend without taking into account older cohorts (i.e. between the latest cohorts in each survey) points in the right direction, but is exaggerated in relation to the trend across cohorts. More generally, it is clear in many cases that there are systematic differences in baseline bias between surveys, resulting in 'parallel' series from different sources. Such bias can reflect differences in sampling frames or how questions are asked or coded.

As in the examples already seen, retrospective series will greatly increase the number of years for which more than one source is available. This means that in any case an aggregate estimate needs to be reported rather than raw data directly. This makes challenge 6 more acute, because a simple average of values in each year does not accurately reflect the differences in sample size (and consequently, sampling variation), bias, and other sources of error. The relative magnitude of these effects must be estimated in order to weight different observations appropriately. Even excluding census samples and observations close to 0% or 100% (where statistical variation is intrinsically smaller), estimates from different surveys and years for the same country and level have sampling errors that differ by up to a factor of 10.

In countries where entry can occur very late or where repetition is common, some children or adolescents in the age group examined may still attend school and the eventual rate of completion may therefore be underestimated. Preliminary analyses suggest that this effect can be large. This is clearly evident in the example in Figure 2. The youngest cohorts in each survey, even though they are already several years above the theoretical age for the final grade, consistently display lower completion than those observed at higher ages in a pattern that is highly consistent across surveys.

Timely completion is important and the indicator should reflect this by remaining focused on the age group 3 to 5 years above the theoretical entry age into the final grade of the level in question. Nevertheless, estimates of the degree to which ultimate completion is underestimated by this are valuable for interpretation and cross-country comparison.

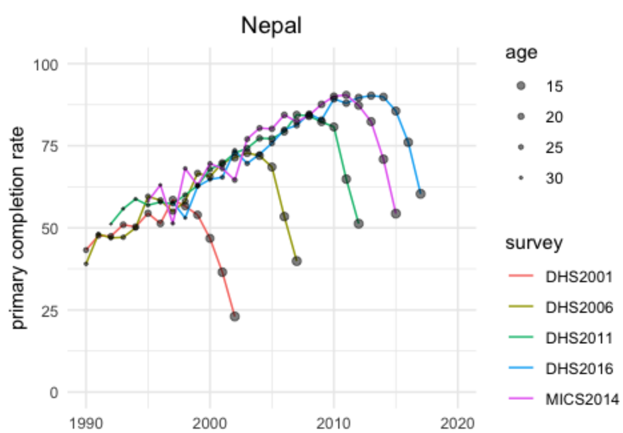

Figure 2: Nepal, primary completion by single *years of age. Cohorts indexed to year in which they reach the middle of the indicator age bracket. The youngest three cohorts in each survey enter the calculation of the completion rate indicator.*

Proposed methodological refinement

Similar to the models endorsed by the IAEG-SDG for the estimation of mortality, the proposed refinement is a hierarchical Bayesian model. Based on observed completion by single years of age from heterogenous survey data, the model directly estimates complete time series of completion by age (to take into account late completion). The existing computation method for the completion rate indicator can then be applied to these estimates. A minor adjustment that is recommended for this final step is to take the unweighted average of the age-specific completion for those 3, 4, and 5 years above the official entry age into the final grade of the level in question.

The model takes into account the following factors:

1. Nonlinearity (floor and ceiling effects at 0% and 100% completion).
2. Country-specific long-term trends.
3. 'Stickiness' of positive and negative changes over time (autocorrelation).
4. Uncertainty of retrospective data increasing with age of reporting cohort.
5. Differences in sampling variation between surveys of different sample size/sample design.
6. Differences in relative bias between different survey series (such as MICS4, or DHS VI).
7. Age misreporting (clustering of reported ages around multiples of 5 in some settings).
8. Late completion up to 8 years above theoretical entry age into the final grade.

Technical details of the model specification are documented in a separate note.

Implications and limitations

Data sources and data availability

The input data requirements and country coverage remain unchanged.

As for the current computation method, the proposed model-based estimation relies on nationally-representative individual-level microdata on school completion collected with censuses and household surveys. With regard to input data, data providers, availability, and limitations are as documented for the current methodology.

The country coverage of estimates remains unchanged (Table 1). The coverage of estimates over time is improved by the proposed refinement: All gaps in country and regional time series are closed, including between the last survey and the present year.

Table 1: Available completion rates by SDG region, 2018 or latest year

Region	Countries with completion rates								
	Primary education			Lower secondary education			Upper secondary education		
	No. of countries	% of countries	% of pop. in region	No. of countries	% of countries	% of pop. in region	No. of countries	% of countries	% of pop. in region
Central and Southern Asia	11	78.6	93.9	11	78.6	93.7	10	71.4	93.8
Eastern and South-Eastern Asia	10	55.6	91.5	10	55.6	91.2	10	55.6	89.7
Europe and Northern America	10	17.5	51.5	41	71.9	100.0	41	71.9	100.0
Latin America and the Caribbean	28	57.1	99.7	28	57.1	99.7	28	57.1	99.7
Northern Africa and Western Asia	12	48.0	75.3	14	56.0	77.0	14	56.0	77.5
Oceania	0	0	0	1	4.0	59.4	1	4.0	35.8
Sub-Saharan Africa	40	78.4	97.6	40	78.4	98.0	40	78.4	96.7
World	111	46.4	88.8	145	60.7	93.8	144	60.3	93.1

Notes: Oceania includes Australia and New Zealand. Population refers to theoretical age for the respective levels of education (primary, lower secondary, upper secondary).
Sources: UIS database, http://data.uis.unesco.org; World Inequality Database on Education (WIDE), https://www.education-inequalities.org.

Quality assurance

With respect to the input data, existing quality assurance processes and needs remain in place. Additional demands on the input data concern the need for data on school completion for older cohorts in their 20s, 30s, and 40s to be reliable.

New quality assurance requirements for the estimated completion rates are sensitivity analyses to various model assumptions and parameters. General quality assurance for the model as such make use of well-established statistical methods such as predictive performance and calibration on hold-out samples.

The availability of short-term ‘projections’ up to the current reporting year raises the question how old the most recent raw input data are required to be. Current thresholds for reporting or dropping the ‘latest available’ may require reconsideration when the question becomes one of projection horizon. In addition, how old the *oldest* data entering the trend calculation are allowed to be arises as an additional consideration.

The existing indicator values already are not sourced from National Statistical Offices or Ministries of Education in individual countries, but they can review and provide feedback. This process could remain unchanged in principle for the model-based estimates. In most cases, model-based estimates for completion in the current year are higher than the ‘latest available’ single survey figure, given broadly expansionary trends. Where the model-based estimate is ‘worse’, additional notes may be required to identify the specific factors leading to this conclusion (e.g. that the most recent survey has been identified as likely exhibiting a positive bias).

While some new quality assurance burden is introduced, there are also gains. In particular, input data will automatically be flagged as potentially erroneous based on its agreement with other data sources. In addition, new recent surveys will further improve the estimates of completion rates for earlier years. The additional time requirement for quality assurance is not so long as to require changes to the data release schedule.

Next steps

This refined approach has already been used to make global estimates (as well as provide a basis for projections to 2030) of primary, lower secondary and upper secondary completion for Meeting commitments: are countries on track to achieve SDG 4?, the joint publication by UIS and the GEM Report prepared on the occasion of the 2019 High-level Political Forum. It is proposed that the Inter-Agency Group on Education Inequality Indicators or a related working group of the Technical Cooperation Group reviews this proposal to finalize technical details and perform further sensitivity analyses. It is also proposed that the refined indicator be considered in conjunction with global indicator 4.1.1 to estimate the percentage of a cohort (rather than of students) that achieve minimum learning proficiency.

References

Alkema, L., and J.R. New (2012). Progress toward global reduction in under-5 mortality: A bootstrap analysis of uncertainty in Millennium Development Goal 4 estimates. *PLOS Medicine* 9(12): e1001355.

Alkema, L., D. Chou, D. Hogan, S. Zhang, A.-B. Moller, A. Gemmill, D. Ma Fat, T. Boerma, M. Temmerman, C. Mathers, L. Say (2016). Global, regional, and national levels and trends in maternal mortality between 1990 and 2015, with scenario-based projections to 2030: a systematic analysis by the UN Maternal Mortality Estimation Inter-Agency Group. *The Lancet*, 387(10017): 462-474.