Technical Cooperation Group
9th meeting
22-24 November 2022

WG/GAML/20

Trends in learning proficiency in the last twenty years: How close are we to reliable regional and global SDG 4.1.1 trend statistics?

Martin Gustafsson
mgustafsson@sun.ac.za
Research on Socio-Economic Policy (ReSEP)
University of Stellenbosch

***DRAFT VERSION** of 15 November 2022*

Global Alliance to Monitor Learning (GAML)

23 November 2022

Trends in learning proficiency in the last twenty years

How close are we to reliable regional and global SDG 4.1.1 trend statistics?

Report prepared for the UNESCO Institute for Statistics

Martin Gustafsson
mgustafsson@sun.ac.za
Research on Socio-Economic Policy (ReSEP)
University of Stellenbosch

***DRAFT VERSION** of 15 November 2022*

In the current version of the report, the learning proficiency trend calculations for both sexes are all complete, though they are subject to change following the UIS's formal review process and inputs from stakeholders in general. Gender-specific estimates will be concluded once the overall methodology has been more widely discussed. There are a few incomplete sections in the report, but they deal largely with the way forward for monitoring proficiency, and do not impact on the estimates of historical trends.

SUMMARY

The milestone 2005 UNESCO Global Monitoring Report, titled *The quality imperative*, marked an important shift towards measuring the learning proficiency of children, not just their attendance in school. The emphasis on proficiency is clearly present in the Sustainable Development Goals, introduced in 2015. SDG indicator 4.1.1 requires the monitoring of reading and mathematics proficiency at three school levels. The 2030 target put forward by the countries of the world was 100% for each of these six indicators.

The current report focusses on estimating regional and global trends for the six 4.1.1 indicators, during the pre-pandemic period 2000 to 2019. The impacts of the pandemic on proficiency have been devastating and exceptional, and are covered in several existing reports. Understanding past proficiency trends, including those from before the pandemic, is of obvious importance for understanding what is possible in the years up to 2030. This understanding can assist in decisions around how adequate or inadequate past efforts to improve learning have been.

The 100% targets for 2030 require global gains to be around 2.7 percentage points a year between 2015 and 2030 – this assumes a starting point of around 60% proficiency in 2015, which is based on actual patterns[1]. Past estimates of global annual gains[2], based on less comprehensive methods than those used for the current report, have ranged from positive 0.7 to *negative* 1.0. Clearly, this would not put the world on track to reach the SDG 4.1.1 targets. The current report confirms that reaching the 2030 targets would be extremely difficult. At the same time, gains in the 2000 to 2019 period appear to have been *positive*, though small.

[1] Figure 18.
[2] Section 9.

The analysis presented here draws from 1,684 proficiency statistics available in the UIS.Stat reporting system for the 2000 to 2019 period, which produce 1,097 two-point trends. It also draws from UIS.Stat statistics on participation in schooling. The latter statistics are important insofar as the practice has been for indicator 4.1.1 statistics to cover just children *within* schooling systems, not those who are out of school. This practice has come about more due to the way data are collected than to any deliberate decision around definitions. Moreover, the relatively high availability of *participation* statistics underscores the fact that *proficiency* statistics should become even more available than is currently the case. The current analysis also uses World Population Prospects data. The share of the world's children accounted for by individual countries and regions is continually changing, and this must be taken into account when national indicator values are aggregated.

Though this report arguably represents the most comprehensive attempt to date to estimate regional and global long-range proficiency trends, this does not mean that the aggregate trends presented in the report are highly reliable. While only 6% of the world's children are in countries where there are no proficiency statistics at all, at any level, 52% of children are in countries where there is no series of comparable values, meaning no trends are detectable[3].

In some regions, trends for the critical early grade reading indicator are mostly absent. For instance, in Sub-Saharan Africa only 29% of children are in countries which *do* measure this[4]. Other regions, such as Central and Southern Asia, have no trend data for this indicator at all.

Much of the lack of trend data is accounted for by India and China, in which 31% of the world's children live and which have no trends for 2000 to 2019, for any level or subject. However, there are some additional 104 countries which also display the same absence of trend data.

A key reason why the just over half of countries which do have trend data cannot be considered representative of all the world's countries is that countries without the data tend to be countries with more serious governance problems. This suggests that these excluded countries would be *less* likely to display proficiency gains. Put differently, the presence of trend data is not randomly spread across the countries of the world. This pattern is true even if world regions are analysed separately[5].

Though the available statistics are likely to over-state the gains made, these statistics point to very modest gains for the 2000 to 2019 period, relative to what is needed to reach the 2030 targets of 100%. In reality, annual improvements in the 2000 to 2019 period have optimistically been around 0.23 percentage points a year, or just one-twelfth of the necessary 2.7 annual gain mentioned above. This figure of 0.23, which refers to end of primary reading[6], is optimistic in the sense that it assumes that where trends have been detected in countries, the years during which measurement took place were randomly selected, meaning countries were not more likely to measure change during periods when improvements were more likely to be occurring. The assumption is optimistic because in reality one could expect better measurement to be occurring when there is a strong focus on improvement initiatives.

Because participation in schooling has improved substantially over the 2000 to 2019 period, it is important to take this into account when examining proficiency trends. It is typically assumed that children who do not complete specific levels of schooling do not achieve the proficiency expected at those levels. Using this assumption, and using a new harmonised dataset of school completion, is is estimated that the abovementioned 0.23 percentage point

[3] Table 7.
[4] Table 8.
[5] Section 6.
[6] Table 8.

gain for end of primary reading becomes 0.33 after completion of primary schooling is taken into account[7]. This is in a context where primary school completion has improved from 77% in 2000 to 86% in 2019[8]. Improvements in primary school completion have thus helped to improve levels of proficiency among children. Even so, annual gains have remained around one-eighth of what they would need to be to reach the 100% targets by 2030.

An annual gain of 0.33 percentage points a year translates to just 3.3 percentage points over ten years. While this may seem unsatisfactory, the historical reality represented by the 0.33 annual gain statistic is important, for two key reasons. Firstly, it suggests that a country achieving this rate of improvement is not failing, it is simply achieving what has *on average* been achieved in the past. Any improvement above this level would be *above average*, and arguably satisfactory, at least relative to past trends. This would be a conservative assumption as the 0.33 annual gain statistic is likely to be an over-estimate, both because it ignores countries with worse governance systems and no data, and ignores the fact that countries which do measure trends are more likely to do so when improvements are occurring.

Secondly, the 0.33 average annual gain underscores how rigorous systems measuring proficiency need to be. Even small changes can be significant.

The end of primary reading indicator receives attention here both because it is an important reflection of the learning process, and because the available data behind this indicator cover a greater proportion of the world's children, compared to the other five SDG 4.1.1 indicators. Specifically, 34% of the world's children are represented by the end of primary reading statistics[9]. This drops to 32% if both proficiency and participation statistics are used[10].

Turning to the other five indicators, all except for end of primary mathematics produce positive annual gains. The mean across the five positive annual gain statistics is 0.4 (before improvements in school completion have been taken into account), and none of the five reach 1.0 percentage point a year[11]. This is at the global level.

The abovementioned 0.33 annual gain for end of primary reading is largely driven by very high annual gain statistics of over 2.0 in the regions Central and Southern Asia and Eastern and South-eastern Asia. The statistics for these two regions must clearly be interpreted with much caution as they draw from just 12% and 16% of the child population respectively. Just one other region in the developing world, Sub-Saharan Africa, displays a positive annual gain for end of primary reading. Here the gain of 0.14 points a year is relatively representative, as it is based on 39% of the child population[12].

What can be done to improve the monitoring of regional and global proficiency trends in the coming years? Clearly, much depends on the capacity and willingness of countries to become more deeply involved in this type of monitoring, be it in the form of international testing programmes or national programmes. There is much that global bodies, such as the UIS, can do to promote this. Engagement with individual countries or individual cross-country testing programmes, such as PASEC, SACMEQ and LLECE, with a view to building capacity and promoting the availability of high quality microdata and metadata, is important. Rigorous analyses of the data by UIS can be considered capacity building insofar as this provides, for instance, national analysts with methods and insights. Clearer procedures whereby countries report on proficiency to UIS, through a bottom-up approach, can result in more national ownership of statistics, but also more reliable trends. This can complement the existing top-

[7] Table 9.
[8] Figure 19.
[9] Table 8.
[10] Table 9.
[11] Table 8.
[12] Table 9.

down approach, whereby around 97% of proficiency statistics in UIS.Stat are drawn from international programmes. International programmes are very valuable, in part because they can enhance measurement rigour and comparability across time and countries, and in part because they are less subject to undue interference by national authorities, in a context where proficiency trends are inevitably sensitive from a political perspective. At the same time, promoting national programmes, which can be tailored to local needs, is important too.

Clearly one thing that is urgently needed is a more standardised approach to viewing indicator 4.1.1 statistics in relation to participation statistics. The two should not be viewed in isolation. Yet how best to bring them together is not a straightforward matter. This report has provided proposed methods for addressing this.

Contents

1 Introduction

National and international debates around progress in schooling systems have over many years shifted towards an emphasis on how well children learn, not just attendance in school. This shift began two decades ago, with the 2005 UNESCO Global Monitoring Report[13], titled *The quality imperative*, marking an important milestone. Learning proficiency subsequently became a central feature of Goal 4 within the Sustainable Development Goals (SDGs), introduced in 2015.

There is a clear demand for regional and global statistics relating to the Sustainable Development Goals (SDGs), including indicator 4.1.1, which deals with the 'proportion of children and young people … achieving at least a minimum proficiency level in (i) reading and (ii) mathematics'. Mandatory disaggregation of this indicator is by level: grades 2/3; end of primary; and lower secondary. This produces six indicators. In addition, each of the six should be disaggregated by sex. The 2030 target put forward by the countries of the world was 100% for each of the six indicators.

Reporting bodies in the UN system have responded to this demand with a variety of estimates in published reports. Producing regional and global statistics is clearly among the responsibilities of the UIS[14]. However, reporting on indicator 4.1.1 *trends* at the regional and global levels remains infrequent and methodologically problematic, as will be explained in this report.

During the COVID-19 pandemic, global estimates of learning losses arising from the disruptions to schooling have been widely quoted. Global and regional trends in this regard have become a matter of interest to a degree not seen before the pandemic. What is needed is to sustain a focus on global, but also regional, proficiency trends, through rigorous methodologies and practices. Global estimates produced during the pandemic were a response to a global emergency, and had to be put together hastily. An opportunity now exists to learn from past attempts and to plan for more reliable estimations of trends.

[13] UNESCO, 2005.
[14] UNESCO, 2017.

Two key research questions inform this report:

1. What do we know and not know about regional and global SDG 4.1.1 trends since around 2005, up to the start of the pandemic in 2020?

2. What are the critical factors needed to improve the availability of regional and global SDG 4.1.1 statistics in the years up to 2030?

Section 2 below introduces the statistics currently available on the UIS.Stat platform needed for learner proficiency trends. Section 3 discusses the reliability of the trends currently seen, with reference to unusual outlier trends and the consistency of data sources. Section 4 explains confidence intervals arising from the sample-based nature of most learning proficiency datasets, and what this means for understanding aggregate regional and global trends.

Section 5 discusses the extent to which international testing programmes, which lie behind 97% of the proficiency statistics available in UIS.Stat, make their microdata and metadata available to a range of analysts. The more such materials become available, and the more analysts use them, including analysts based in national institutions, the greater the certainty around levels of data quality becomes.

Section 6 examines statistically which types of countries have and do not have SDG 4.1.1 statistics pointing to proficiency levels and trends. Currently, just 5% of school-age children are in countries where there are no statistics, at any schooling level, on proficiency. However, 52% of school-age children are in countries for which a *trend* cannot be established for the two decades preceding the pandemic through the available UIS statistics. Moreover, the monitoring of trends in developing countries with respect to the acquisition of reading in the early grades, an especially critical indicator and a prerequisite for subsequent schooling, is often weak. For instance, in Sub-Saharan Africa there is essentially no such monitoring for 71% of children over two decades.

Section 7 looks at how the concentration of the population in specific countries per world region affects the ease with which regional averages can be estimated. Section 8 introduces national and regional child population trends, a matter that can influence global trends to a large degree.

Section 9 presents the methods used in the current report to aggregate national proficiency statistics into regional and global trends. A clear limitation is the low availability of trend statistics, especially in certain regions and for specific schooling levels. Moreover, countries without trend statistics appear to be countries with larger political and social problems, or with less organised schooling systems, suggesting that existing statistics will produce overly positive aggregate trends.

Sections 10 and 11 present aggregate trends where both sexes are combined. Even with the risk that aggregate trends will be overly positive, regional and global proficiency trends are remarkably flat for the period 2000 to 2019. However, this should be seen against the fact that *participation* in schooling, at least at the primary level, has steadily improved. This suggests strongly that although official SDG 4.1.1 statistics, which refer only to children *attending* school have not improved markedly, proficiency levels in the *child population* have improved.

Section 12 examines what the statistics say about differences in trends across gender.

Section 13 presents a basic methodology for combining proficiency and participation statistics with a view to obtaining estimated trends for proficiency in the *child population.*

Finally, section 14 addresses the question of how the availability of regional and global trends for indicator 4.1.1 can be improved. Clearly, this is largely about addressing gaps at the national level, in particular those pertaining to China and India, which jointly account for 31% of the world's children. However, there are further approximately 104 smaller countries for which trend statistics also need to be improved. This section also addresses how trends are reported on, and specifically the need for a more systematic integration of proficiency statistics emerging from testing programmes and statistics on school participation.

A first and second appendix provide aggregations, along the lines of the regional aggregations, according to World Bank income category and according to whether a country is in the Global Partnership for Education (GPE) or not. A third appendix provides an analysis of PASEC microdata, in part to demonstrate the utility of such analysis for probing data reliability questions.

2 The availability of national trend statistics

This report draws largely from what is labelled the '4 March 2022 Release' indicator 4.1.1 statistics in the SDG section of UIS.Stat. These data were downloaded in September 2022. The 228 countries reflected in Table 1 below are those found in the UIS.Stat demographic data tables after three tiny countries with missing population values had been removed[15]. Using only statistics that combine both sexes, there are 1,684 proficiency statistics in the 4.1.1 data, spread across 160 countries. Here and elsewhere in the report, any statistics after 2019 are ignored, as the focus is on the pre-pandemic period. Had 2020 and 2021 statistics been included, there would have been 1,696 proficiency statistics. The 1,684 statistics reflected in Table 1 produce 1,097 level- and subject-specific two-point trends spread across 120 countries. To illustrate the meaning of a two-point trend, if a country had proficiency statistics for the end of primary reading for 2009, 2013 and 2017, this would produce two two-point trends: one for 2009 to 2013, and another for 2013 to 2017.

These 120 countries with trends are 53% of all countries, and account for 48% of the world's children (this will be seen in subsequent Table 7).

More than half of the 52% of children in countries with no trend are from China and India, two countries with no trend in the 2000 to 2019 period. These two countries account for 31% of the world's children. This still leaves 104 countries other than China and India which do not have trend data, at least not for the two decades preceding the pandemic.

Table 1: Availability of national trends

	Countries	Number of 4.1.1 level statistics	Countries covered in previous column	Number of 4.1.1 trends	Countries covered in previous column	% with trend
Sub-Saharan Africa	52	215	40	80	29	56
Northern Africa and Western Asia	24	199	21	135	20	83
Eastern and South-eastern Asia	18	143	16	86	8	44
Central and Southern Asia	14	66	11	24	4	29
Oceania	21	59	6	40	2	10
Latin America and the Caribbean	48	305	24	192	17	35
Europe and Northern America	51	697	42	540	40	78
World	228	1,684	160	1,097	120	53

As will be seen in section 10, around 100% of the world's children are in countries with trend data for *participation* in schooling. This is a key reason why reliable regional and global trends for learning proficiency, or the quality of schooling, are far more difficult to estimate

[15] Holy See, Nieu, Tokelau.

than trends for participation. Put differently, while the Sustainable Development Goals have shifted the emphasis to where it is needed, namely how well children learn, these goals pose a much larger monitoring challenge than the preceding Millennium Development Goals, whose emphasis with respect to schooling was largely on participation.

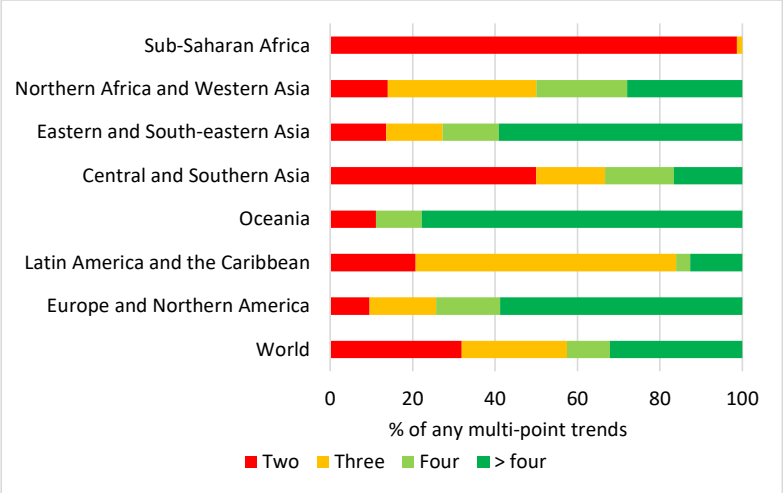
The 1,097 two-point trends referred to previously produce 395 country-, level- and subject-specific multi-point trends, where for instance points in 2009, 2013 and 2017 would produce *one* (mostly) non-linear trend. Within the 395 would be two-point trends which are not extended beyond the two points. How these 395 trends are divided across regions is shown in Table 2. The average number of years covered per trend is also shown. Sub-Saharan Africa clearly lacks longer trends. The ideal would be to have a high percentage of countries with trends (last column of Table 1), but also a high average number of years covered (last column of Table 2). In this regard, apart from the largely developed countries in Europe and Northern America, the region Northern Africa and Western Asia stands out as being relatively well endowed with statistics required to measure trends.

Table 2: Length of all multi-point trends

	Number of level- and subject specific multi-point trends	Average coverage per trend in years
Sub-Saharan Africa	79	5
Northern Africa and Western Asia	50	11
Eastern and South-eastern Asia	22	14
Central and Southern Asia	12	8
Oceania	9	16
Latin America and the Caribbean	87	12
Europe and Northern America	136	14
World	395	11

Figure 1 below provides further details with regard to the length of multi-point trends, by focussing on the number of points per trend. In Sub-Saharan Africa, nearly all trends are just two-point trends.

Figure 1: Number of points per trend by region



3 National trends with some reliability controls

For the purposes of the current report, there are three key reasons why proficiency trend statistics can become unreliable[16]. Firstly, there may be sampling problems. As will be seen, most 4.1.1 statistics are from sample-based testing systems. Apart from the obvious risk that samples may not be nationally representative at different points in time, sampling problems include inconsistencies with respect to when tests are administered. Learners who have just entered Grade 6 and learners who are halfway through Grade 6 are not comparable, even in a context where samples are nationally representative. Changing levels of grade repetition can also influence the comparability of groups of learners from different years. Secondly, what occurs in the testing venue can change over time. A very different test design, different time limits, and a different type or degree of assistance to children by test administrators could apply. Thirdly, there can be problems in the at the data processing or in the statistical approaches intended to improve the comparability of results.

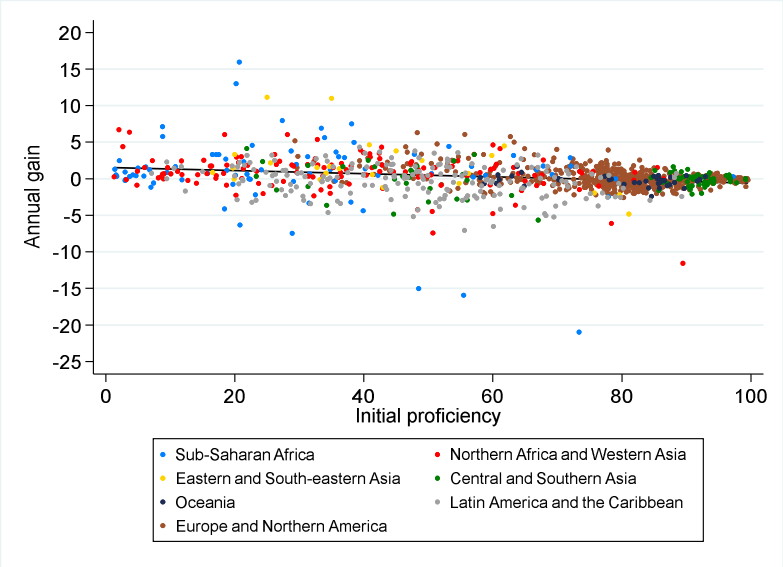
Exceptionally steep upward or downward trends could point to comparability problems flowing from any of the three factors discussed above. Figure 2 below reflects the 1,097 two-point trends referred to in Table 1. Of these trends, 595 trends are greater than zero. None are exactly zero, meaning the remaining 502 are negative. Only slightly more than half of the trends display an improvement.

Of the eight annual gains which exceed 10 percentage point in absolute terms, five are from Sub-Saharan Africa. It is clear from the graph that both positive and negative outliers from this region are fairly common. As might be expected, trends with a lower initial point tend to be associated with a higher gain, though the differences are small. The trendline points to a typical gain of 1.3 percentage points a year where the point of departure is 10% of learners being proficient, and a gain of -0.4 percentage points, in other words a loss, where the point of departure is 90% of learners being proficient.

An earlier UIS analysis of trends in Africa concludes that the continent has displayed gains which are exceptionally large, even if the lower point of departure in Africa is taken into account. To some extent, analysis of both the test and background data associated with typical testing programmes can evaluate whether inconsistencies with respect to sampling, test administration and data processing unduly influenced trends. Appendix 3 below examines PASEC microdata and concludes that sampling and the conversion of raw data to trend statistics do not appear to have unduly affected trends in PASEC, though questions with regard to individual countries do arise. It is clear that providing a broad range of analysts access to a testing programme's data can enhance certainty around strengths and current limitations of particular programmes.

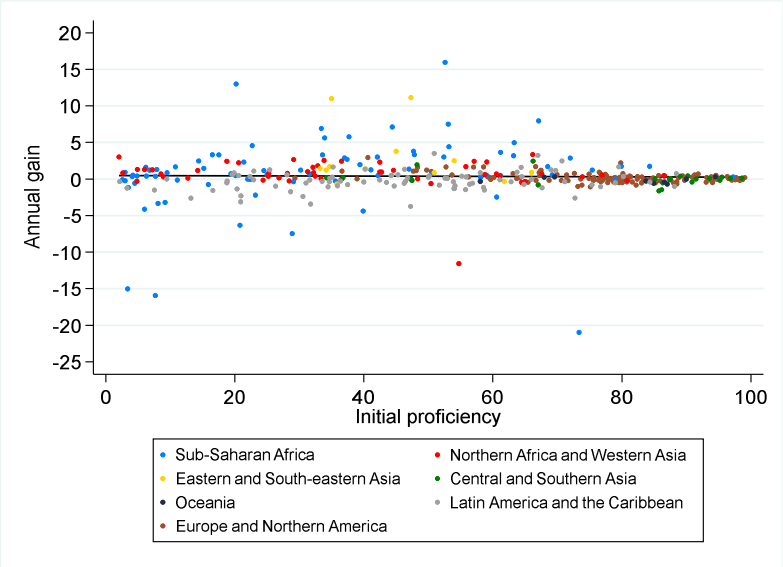
[16] UIS, 2017a, 2018.

Figure 2: Outlier annual gains among two-point trends



If the 395 multi-point trends reflected in Table 2 are graphed, along the lines of Figure 2, the result is the following Figure 3. Of the 395 trends, 234 or over half are positive and 16 display absolute annual gains (or losses) exceeding five percentage points.

Figure 3: Outlier annual gains among 395 multi-point trends



The simple unweighted mean for the 395 points shown in Figure 3 is a gain of 0.4 percentage points a year. *If the target of proficiency of 100% were to be obtained by 2030, the average annual gain would need to be around 2.7 percentage points.* This assumes a global average level of proficiency of 60% in 2015, the approximate 2015 reality for end of primary reading (Figure 18 below). This represents a speed of improvement around seven times faster than the average seen in Figure 3. Put differently, only 30 of the 395 points in Figure 3 represent annual gains of 2.7 percentage points or more. As already pointed out in UIS (2020), while reaching the 100% targets in 2030 is a noble aspirational goal, this would only be possible if the entire world achieved what has historically been exceptional rates of improvement.

Earlier UIS work suggests that annual gains in percentage proficient statistics are unlikely to exceed around three percentage points[17]. If the assumption were used that a change exceeding five percentage points a year should be considered possibly unreliable, then the number of two-point trends would be reduced from 1,097 to 1,064.

Turning to the consistency of data sources behind the 4.1.1 statistics, UIS.Stat now provides data sources as an easily accessible additional variable when data are downloaded. All of the 1,684 proficiency statistics referred to in Table 1 had some source described in the metadata variable. In all, there were 84 unique metadata descriptions. A normalised set of metadata descriptions were generated, following two analytical steps. First, where original metadata descriptions differed from each other only with respect to the grade (or level) and subject, one common descriptor was generated. Thus, several original descriptors all describing the international PASEC programme were changed to 'PASEC'. Moreover, any national monitoring programme tracking a specific level and subject was given the name of the country. Where a monitoring programme provided statistics for just one year, in other words without producing a trend, the new metadata descriptor was left blank. The exercise of normalising metadata descriptors, and excluding metadata not associated with a two-point trend, resulted in the ten descriptors seen in Table 3 below. Of the 1,097 two-point trends referred to in Table 1, 1,086 had the same descriptors for both points in time. The 11 cases with inconsistent data were from five African countries: Ethiopia, Kenya, Lesotho, Uganda, and Zambia. For instance, in Kenya SACMEQ was used for end of primary mathematics in 2013, while a national assessment was used for the corresponding monitoring in 2019. In the case of Ethiopia, national assessments were used for both 2016 and 2019 statistics, but these were clearly different in their approach, according to the metadata, and produced very different proficiency statistics. Of the five countries, Kenya and Uganda also had two-point trends where the metadata *were* consistent.

If both reliability controls referred to above, namely the exclusion of overly steep trends and trends with inconsistent sources, the 1,097 two-point trends are reduced to 1,059.

Table 3: Metadata behind the trends

	Number of two-point trends	% of population-weighted two-point trends
LLECE	116	12.6
PASEC	44	4.5
PIRLS	100	8.3
PISA	598	54.4
SACMEQ	18	1.6
TIMSS	200	15.6
Total for international programmes	1,076	96.9
Bangladesh	4	2.1
Kenya	2	0.5
Kyrgyzstan	2	0.1
Uganda	2	0.5
Total for national programmes	10	3.1
Grand total	1,086	100.0

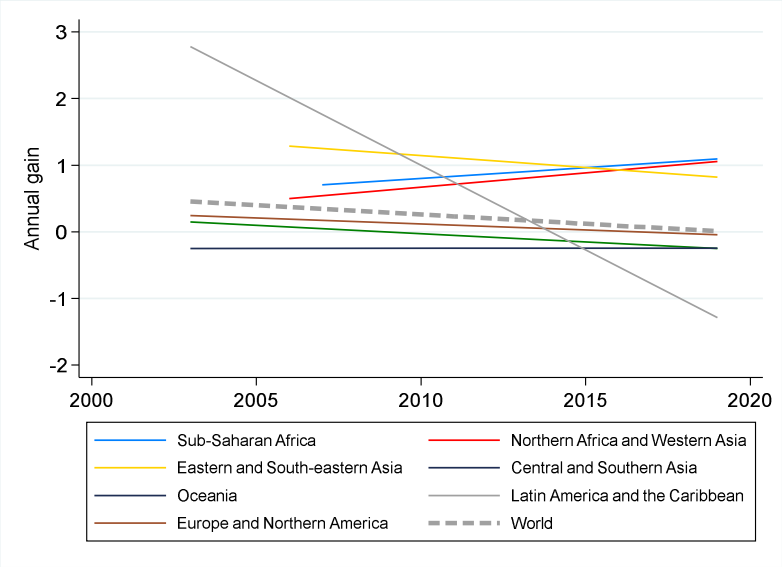
Of the 395 multi-point trends referred to in Table 2, 384 have consistent data sources within each trend. The difference between the two values is accounted for by the 11 two-point trends with inconsistent data sources discussed previously. If both reliability controls are applied to the 395 multi-point trends, the number is reduced to 374. For the purposes of this calculation,

[17] UIS, 2019.

the slope across all points in a multi-point trend, obtained through a bivariate regression, had to exceed five in absolute terms for the multi-point trend to be excluded.

Figure 4 below draws just from the 1,059 reliable two-point trends and examines through simple regression trendlines whether gains became larger or smaller over time. There is no weighting by population in this graph. Globally, gains declined slightly over time, which would be consistent with diminishing gains the higher the baseline, something seen in Figure 1. The predicted global gain values are 0.45 percentage points a year in 2003 and just 0.01 percentage points a year in 2019. Just two world regions display increasing gains over time: Sub-Saharan Africa and Northern Africa and Western Asia. The Latin America and the Caribbean region displays sharply declining gains. Details behind this will be discussed in section 11.

Figure 4: Changes in gains over time



Note: Initial years are not covered by the trendlines as they begin at the earliest end *point in the available two-point trends.*

From Table 3 above it is clear that international programmes predominate as data sources. The last column in the table weights each two-point trend by the child population of each country. Using this approach, 97% of data sources can be said to be from international programmes. This arguably increases the reliability of trends, as large international programmes benefit from an array of skilled technical people from different countries, and would be far less subject to political interference by individual national governments.

4 Confidence intervals around sample-based statistics

The current report presents statistics without confidence intervals, though virtually all the testing behind the assessment programmes reflected in Table 3 above draws from samples of schools and learners in each country. This means that published national proficiency statistics mostly come with margins of error, and that these margins of error can be calculated when national statistics are aggregated to the regional or global level. As explained in UIS (2020), margins of error for regional or global aggregate proficiency statistics are narrower than those at the national level. For instance, a global statistic such as that 57% of children in countries with testing are proficient comes with a 95% confidence interval of around 5.0 percentage points, essentially meaning that one can be 95% certain that the true value lies between 54.5%

and 59.5%. This degree of uncertainty exists around the global annual proficiency statistics referred to in the current report.

How do confidence intervals around individual annual means impact on confidence intervals for the improvement slope? To provide an idea of this, simulations in a statistical package will reveal that given the abovementioned confidence interval of 5.0, given a period of ten years, and given an annual improvement of 0.5 percentage points a year, the slope of 0.5 percentage points will come with a confidence interval of around 0.5 percentage points, meaning one can be 95% certain that the true slope lies within the range of approximately 0.25 to 0.75 percentage points.

5 Public availability of micro- and metadata and implications for reliability

To come.

6 Characteristics of countries with greater 4.1.1 data availability

Given that around half of the world's children are in countries where no official SDG proficiency trends are available (see details in section 11), a key question is whether this half of the world's children are in countries with a lower likelihood of proficiency improvements, compared to the half who are in countries with proficiency trend statistics. Put differently, how representative is the half of the world *with* trend data of the entire world?

To answer this question, a simple measure of trend availability was constructed. The measure used each country's longest, in terms of number of data points, level- and subject-specific trend as an indicator. If a country had no SDG 4.1.1 values at all, zero was assigned to the country, while 1 means some 4.1.1 values exist but no trend is available, 2 means there is at least one two-point trend, 3 means the longest trend has three points, and so on. Any number of points above 5 was converted to 5 as the aim was to explore differentiation among more 'data poor' countries. In constructing the measure, only the 374 trends found after the reliability controls were considered (see section 3) were used.

Table 1 refers to five possible predictors of the availability of proficiency data. Country size is represented by the population aged 0 to 14 (the source here is UIS.Stat – see section 8). It is possible that very small countries lack the capacity or financing to participate in the relevant international testing programmes. The most recent pre-2020 measure of the percentage of children of primary school age who are in school was used an indicator of educational development (the source here is also UIS.Stat – see discussion in section 10)[18]. Countries with low levels of participation could be under-prioritising education and may therefore not be interested in monitoring proficiency. The average annual improvement between around 2010 and 2021 in this indicator was also calculated. The 2019 income per capita with purchasing power parity, as reflected in the World Development Indicators[19], was used as a measure of a country's general level of development. Finally, of the six main indicators within the Worldwide Governance Indicators[20] programme, the indicator 'government effectiveness' was considered the most likely predictor of a country's preparedness to monitor SDG 4.1.1. For all countries, the governance situation in 2019 was used. The governance indicator carries values ranging from -2.5 to 2.5, with greater values representing more government effectiveness.

[18] As will be seen in section 10, completion of primary is the optimal indicator in terms of coverage of the world's children is concerned. However, indicators of being in school are optimal in terms of maximising the number of countries analysed.
[19] https://databank.worldbank.org/source/world-development-indicators
[20] https://databank.worldbank.org/source/worldwide-governance-indicators

Table 4: Means and standard deviations of variables used in the regression

Variable	Mean	Std. dev.	Countries with non-missing values
Points in trend	2.15	2.00	228
Child population (thous.)	8758	31020	228
% in school	92.52	10.70	190
Annual change in above	0.00	1.35	178
PPP income per capita	22153	22765	192
Government effectiveness	-0.03	0.99	206

It should be noted that although the annual change in the percentage of children in school ('Annual change in above') carries a mean of zero, there were somewhat more countries with an upward trend than a downward trend – 98 against 78 respectively.

The outputs of four regression analyses are shown in Table 5. In Model A, which region a country is in emerges as a strong predictor of the 'data-richest' series found. The coefficients on the regional dummies, Europe and Northern America being the omitted reference region, are all negative and nearly all statistically significant. Among the remaining potential predictors, just two are statistically significant at at least the 10% level: higher participation in primary schools and greater government effectiveness are associated with better trend data. Model B allows for more countries to be covered through the exclusion of variables with relatively high degrees of missing values. Model C excludes all countries from Europe and Northern America in order to bring about a stronger focus on developing countries. Finally, Model D excluded just those countries with fewer than 100,000 children aged 0 to 14, in other words particularly small countries with little influence on population-weighted aggregate statistics. Across all four models government effectiveness emerges as a statistically significant predictor. If income per capita is replaced by its natural logarithm, to compress differences among rich countries, income per capita remains statistically insignificant and government effectiveness remains a significant predictor.

Table 5: Regressions on points in trend

	A	B	C	D
Dependent variable → Points in trend				
Constant	1.219	3.679***	-1.730	0.390
Child population (thous.)	0.000	0.000	0.000	0.000
% in school	0.025*		0.028*	0.032**
Annual change in above	0.080		-0.048	0.012
PPP income per capita	0.000		0.000	0.000
Government effectiveness	0.589**	0.676***	0.467*	0.662***
Sub-Saharan Africa	-1.572***	-1.793***	1.046**	-1.207***
Central and Southern Asia	-2.195***	-1.895***	0.375	-1.976***
Eastern and South-eastern Asia	-1.589***	-1.544***	0.851	-1.488***
Latin America and the Caribbean	-1.712***	-1.933***	0.804*	-0.541
Oceania	-2.567***	-2.808***		-1.848***
Northern Africa and Western Asia	-0.023	0.017	2.465***	-0.019
N	166	206	124	141
Adjusted R squared	0.512	0.450	0.343	0.620

*Note: *** indicates that the estimate is significant at the 1% level of significance, ** at the 5% level, and * at the 10% level. Observations (countries) are unweighted.*

It can be concluded that countries without SDG 4.1.1 trend statistics are unlike those with such statistics in important ways. The former tend to display lower levels of government effectiveness, even after controlling for factors such as country size and per capita income.

Table 6 indicates that without any filtering of countries, across all regions other than Latin America and the Caribbean, having SGD 4.1.1 trend statistics with two or more points is

associated with more government effectiveness. Latin America and the Caribbean also fits this pattern if very small countries are excluded (in the case of this region this would largely be Caribbean islands). Thus, even within regions countries with proficiency trends are unlikely to be representative of all countries in the region. Insofar as government effectiveness strengthens the likelihood that improvements in proficiency will be seen, countries with trend data are likely to display better progress than countries without trend data.

Table 6: Government effectiveness within regions

	No filters (206 countries)		Only where child population exceeds 100,000 (171 countries)	
	No trends	With trends	No trends	With trends
Sub-Saharan Africa	-0.970	-0.594	-0.985	-0.639
Northern Africa and Western Asia	-1.349	-0.008	-1.349	-0.008
Eastern and South-eastern Asia	-0.151	1.011	-0.315	0.988
Central and Southern Asia	-0.573	-0.462	-0.573	-0.462
Oceania	-0.296	1.621	-0.537	1.621
Latin America and the Caribbean	0.076	-0.112	-0.652	-0.112
Europe and Northern America	0.369	0.967	-0.409	0.913
World	-0.403	0.262	-0.741	0.209

7 Concentrations of populations with specific countries

It is useful to assess the concentration of children within specific countries per region, as this provides an idea of how many countries, in particular smaller countries, can remain with no statistics while credible regional aggregates can still be calculated. Figure 5 illustrates the global situation. Just two countries, India and China, account for 31% of the world's children. Just 10% of the the world's countries largest countries account for 70% of the world's children.

Figure 5: Global concentration of populations in countries

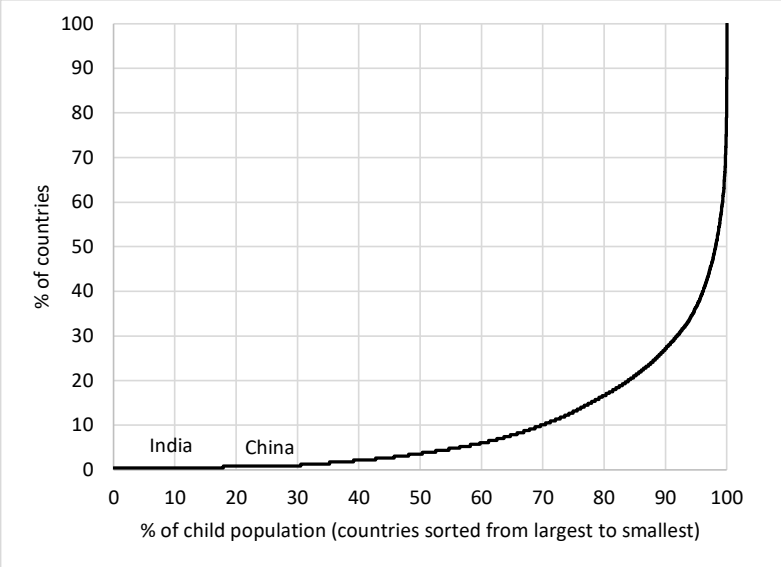
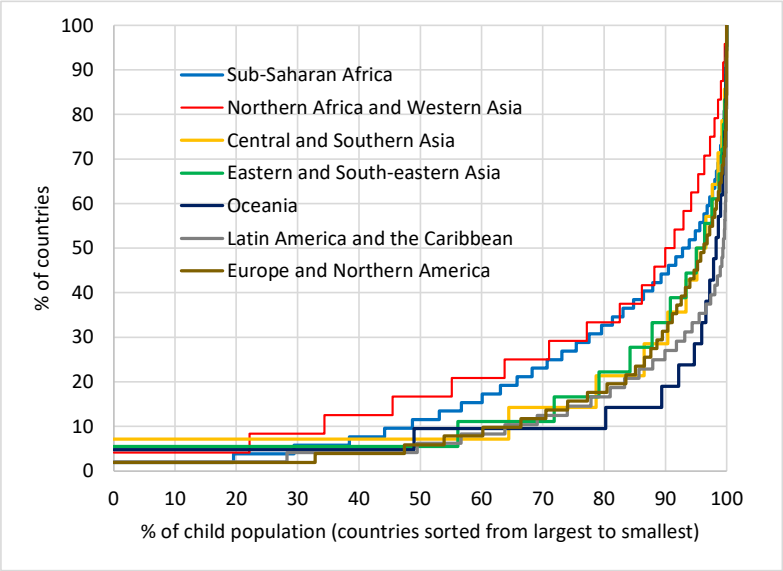


Figure 6 produces similar curves for regions. The region where children are most spread out across countries is Northern Africa and Western Asia. At the other end of the spectrum, child populations are highly concentrated in just a few countries in Oceania.

Figure 6: Regional concentration of populations in countries



8 Trends with respect to child populations

Indicator 4.1.1 as currently used expresses the percentage of *enrolled learners* who reach minimum proficiency levels, not the percentage of the child population. This practice has come about more due to the way data are collected than to any deliberate decision around definitions. The wording of the 4.1.1 indicator is arguably ambiguous when it comes to the question of whether attending learners of an age cohort of children should be the denominator.

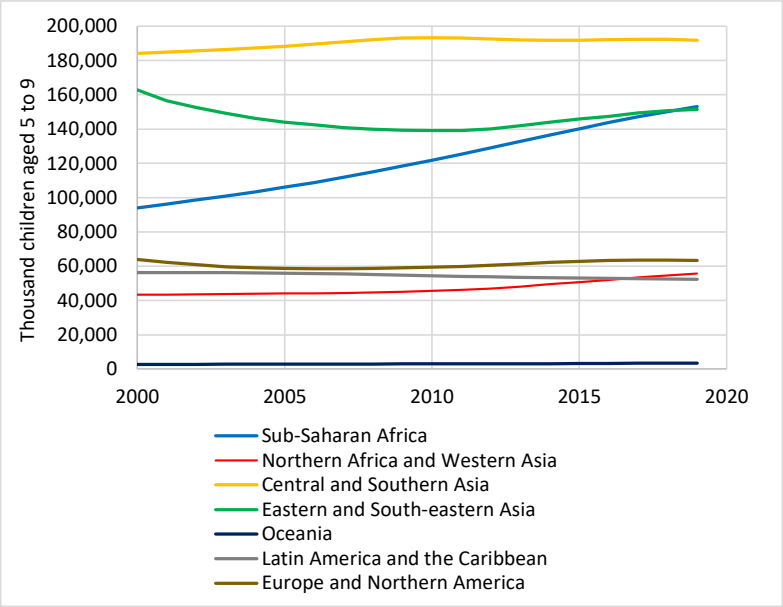
Given the way the indicator is calculated, it would have been consistent to weight country statistics using *enrolments* at specific levels. However, a good enrolment dataset, by grade, covering all countries in the period 2000 to 2019 was not found. Instead, historical population totals in five-year age bands published as part of the World Population Prospects 2022[21] was used. This is likely to approximate well what would have been obtained had weighting by enrolment been used. In particular, the major effect of a rising number of children in Sub-Saharan Africa is likely to be captured as well using population statistics as using enrolment statistics.

The historical World Population Prospects data are used only for the calculation of aggregate proficiency trends. In other situations where population is used to weight countries in the report, the most recent age 0 to 14 total per country available in UIS.Stat in 2022 was used. The most recent year per country was mostly 2022.

As can be seen in Figure 7, children aged 5 to 9 increased in Sub-Saharan Africa between 2000 and 2019 by over 50%, an increase that clearly exceeds that of the other regions.

[21] https://population.un.org/wpp

Figure 7: Trend for children aged 5 to 9



Population of ages 5 to 9 was used for weighting grades 2/3, ages 10 to 14 were used for end of primary, and 15 to 19 for lower secondary. At the secondary level, where the population who are are out of school is higher than at the primary level, is likely to produce *levels* of proficient youths which is over-stated if population instead of enrolments are used. However, the distortion would be less serious with regard to *trends*, assuming a relatively constant secondary school participation rate. Clearly this is one area where there is room for improvement in future estimation of regional trends, assuming that sufficiently comprehensive and reliable enrolment numbers, or participation rates, are obtained.

9 A methodology for aggregating levels and trends

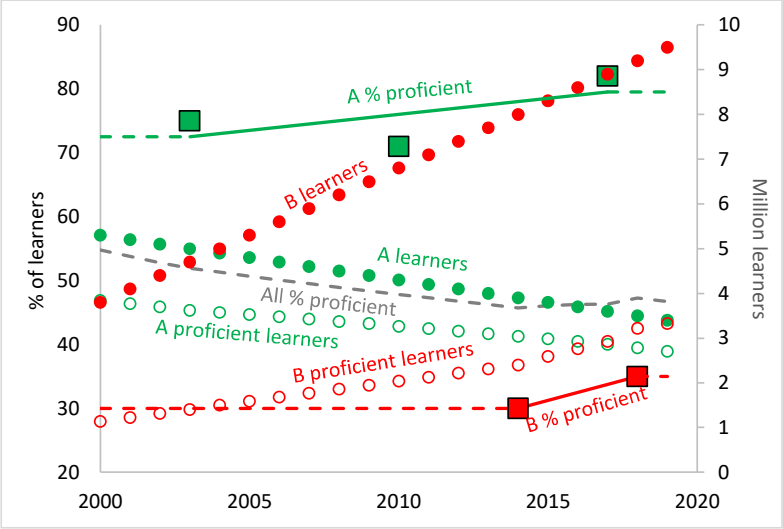
Not much work has occurred to gauge pre-pandemic learning proficiency trends at the regional and global levels. UIS (2021b: 44), a report focussing the impacts of the pandemic on learning, assumes that prior to the pandemic global learning proficiency was improving by around 0.7 percentage points a year. This is based in part on trends among developing countries seen in PIRLS[22], and would be line with some work by the World Bank[23]. Improvements of around 0.7 percentage points a year are also seen in rough (and completely linear) global estimates for all six 4.1.1 indicators available on the in the SDG section of UIS.Stat[24]. What appears in the current report is the most comprehensive attempt to date to gauge regional and global pre-pandemic trends. Despite, this even the work presented here suffers from several limitations, such as serious gaps in the source data. The details of this are discussed extensively.

Figure 8 below illustrates one of two methods which will be followed in the current report to calculate aggregate trends for groups of countries. In both methods, only countries with trends based on consistent sources are used to feed into the aggregate trend statistics.

[22] UIS, 2019.
[23] World Bank, 2019: 18. Note, however, that World Bank (2022: 9) concludes that learning proficiency has declined, even before the pandemic. Specifically, a decline of around 1.0 percentage point a year globally for the period 2015 to 2019 is found.
[24] Accessed November 2022. Regional trends are also linear and largely align to estimates in UIS (2021).

Figure 8 illustrates two hypothetical countries, A and B, the only two countries in the group for which an aggregate trend must be estimated. Country A has proficiency statistics for one level and subject in three years: 75% in 2003; 71% in 2010; and 82% in 2017. These three points are indicated by square green markers. A simple linear trend, obtained by regressing year on proficiency, reflects an upward trend. Before 2003 and after 2017 the trend is assumed to be flat, and at the two levels of the ends of the trendline. The flat initial and final trends are what make this the 'restricted' trend method. The curve 'A learners' indicates that the learner population relevant to the level is in decline, from 5.3 million in 2000 to 3.4 million in 2019 – this trend should be read against the right-hand axis. As discussed in section 8, age-specific population statistics will be used to represent learner populations. The 'A proficient learners' is the population for each year multiplied by the proficiency statistic for each year, the proficiency statistic being read off the 'A % proficient' curve.

Figure 8: Aggregation method with restricted trend



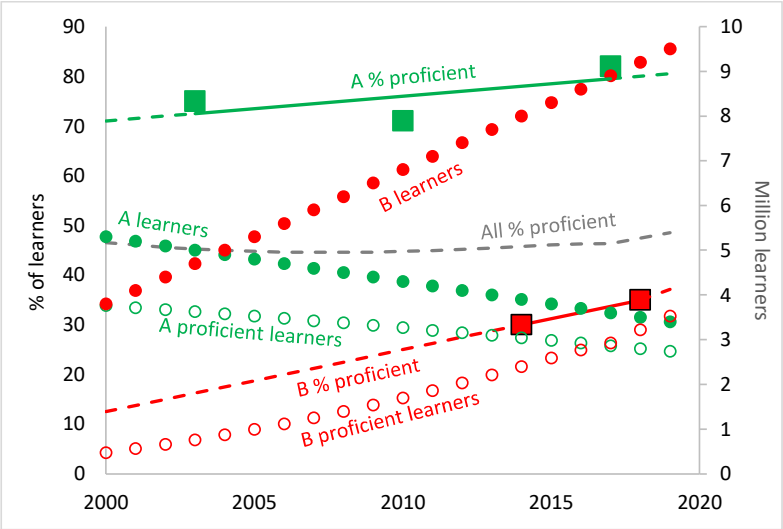
Country B, with a larger and increasing learner population, and low but improving proficiency levels, is handled in the same way as Country A. Because Country B only has two proficiency statistics, 30% in 2014 and 35% in 2018, the simple connecting line represents the trend. As in the case of Country A, for Country B initial and final trends, for which no proficiency statistics exist, are assumed to be flat.

It is clear that although Country A's trend for its proficiency statistics is an upward one, the trend for the *number* of proficient children is downward, because the child population is declining. Adding the proficient learners across A and B and dividing by the total learners across A and B, for each year, produces the aggregate trend 'All % proficient'. In this example, that trend is *negative*, despite the fact that both countries display positive trends for their proficiency statistics. This is because of changing learner population numbers. As population growth is strong in a country with a generally low level of proficiency, Country B, and as population is declining in a country with a generally high level of proficiency, Country A, the overall effect is one of a decline in the overall percentage of learners who are proficient. As will be seen in section 10, demography does not influence trends to the extent illustrated in this example, yet to some extent the effects illustrated in the example play a role.

Figure 9 illustrates the second aggregation approach, one which uses a 'continuous trend'. The only difference compared to the previous approach is that here initial and final trends follow the slope for the period where proficiency statistics do exist. Put differently, each country has a completely linear trend from 2000 to 2019. In this example, the aggregate trend

is roughly a flat one. The clear decline in 'All % proficient' seen in Figure 8 is not visible in Figure 9.

Figure 9: Aggregation method with continuous trend



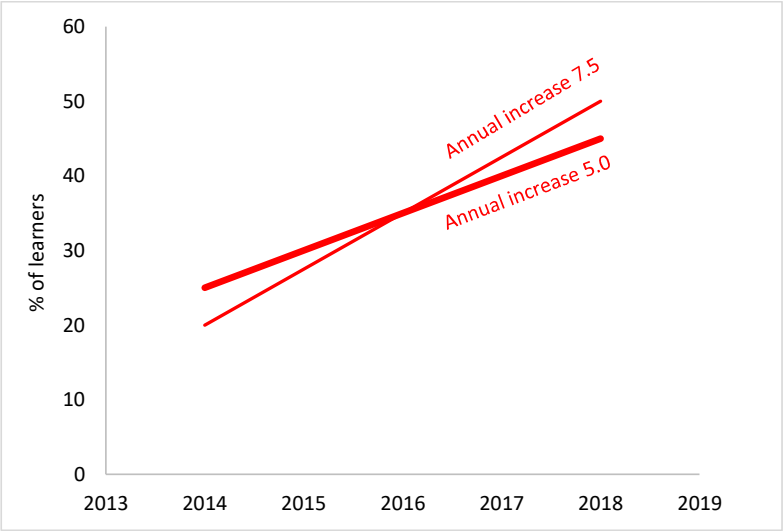
Which of the two approaches is likely to be most reliable? The answer to this question depends strongly on whether the timing of the monitoring of proficiency within countries is random or not. If the years when proficiency was measured is fairly random, then the continuous approach can be considered most reliable. To illustrate, there could be an international testing programme that tested in 2005, 2011 and 2017. There could moreover be a country that participated in just two consecutive assessments, either 2005 and 2011 or 2011 and 2017. If the timing of the country's assessments were not related to the likelihood of there being a proficiency improvement, it would not matter for the purposes of gauging trends when the country measured proficiency. However, it is very possible that countries will pay more attention to monitoring proficiency when there is simultaneously a strong focus on improving learning. This would make the timing of measurement far more non-random. Despite these reservations, the discussion that follows will consider slopes seen in the continuous approach as a somewhat more reliable reflection of reality than the restricted approach. The restricted approach has its own limitations. If a country monitored proficiency in, for instance, 2010 and 2015, and saw an improvement in this period, the assumption that outside this period there would be no change at all is unrealistic. The trends seen from the two approaches can be thought of as upper and lower bounds of what the true slope is.

A key feature of the UIS.Stat proficiency dataset used is that in the 11 cases where the data source is inconsistent over time (see the discussion in section 3), *the trend always consists of just two points*. This means that in aggregating trends, and if data source inconsistencies are to be excluded, no decisions need to be made around which point to exclude. Both points are deemed not able to produce trends. Had there been situations where, for instance, data sources X, Y, X and Y had been used sequentially across four different years, a decision would have to be taken around whether to calculate a trend on the basis of source X or source Y. As it turns out, such situations do not arise.

There are 384 multi-point trends available for the various aggregations, after 11 cases with inconsistent source data have been removed (see section 9). Of the 384, ten have gains or losses which exceed five percentage points a year. Figure 10 below illustrates what occurs in these cases. A trend with an annual increase of 7.5 a year would be made less steep, to the point at which the slope becomes 5.0, by pivoting the trendline on its midpoint. The assumption is thus made that, for instance, an annual increase of 7.5 probably represents a

mismeasurement, but that a considerable increase did in fact occur, though it is not as large as the reported increase.

Figure 10: Attenuating extreme changes



In the case of both methods described above, any annual proficiency statistic emerging as less than zero in the calculation, is adjusted to zero. Similarly, a ceiling of 100% applies.

10 Graphs with national, regional and global trends

The following eight composite graphs represent the seven world regions, and the world. The objective of these graphs is to provide a fairly detailed visualisation of the indicator 4.1.1 statistics. They include a few explicit statistics. These statistics, plus additional statistics, are presented in section 11 in tables.

Guidance to reading the composite graphs is as follows:

- The points in the graphs are countries with proficiency statistics, their diameter being proportional to the population aged 0 to 14 (the most recent statistic was used, meaning diameters do not change for a country over time). Grey points represent countries with no second point, meaning no trend is discernible.

- Lines represent *two-point* trends. The very few dotted lines would represent trends where different data sources feed into the two endpoints. Red lines, which are more common, represent trends which are steeper than would typically be considered credible – any annual change exceeding 5.0 percentage points, either positive or negative, is considered possibly problematic. Lines are coloured red where the slope for an entire series, which could consist of more than two points, and hence more than two line segments, exceeds a steepness of absolute 5.0. Red lines reflect the situation before the adjustments discussed in section 9 are applied. Red and dotted lines would only appear in the first six graphs of each figure, relating to indicator 4.1.1.

- Green lines reflect aggregate trends. In the case of the first six graphs of each figure, light green lines represent the 'continuous' method described in section 9, while dark green lines represent the 'restricted' approach. It is clear that the two approaches do not produce very different patterns in most graphs.

- The three statistics per graph in red text begin with a percentage of children with any point statistic. The second statistic is the percentage of children in countries with a usable trend which could be used for the aggregation. In other words, trends other than those with inconsistent sources apply. The third statistic is a population-weighted measure of the availability of trends over the years covered by the data. This range of years is mostly shorter than the full 2000 to 2019 scope. To illustrate, if the earliest statistics for end of primary reading in a region are from 2003, and the latest from 2018, then that range of 15 years is considered. If in Country A there are statistics for 2003 and 2015, that means Country A has a trend available for 12 years, or 80% of the range. If in Country B the earliest and most recent statistics are for 2015 and 2019, just four years, or 27% of the range, are covered. The population-weighted mean of the 80% and 27% would be found. Because different graphs within the figures have statistics spanning different ranges of years, the indicator values are indicative rather than strictly comparative.

- The participation statistics illustrated in the last two graphs of each figure all extend back to 2010 only. That is the range within the original UIS.Stat data currently accessible. The names of the two indicators within UIS.Stat are 'Out-of-school rate for children of primary school age' and 'Completion rate, primary education' (the latter of these two should be distinguished from the VIEW completion statistics discussed in section 13). The two indicators are packaged as part of SDG indicators 4.1.4 and 4.1.2 respectively. For the first of these two indicators, the difference between the indicator value and 100% was found, which would be the percentage of children who are *not* out of school. Only one kind of trendline is included in the participation graphs, a linear trendline in dark green. The slope of this trendline is the population- and duration-weighted mean of the various two-point slopes found in the data. The two participation graphs are included for two reasons. Firstly, they illustrate how low the availability of trend statistics is in the case of learning proficiency, relative to participation. Secondly, they point to improvements which are often steeper than the improvements seen with respect to proficiency, keeping in mind that the proficiency statistics refer to children and youths who are *in* school.

Figure 18 below depicts the global situation. What are the striking features? Globally, the availability of proficiency statistics improves the higher the schooling level. This is despite the fact that monitoring learning in the initial grades is widely seen as especially important, as this is where the foundations for educational success are laid. (The 'trend coverage' statistics are easiest to see in Table 8.) The green aggregate curves in Figure 18 do not point to the strong upward slopes that would be necessary if the SDG targets of 100% by 2030 were to be achieved. The trends are relatively flat. At the same time, the two participation graphs in Figure 18 display clear upward trends. This suggests that the percentage of *children in the population* who are proficient is likely to be increasing. Figure 18 allows for rough estimations of the magnitudes of these changes. Completion of primary schooling improved from 75% to 89% between 2010 and 2019 – this is what the aggregate curve in the completion graph in Figure 18 points to. Over this 2010 to 2019 period, reading proficiency at the end of primary remained at around 60%. Assuming that the additional children who completed primary schooling were from the most socio-economically disadvantaged segments of society, it can be concluded that the percentage of proficient children in the population improved from roughly 45% (60% × 75%) to 53% (60% × 89%), a substantial gain of eight percentage points in the population, almost a percentage point per year. This is explored further in section 13.

Figure 11: Trends in Sub-Saharan Africa

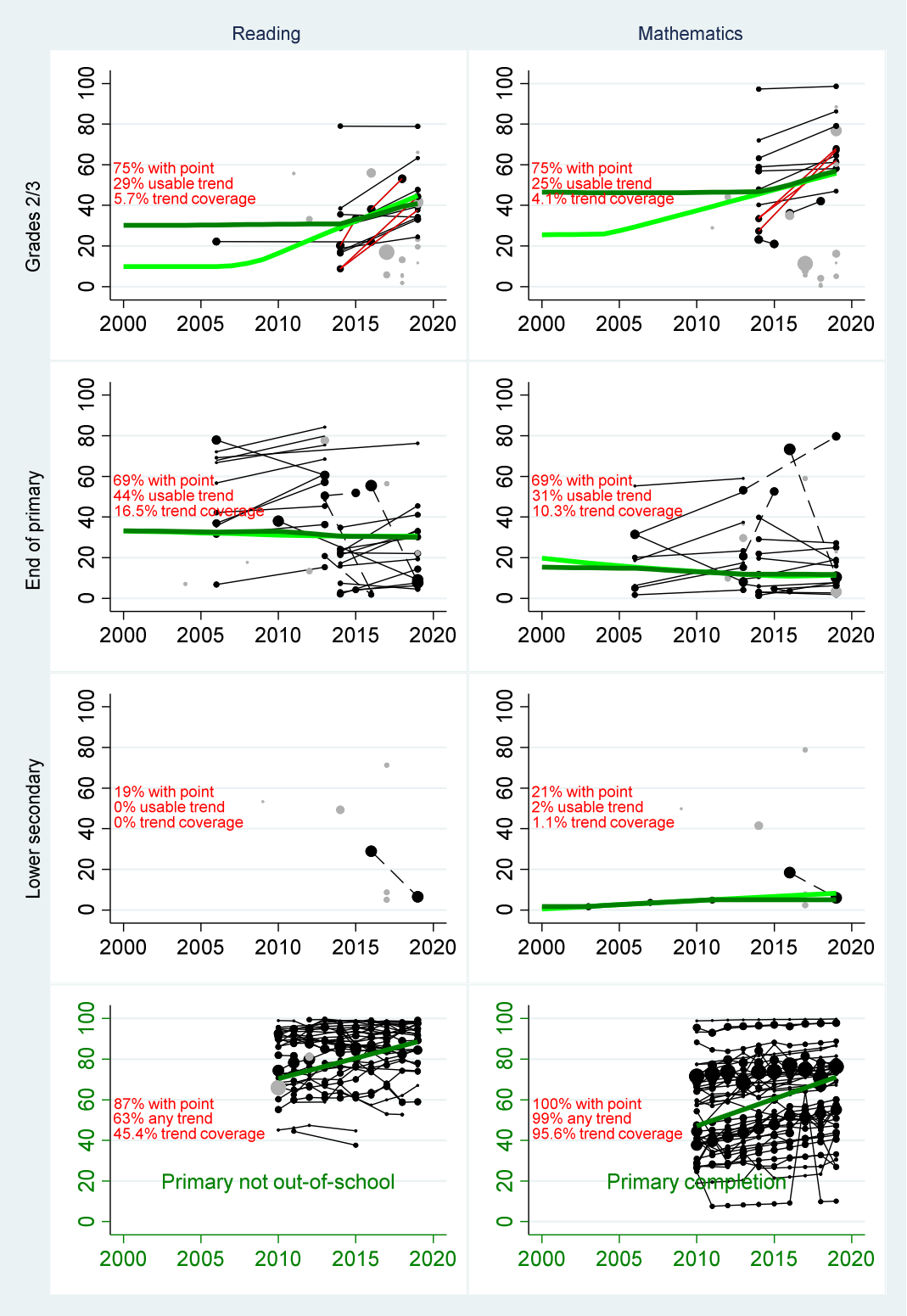


Figure 12: Trends in Northern Africa and Western Asia

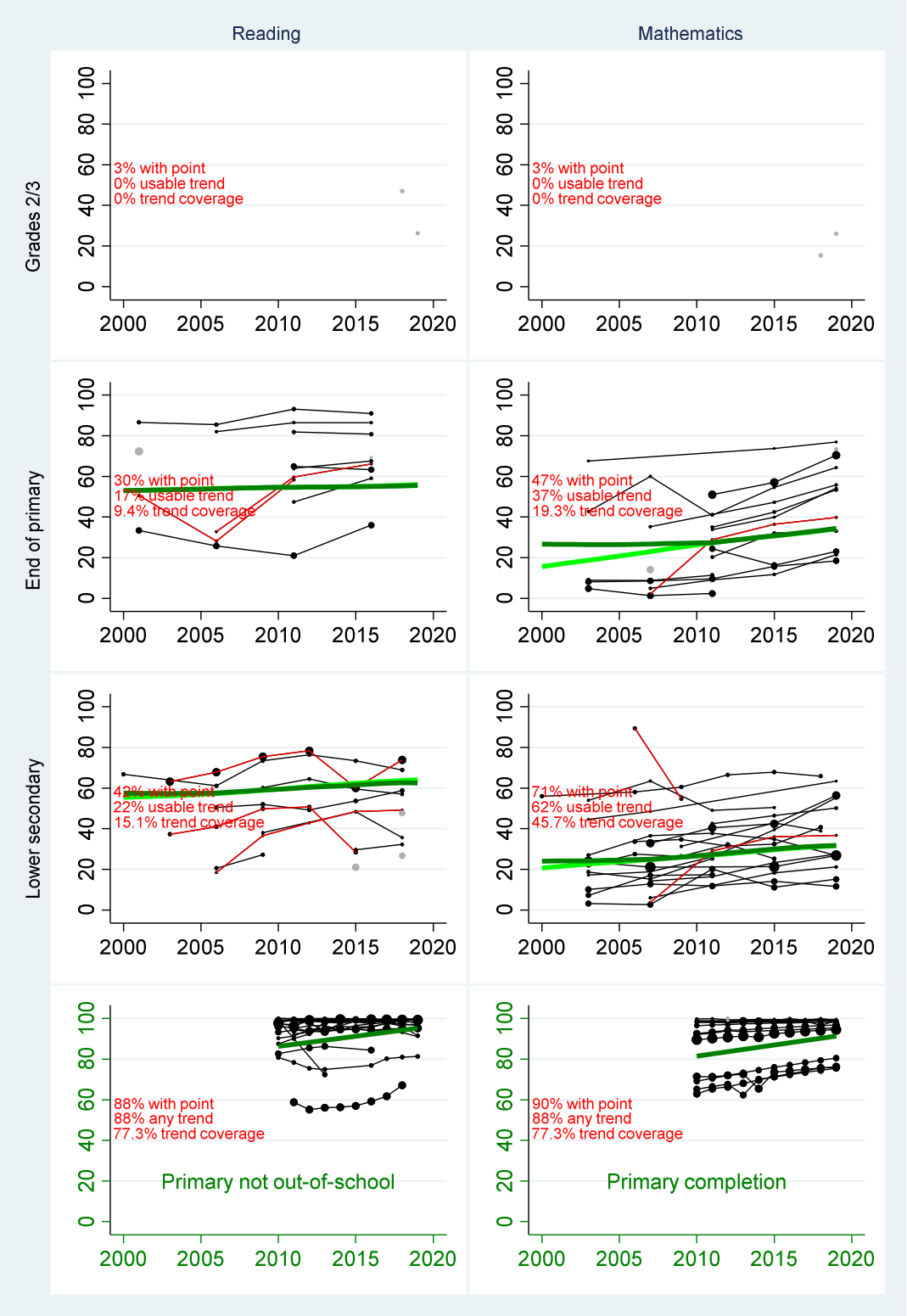


Figure 13: Trends in Central and Southern Asia

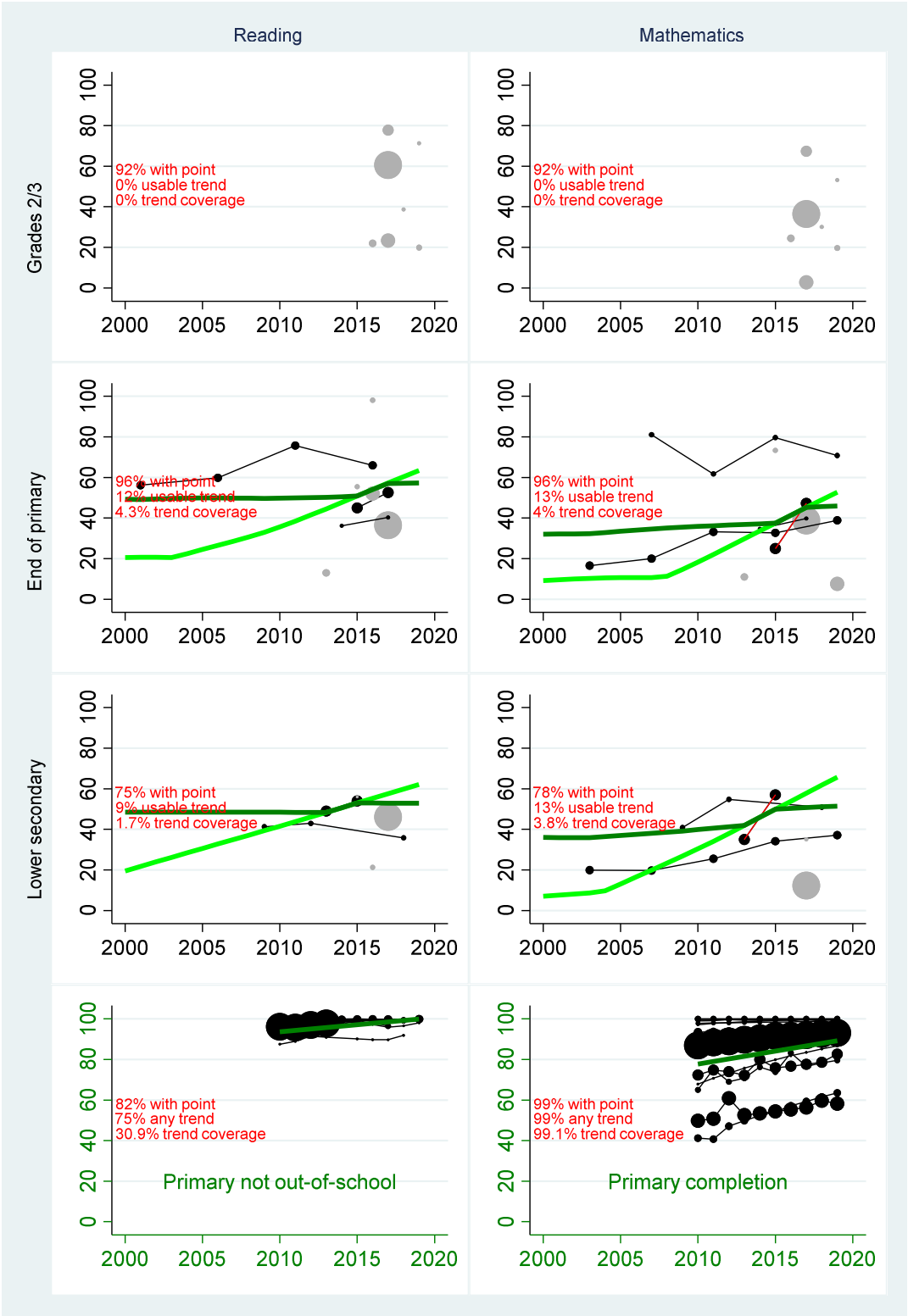


Figure 14: Trends in Eastern and South-eastern Asia

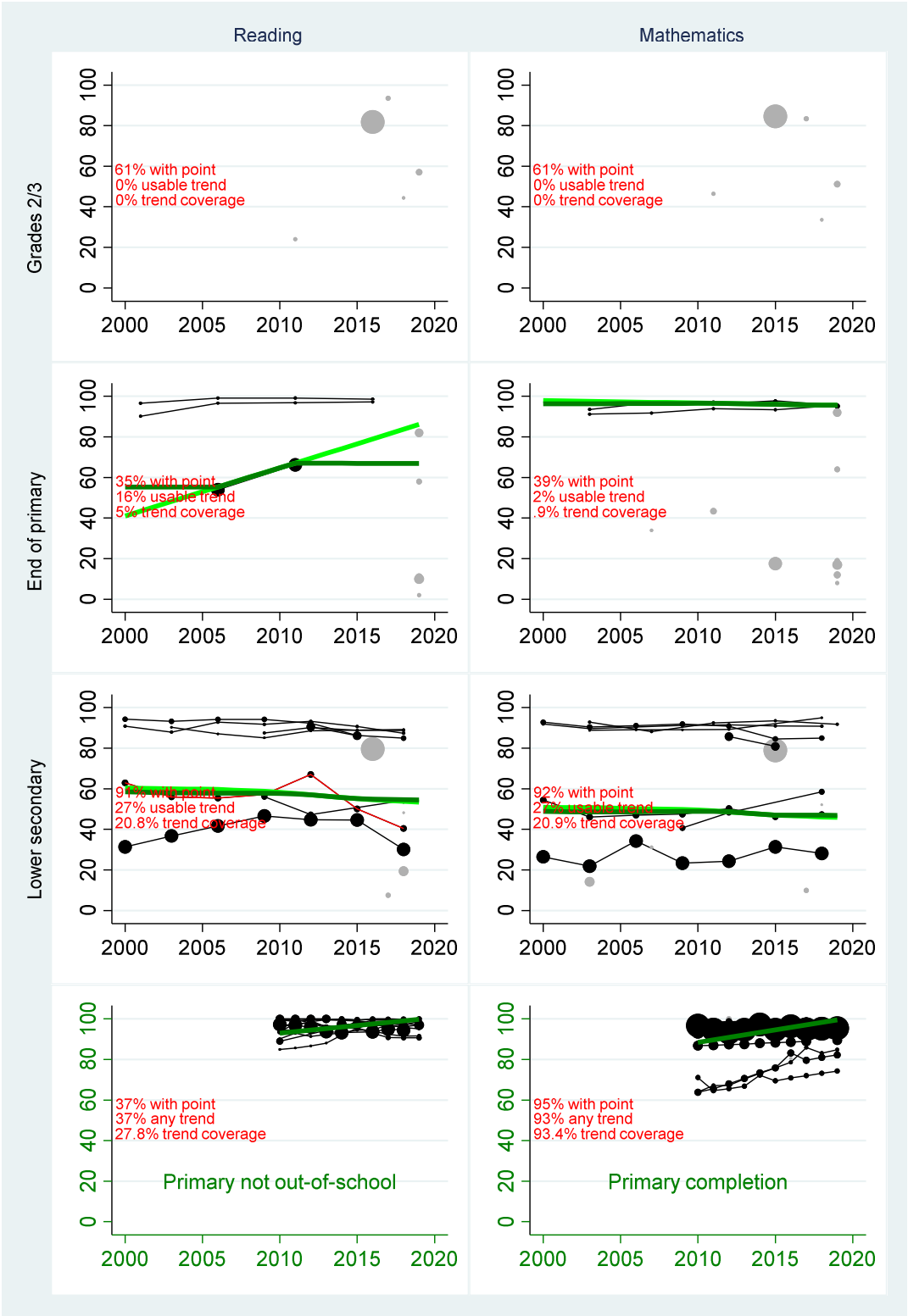


Figure 15: Trends in Oceania

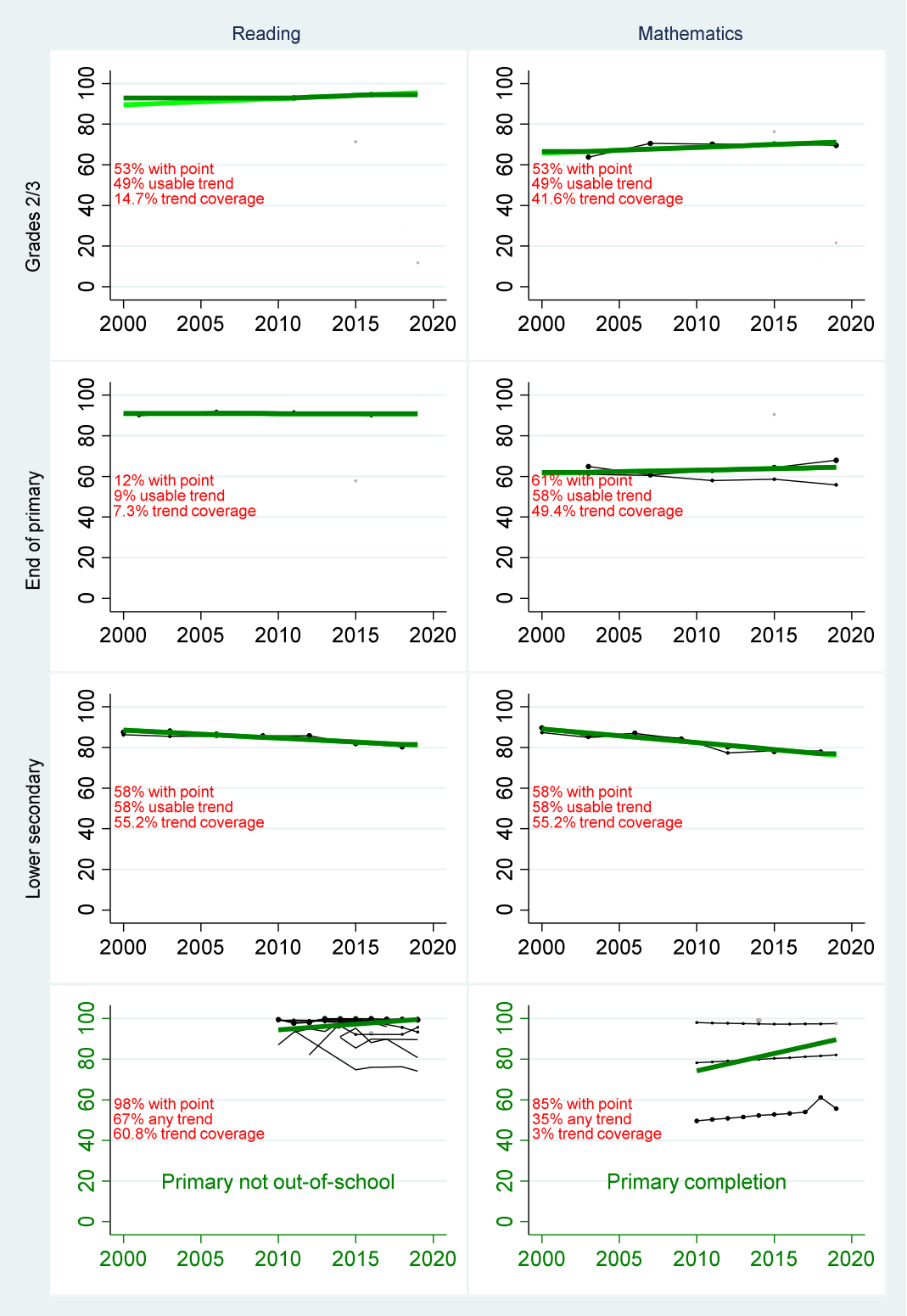


Figure 16: Trends in Latin America and the Caribbean

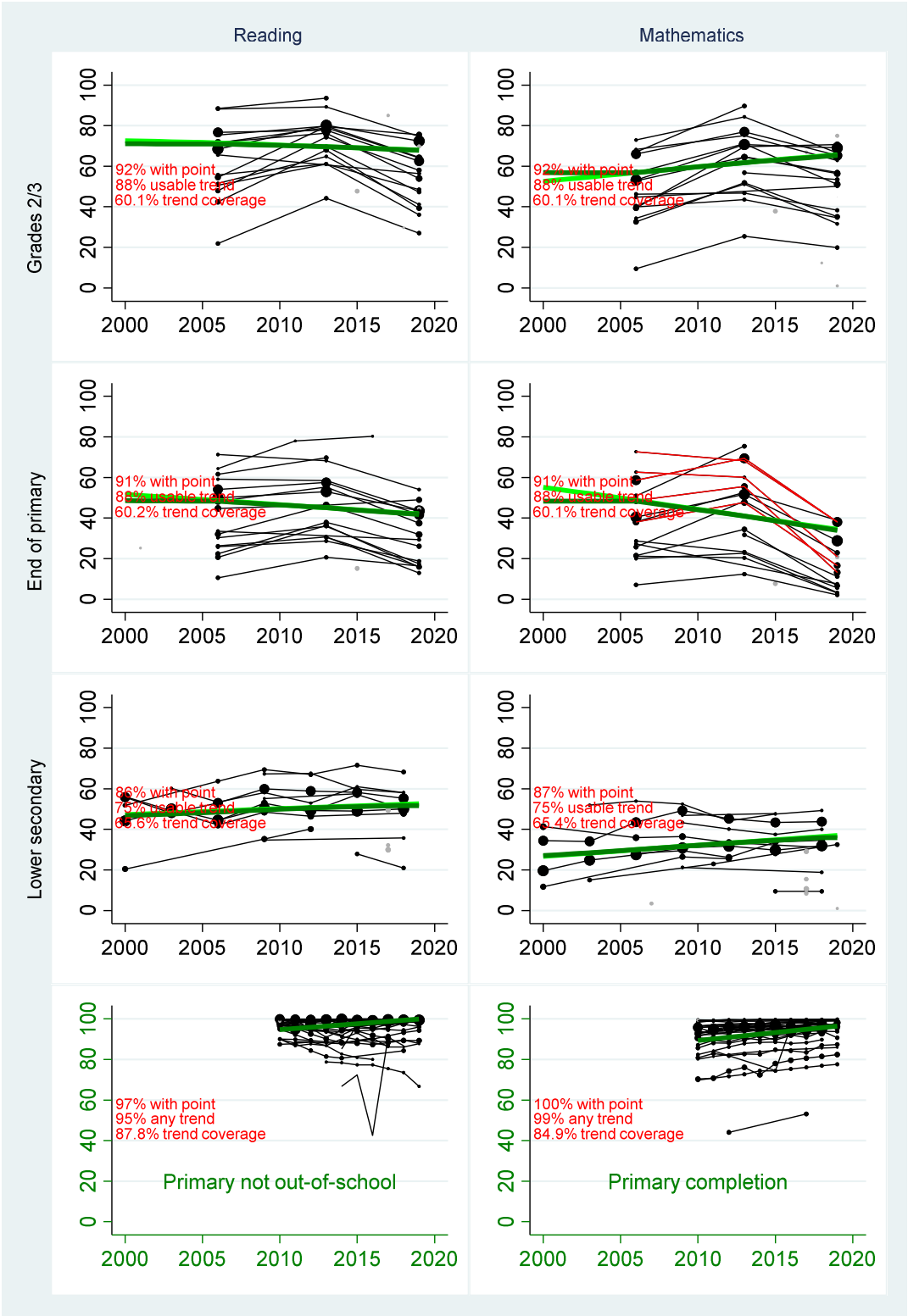


Figure 17: Trends in Europe and Northern America

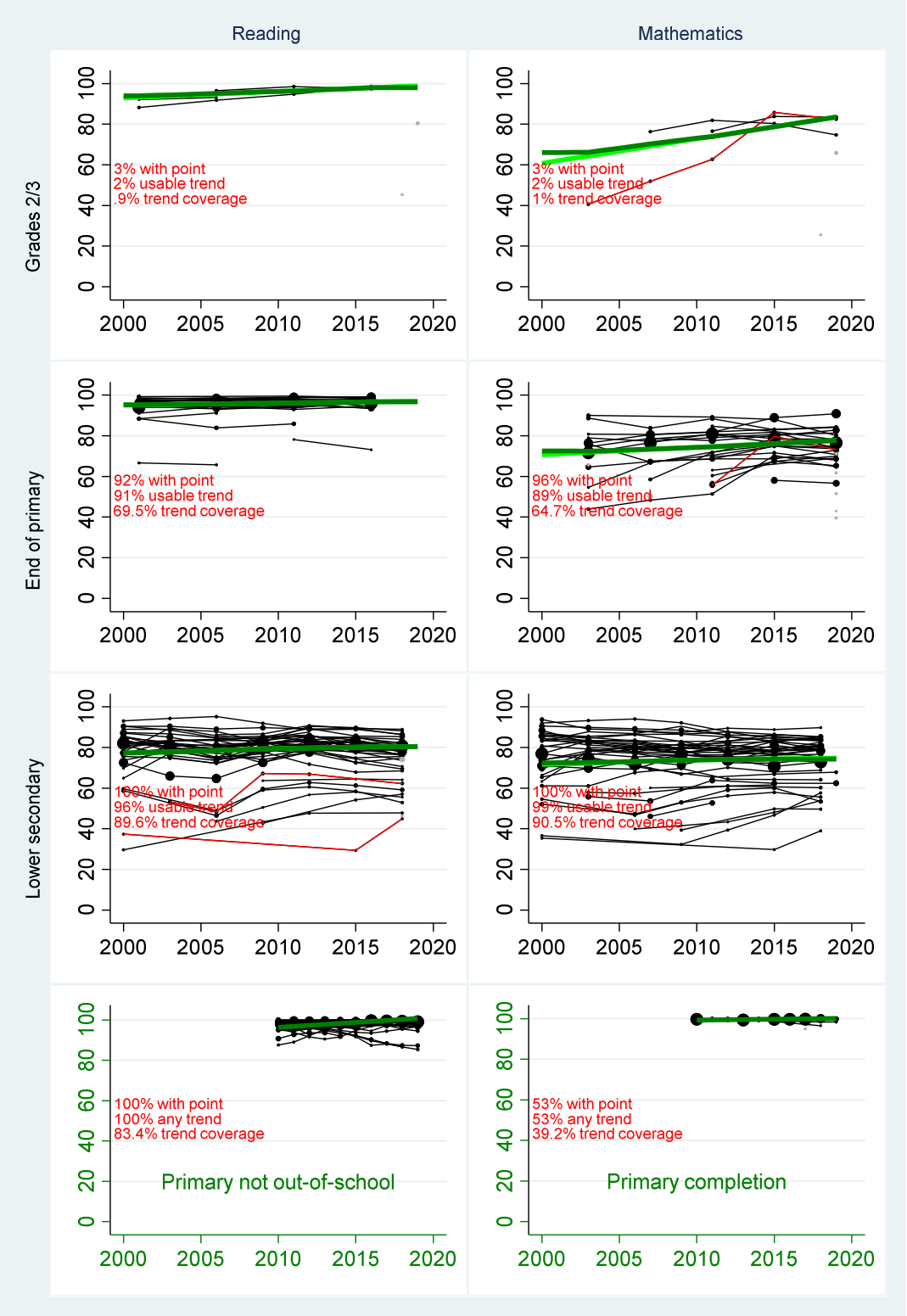
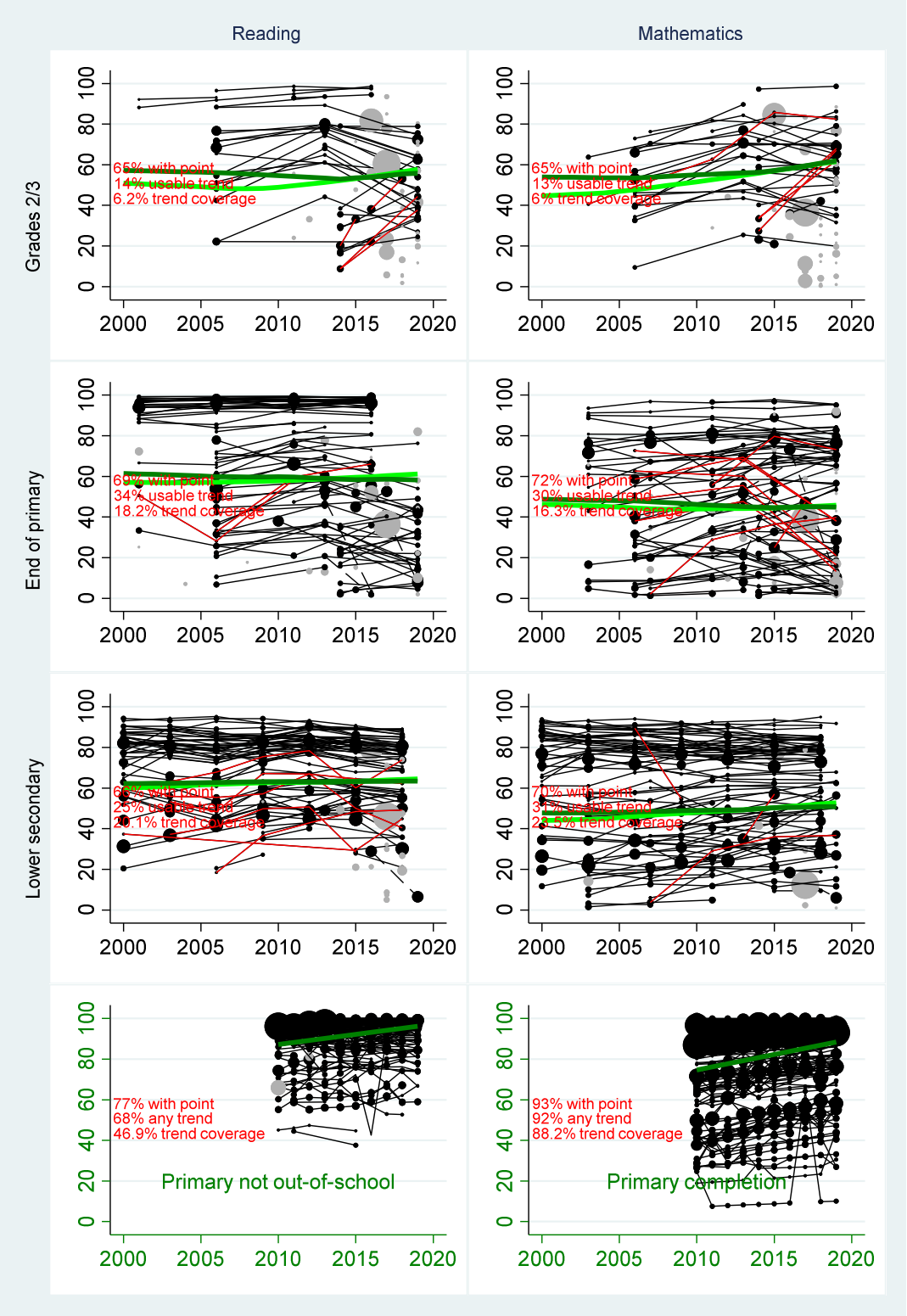


Figure 18: Global trends



11 Tables with regional and global trends

Table 7 below provides statistics per world region. Here there is no separation between the six SDG 4.1.1 indicators, in other words by level and subject.

The column '% of children covered by the largest quarter of countries' draws from the kind of analysis presented in section 7 above. According to the figures in this column, Oceania is the region with the greatest concentration of children in a few countries. The largest one quarter of countries, in terms of child populations, account for 95% of all children in the region. More specifically, Australia on its own accounts for 48% of the region's children. Even in the region with the lowest country-specific concentrations, Northern Africa and Western Asia, the largest quarter of countries account for 71% of all children.

The range in years for which 4.1.1 proficiency statistics are available in UIS.Stat is similar across regions.

Oceania is the region with the lowest percentage of children in countries with at least one proficiency statistic: the statistic for the region is 62%. This is largely because Papua New Guinea, which accounts for 31% of the region's children, has no 2000 to 2019 proficiency statistics in UIS.Stat. However, with respect to the availability of any trend data, regardless of consistency problems, Central and Southern Asia is the region that fares worst – here only 13% of children are in countries with some trend. The final column of the table indicates that virtually all the world's children are in countries where participation in schooling is measured.

The next table, Table 8, presents statistics per region, per 4.1.1 indicator, and with respect to the two participation indicators discussed in section 10. The three columns whose values are represented within the graphs of section 10 are the following: % of population with any data point; % of population with a usable trend; Trend coverage (%). The column '% of population with a usable trend' is the result of removing trends with inconsistent data sources across two years (see the discussion in section 9).

The year range in Table 8 is the range of all statistics for the relevant SDG 4.1.1 indicator, whether they produce a trend or not.

The calculation of the 'Trend coverage' values was described in section 10. These values provide a sense of how extensive the reporting on trends is within the year range of the first two columns of Table 8. The 'Average level within year range' is the average vertical height of the dark green trendlines in the section 10 graphs. This means the restricted approach is used.

The two 'Annual percentage point gain' columns at the end of Table 8 present the slopes of the light green and dark green aggregate curves respectively of section 10. The final annual gain statistic is the more conservative one, as it assumes that outside the range of years for which a country has statistics, the trend is flat.

The higher the value in the column '% of population with a usable trend', the more the gain statistics of the last two column are likely to be reliable. It is noteworthy that of the 42 rows of Table 8 representing the seven regions and six SDG 4.1.1 statistics, just three have a percentage of the population with a reliable trend exceeding 50% *and* have a trend that is greater than positive 0.50, using the continuous approach of the second-last column. An annual gain that is over positive 0.50 translates to one percentage point gain every two years. Two of these three relatively reliable and steep positive trends are in the region Latin America and the Caribbean. To illustrate, in this region 88% of children are in countries with grades 2 to 3 mathematics trends, and the aggregate slope is positive 0.69, using the continuous approach. The upward trend can be seen in Figure 16 above. Compared to the three relatively

reliable and steep *positive* trends, there are three relatively reliable and steep *negative* trends, using similar criteria. Two of these are in Latin America and the Caribbean (both at the end of primary), and one in Oceania. The contradictory patterns in Latin America and the Caribbean could point to reliability problems with the statistics, even if a relatively high percentage of children are covered. This seems to be confirmed by the first four graphs in Figure 16. It is unlikely that so many countries in the region would experience the up and then down patterns demonstrated by available statistics.

Given how important it is to establish the foundations for learning through reading in the early grades, what stands out is the low level of monitoring at this level, and for reading. It is among the least monitored of the six 4.1.1 indicators. Three regions – Northern Africa and Western Asia; Eastern and South-eastern Asia; and Central and Southern Asia – display no trend data for reading in the early grades.

In Table 8 some of the annual gain statistics are close to, or higher than 1.0, using the continuous approach, though this is never found in combination with a population with a usable trend exceeding 50%. Yet such gains are noteworthy, especially when they occur both in reading and mathematics. This is true for the lower primary gains in Sub-Saharan Africa and end of primary gains in Central and Southern Asia – in both these regions and in both subjects annual 'continuous' gains of around 2.0 are seen. These statistics are based on around 30% of the child population in Sub-Saharan Africa and around 12% in Central and Southern Asia[25].

The annual percentage point gains for reading at the end of primary in the Sub-Saharan Africa region are negative in Table 8, despite substantial gains in this area being reported elsewhere[26]. The Table 8 patterns for this region, subject and level are largely explained by sharp downward trends in two large countries, Democratic Republic of the Congo (2010 to 2019) and United Republic of Tanzania (2006 to 2013). If these two countries are excluded, the slope for end of primary reading in Sub-Saharan Africa becomes 0.97, instead of the negative 0.15 seen in Table 8 for the continuous approach.

[25] To illustrate, the mathematics trend is based on three countries in the Central and Southern Asia region: Bangladesh; Iran (Islamic Republic of); and Kazakhstan.
[26] UIS, 2021a.

Table 7: Overall breakdown by region

	Countries	Population aged 0 to 14 (thous.)	% of children covered by the largest quarter of countries	Year range for 4.1.1 statistics		% of population with any 4.1.1 data point	% of population with any 4.1.1 trend	% of population with any participation trend
Sub-Saharan Africa	52	478,492	73	2003	2019	93	67	100
Northern Africa and Western Asia	24	162,229	71	2000	2019	78	69	98
Central and Southern Asia	14	553,764	90	2001	2019	98	13	100
Eastern and South-eastern Asia	18	450,430	88	2000	2019	96	27	95
Oceania	21	10,261	95	2000	2019	62	58	98
Latin America and the Caribbean	48	155,744	90	2000	2019	92	88	100
Europe and Northern America	51	185,918	88	2000	2019	100	99	100
World	228	1,996,838	89	2000	2019	94	48	99

Table 8: Region breakdown with details by statistic

	Year range		% of population with any data point	% of population with any trend	% of population with a usable trend	Trend coverage (%)	Average level within year range	Annual percentage point gain (continuous approach)	Annual percentage point gain (restricted approach)
Sub-Saharan Africa									
Grades 2/3 reading	2006	2019	75	29	29	5.7	32	1.94	0.43
Grades 2/3 mathematics	2011	2019	75	25	25	4.1	48	1.75	0.39
End of primary reading	2004	2019	69	60	44	16.5	32	-0.15	-0.18
End of primary mathematics	2006	2019	69	52	31	10.3	13	-0.47	-0.24
Lower secondary reading	2009	2019	19	10	0	0.0	-	-	-
Lower secondary mathematics	2003	2019	21	12	2	1.1	4	0.41	0.22
Primary aged not out-of-school	2010	2019	87	63		45.4	79	2.0	
Primary completion	2010	2019	100	99		95.6	59	2.7	
Northern Africa and Western Asia									
Grades 2/3 reading	2018	2019	3	0	0	0.0	-	-	-
Grades 2/3 mathematics	2018	2019	3	0	0	0.0	-	-	-
End of primary reading	2001	2016	30	17	17	9.4	54	0.14	0.13
End of primary mathematics	2003	2019	47	37	37	19.3	29	0.98	0.40
Lower secondary reading	2000	2018	42	22	22	15.1	60	0.49	0.33
Lower secondary mathematics	2000	2019	71	62	62	45.7	27	0.58	0.46
Primary aged not out-of-school	2010	2019	88	88		77.3	91	1.0	
Primary completion	2010	2019	90	88		77.3	86	1.1	
Central and Southern Asia									
Grades 2/3 reading	2016	2019	92	0	0	0.0	-	-	-
Grades 2/3 mathematics	2016	2019	92	0	0	0.0	-	-	-
End of primary reading	2001	2017	96	12	12	4.3	51	2.42	0.36
End of primary mathematics	2003	2019	96	13	13	4.0	37	2.39	0.70
Lower secondary reading	2009	2018	75	9	9	1.7	50	2.24	0.27
Lower secondary mathematics	2003	2019	78	13	13	3.8	42	3.32	0.94
Primary aged not out-of-school	2010	2019	82	75		30.9	97	0.7	
Primary completion	2010	2019	99	99		99.1	83	1.3	
Eastern and South-eastern Asia									
Grades 2/3 reading	2011	2019	61	0	0	0.0	-	-	-
Grades 2/3 mathematics	2011	2019	61	0	0	0.0	-	-	-
End of primary reading	2001	2019	35	16	16	5.0	62	2.37	0.85
End of primary mathematics	2003	2019	39	2	2	0.9	96	-0.12	-0.03
Lower secondary reading	2000	2018	91	27	27	20.8	57	-0.40	-0.22

	Year range		% of population with any data point	% of population with any trend	% of population with a usable trend	Trend coverage (%)	Average level within year range	Annual percentage point gain (continuous approach)	Annual percentage point gain (restricted approach)
Lower secondary mathematics	2000	2019	92	27	27	20.9	48	-0.29	-0.11
Primary aged not out-of-school	2010	2019	37	37		27.8	96	0.8	
Primary completion	2010	2019	95	93		93.4	94	1.2	
Oceania									
Grades 2/3 reading	2011	2019	53	49	49	14.7	93	0.31	0.10
Grades 2/3 mathematics	2003	2019	53	49	49	41.6	69	0.28	0.26
End of primary reading	2001	2016	12	9	9	7.3	91	-0.01	-0.01
End of primary mathematics	2003	2019	61	58	58	49.4	63	0.15	0.14
Lower secondary reading	2000	2018	58	58	58	55.2	85	-0.39	-0.38
Lower secondary mathematics	2000	2018	58	58	58	55.2	83	-0.67	-0.67
Primary aged not out-of-school	2010	2019	98	67		60.8	97	0.6	
Primary completion	2010	2019	85	35		35.0	82	1.7	
Latin America and the Caribbean									
Grades 2/3 reading	2006	2019	92	88	88	60.1	70	-0.25	-0.17
Grades 2/3 mathematics	2006	2019	92	88	88	60.1	60	0.69	0.51
End of primary reading	2001	2019	91	88	88	60.2	46	-0.51	-0.38
End of primary mathematics	2006	2019	91	88	88	60.1	43	-1.10	-0.82
Lower secondary reading	2000	2018	86	75	75	65.6	50	0.29	0.28
Lower secondary mathematics	2000	2019	87	75	75	65.4	32	0.54	0.49
Primary aged not out-of-school	2010	2019	97	95		87.8	97	0.6	
Primary completion	2010	2019	100	99		84.9	93	0.8	
Europe and Northern America									
Grades 2/3 reading	2001	2019	3	2	2	0.9	96	0.31	0.24
Grades 2/3 mathematics	2003	2019	3	2	2	1.0	73	1.20	0.98
End of primary reading	2001	2016	92	91	91	69.5	96	0.10	0.09
End of primary mathematics	2003	2019	96	89	89	64.7	75	0.37	0.29
Lower secondary reading	2000	2018	100	96	96	89.6	79	0.18	0.18
Lower secondary mathematics	2000	2019	100	99	99	90.5	74	0.17	0.11
Primary aged not out-of-school	2010	2019	100	100		83.4	99	0.5	
Primary completion	2010	2019	53	53		39.2	100	0.1	
World									
Grades 2/3 reading	2001	2019	65	14	14	6.2	55	0.35	-0.17
Grades 2/3 mathematics	2003	2019	65	13	13	6.0	56	0.93	0.38
End of primary reading	2001	2019	69	38	34	18.2	60	0.23	-0.16

	Year range		% of population with any data point	% of population with any trend	% of population with a usable trend	Trend coverage (%)	Average level within year range	Annual percentage point gain (continuous approach)	Annual percentage point gain (restricted approach)
End of primary mathematics	2003	2019	72	35	30	16.3	47	-0.04	-0.25
Lower secondary reading	2000	2019	66	28	25	20.1	63	0.22	0.06
Lower secondary mathematics	2000	2019	70	33	31	23.5	49	0.47	0.18
Primary aged not out-of-school	2010	2019	77	68		46.9	92	1.0	
Primary completion	2010	2019	93	92		88.2	82	1.5	

12 Gender-specific aggregations

To come.

13 Combining proficiency and participation statistics

In order to explore trends in the percentage of *children in the population* who are proficient, a newly produced dataset of school completion of by UNESCO was used. This dataset, which is a product of imputations and corrections explained in Dharamshi *et al* (2021), has annual primary completion values over around thirty years for 164 countries, or 98.4% of the global child population – substantial missing countries are Saudi Arabia, Sri Lanka and Democratic People's Republic of Korea. The statistics for all countries begin in 1990, and for 90% of children they reach at least 2019 (for 97% they reach at least 2015). The dataset was retrieved off the VIEW[27] portal of UNESCO[28]. The primary completion rates in the data reflect completion in the population aged three to five years above the official end-of-primary age.

The historical patterns for 2000 to 2019 seen in this dataset are shown in Figure 19 below. Trends are weighted by the annual child population aged 10 to 14. The global increase during 2010 to 2019, from 82% to 86%, is less steep than what was seen in earlier Figure 18. Yet the changes are substantial, a key advantage with the VIEW statistics being that they extend before 2010. The Figure 19 improvement for the entire 2000 to 2019 period is from 77% to 86% globally.

Figure 19: Completion of primary schooling

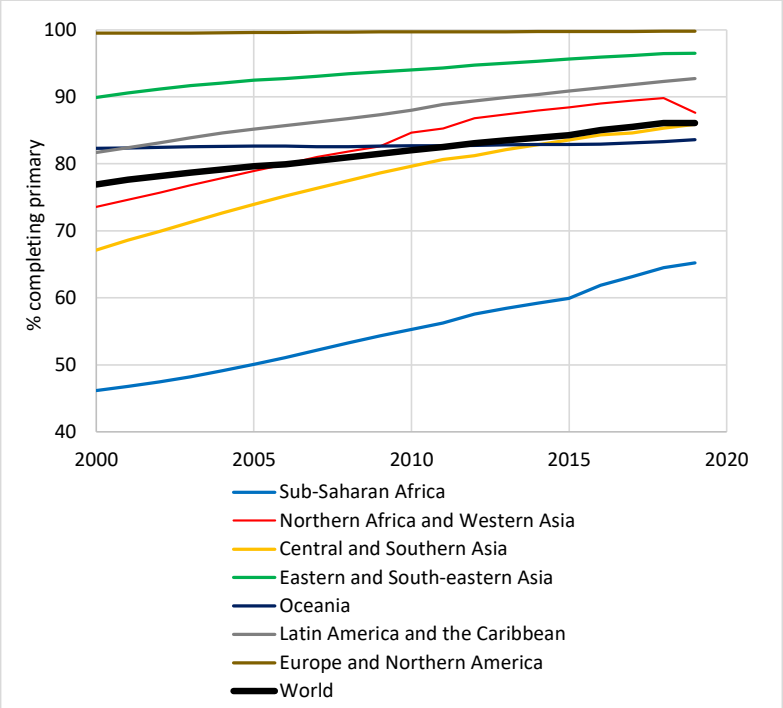


Table 9 below is meant to be compared to the end of primary values of earlier Table 8. For Table 9, proficiency statistics used were just those where there was a VIEW completion statistic for the same country and year. This explains why, for instance, 32% of the world's

[27] Visualizing Indicators of Education for the World.
[28] Retrieved November 2022, https://education-estimates.org. Just the primary completion rates (the 'CR Indicator' version) were retrieved.

children had usable data for end of primary reading in Table 9, against 34% in Table 8. Within each country with the required trend data, the level of proficiency *in the child population* was calculated. For instance, if 40% of enrolled children were proficient, and 80% of children completed primary schooling, then 32% of the population (40% times 80%) were assumed to be proficient. The assumption of UIS (2017b) was thus used that children who do not complete primary schooling do not display the learning proficiency expected of children who complete primary. A region such as Sub-Saharan Africa, with relatively low levels of primary completion, even in more recent years, would display a large gap between the enrolment- and population-based levels of proficiency: for reading the two statistics would be 32% (Table 8) and 23% (Table 9).

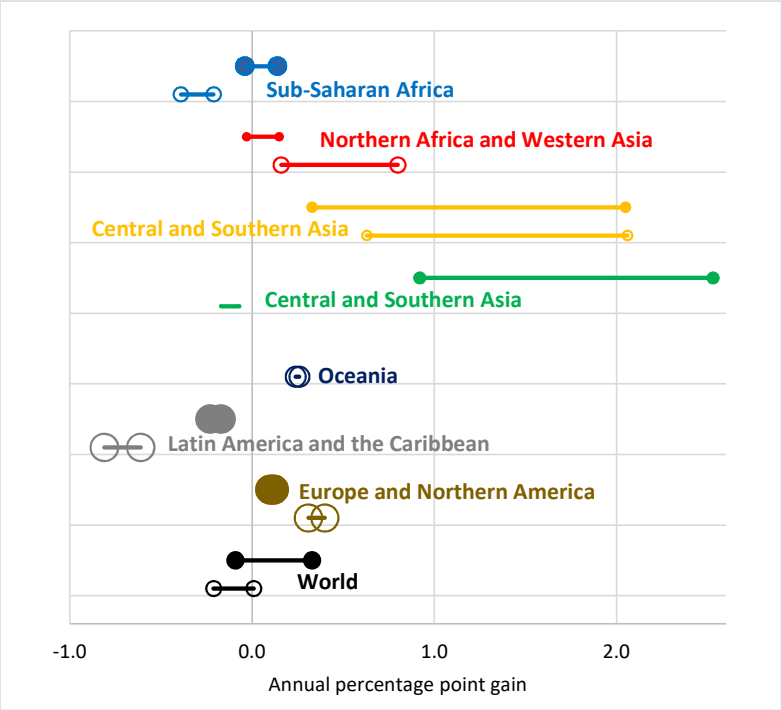
The percentage point gain statistics shown in Table 9 refer to the percentage of an age cohort of the child population, not the percentage of enrolled children. Thus, for instance, a percentage point gain in Sub-Saharan Africa in Table 9 would refer to more children than a percentage point gain in Table 8. What does the 0.33 percentage point annual gain for reading and the world, using the 'continuous approach', mean? Roughly, it means that the percentage of proficient children increased by around seven percentage points over the twenty-year period, from 51% to 58% (the average in Table 9 is 55%). The 0.33 is substantially less than the 'almost a percentage point per year' put forward speculatively at the end of section 10. This can be entirely explained by the less steep completion improvements seen in Figure 19 compared to Figure 18. It can be assumed that the slower improvements in completion seen in Figure 19 are more reliable as they are the result of a detailed exercise to impute and correct statistics. The decline at end right-hand end of the Northern Africa and Western Asia trend seen in Figure 19 is due to a major country with high primary completion, Turkey, not having a 2019 value. Completion statistics behind Figure 19 have values for Syrian Arab Republic only up to 2011, meaning school participation disruptions due to the war in that country would not be reflected. In the case of Syrian Arab Republic, there are no SDG proficiency trends beyond 2011, meaning the absence of completion statistics would have no effect on Table 9. However, the absence of 2019 completion statistics for Turkey would flatten the region's end of primary and lower secondary trends a little as Turkey's improvements in grades 4 and 8 between 2015 and 2019 in TIMSS would not be used for Table 9 (though these improvements would influence the statistics in Table 8).

Table 9: Population-focussed region breakdown with end of primary details

	Year range		% of population with a usable trend	Average level within year range	Annual percentage point gain (continuous approach)	Annual percentage point gain (restricted approach)
Sub-Saharan Africa						
End of primary reading	2004	2019	39	23	0.14	-0.04
End of primary mathematics	2006	2019	26	8	-0.39	-0.21
Northern Africa and Western Asia						
End of primary reading	2001	2016	8	32	-0.03	0.15
End of primary mathematics	2003	2019	29	27	0.80	0.16
Central and Southern Asia						
End of primary reading	2001	2017	12	43	2.05	0.33
End of primary mathematics	2003	2019	13	31	2.06	0.63
Eastern and South-eastern Asia						
End of primary reading	2006	2019	16	57	2.53	0.92
End of primary mathematics	2007	2019	1	97	-0.17	-0.07
Oceania						
End of primary reading	-	-	-	-	-	-
End of primary mathematics	2003	2019	49	63	0.26	0.24
Latin America and the Caribbean						
End of primary reading	2001	2019	87	42	-0.23	-0.17
End of primary mathematics	2006	2019	87	40	-0.81	-0.61
Europe and Northern America						
End of primary reading	2001	2016	91	96	0.12	0.10
End of primary mathematics	2003	2019	86	74	0.40	0.31
World						
End of primary reading	2001	2019	32	55	0.33	-0.09
End of primary mathematics	2003	2019	28	45	0.01	-0.21

Figure 20 below draws largely from the last two columns of Table 9. The predominance of slopes greater than zero at the region level is clear. Only one region has both continuous slope statistics below zero: Latin America and the Caribbean. However, the reliability of trends in this region may not be reliable, as discussed in section 11.

Figure 20: Population-focussed region breakdown with end of primary details



Note: Solid markers reflect reading, markers with no fill reflect mathematics. The lower (left-hand) marker uses the 'restricted approach' while the higher (right-hand) marker uses the 'continuous approach'. Marker areas are proportional to '% of population with a usable trend'.

14 The way forward for regional and global monitoring of proficiency

To come.

References

CONFEMEN (2020). *PASEC 2019: Qualité des systèmes éducatifs en Afrique subsaharienne francophone*. Dakar.

Dharamshi, A., Baraket, B., Alkema, L. & Antoninis, M. (2021). *Adjusted Bayesian completion rates (ABC) estimation*. Toronto: University of Toronto.

Moscoviz, L. & Evans, D.K. (2022). *Learning loss and student dropouts during the COVID-19 pandemic: A review of the evidence two years after schools shut down*. Washington: Center for Global Development.

UIS (2017a). *Principles of Good Practice in Learning Assessment*. Montreal.

UIS (2017b). *Counting the number of children not learning: Methodology for a global composite indicator for education*. Montreal.

UIS (2018). *Costs and benefits of different approaches to measuring the learning proficiency of students (SDG Indicator 4.1.1)*. Montreal.

UIS (2019). *How fast can levels of proficiency improve? Examining historical trends to inform SDG 4.1.1 scenarios*. Montreal.

UIS (2020). *Evidence-based projections and benchmarks for SDG Indicator 4.1.1*. Montreal.

UIS (2021a). *Assessing learning proficiency levels and trends for Sustainable Development Goal 4.1: A focus on Africa.* Montreal.

UIS (2021b). *Pandemic-related disruptions to schooling and impacts on learning proficiency indicators: A focus on the early grades*. Montreal.

UNESCO (2005). *Education for All global monitoring report 2005: The quality imperative*. Paris.

UNESCO (2017). *Revised medium-term strategy 2017-2021 of the UNESCO Institute for Statistics (UIS)*. Montreal.

World Bank (2019). *Ending learning poverty: What will it take?* Washington.

World Bank (2022). *The state of global learning poverty: 2022 update*. Washington.

Appendix 1: Trends by World Bank country income category

The following three tables provide breakdowns by World Bank country income categories, as calculated by the World Bank for 2022. The methods are those explained in earlier sections 11 and 13.

Table 10: Overall breakdown by World Bank income group

	Countries	Population aged 0 to 14 (thous.)	% of children covered by the largest quarter of countries	Year range for 4.1.1 statistics		% of population with any 4.1.1 data point	% of population with any 4.1.1 trend	% of population with any participation trend
Low income countries	28	296,867	59	2003	2019	89	72	98
Lower middle income countries	54	986,780	86	2000	2019	96	34	100
Upper middle income countries	53	505,535	92	2000	2019	96	45	99
High income countries	77	199,311	88	2000	2019	92	91	92

Table 11: World Bank income group breakdown with details by statistic

	Year range		% of population with any data point	% of population with any trend	% of population with a usable trend	Trend coverage (%)	Average level within year range	Annual percentage point gain (continuous approach)	Annual percentage point gain (restricted approach)
Low income countries									
Grades 2/3 reading	2012	2019	54	21	21	2.7	33	1.84	0.40
Grades 2/3 mathematics	2012	2019	54	21	21	3.4	46	0.62	0.34
End of primary reading	2006	2019	78	66	40	15.3	20	-0.68	-0.35
End of primary mathematics	2003	2019	82	55	29	9.9	9	0.02	0.08
Lower secondary reading	2014	2019	28	16	0	0.0			
Lower secondary mathematics	2003	2019	30	18	2	0.9	15	1.26	0.67
Primary aged not out-of-school	2010	2019	73	72		50.8	79	2.3	
Primary completion	2010	2019	98	98		92.1	50	2.7	
Lower middle income countries									
Grades 2/3 reading	2006	2019	72	7	7	1.4	38	1.92	0.25
Grades 2/3 mathematics	2006	2019	72	7	7	1.3	50	2.32	0.40
End of primary reading	2001	2019	78	23	23	8.3	51	1.55	0.32
End of primary mathematics	2003	2019	80	18	16	5.8	25	0.89	0.10
Lower secondary reading	2000	2018	59	14	14	8.2	52	0.61	0.04
Lower secondary mathematics	2000	2019	67	23	23	14.2	37	0.99	0.25
Primary aged not out-of-school	2010	2019	91	75		43.4	92	0.9	
Primary completion	2010	2019	99	99		98.9	82	1.5	
Upper middle income countries									
Grades 2/3 reading	2006	2019	82	28	28	19.1	65	-0.20	-0.17
Grades 2/3 mathematics	2006	2019	82	25	25	17.3	60	0.69	0.52
End of primary reading	2001	2019	42	35	35	23.6	55	-0.19	-0.25
End of primary mathematics	2003	2019	45	36	36	24.8	51	-0.52	-0.48
Lower secondary reading	2000	2018	91	38	38	32.2	55	0.18	0.15
Lower secondary mathematics	2000	2019	91	38	38	31.6	42	0.39	0.31
Primary aged not out-of-school	2010	2019	44	44		39.3	97	0.6	
Primary completion	2010	2019	99	99		90.2	95	1.1	
High income countries									
Grades 2/3 reading	2001	2019	7	7	7	2.9	89	0.29	0.12
Grades 2/3 mathematics	2003	2019	7	7	7	4.5	70	1.19	0.72
End of primary reading	2001	2019	84	82	82	59.0	92	0.09	0.06
End of primary mathematics	2003	2019	87	85	85	57.6	71	0.26	0.15
Lower secondary reading	2000	2018	90	86	86	79.6	81	-0.02	-0.02

	Year range		% of population with any data point	% of population with any trend	% of population with a usable trend	Trend coverage (%)	Average level within year range	Annual percentage point gain (continuous approach)	Annual percentage point gain (restricted approach)
Lower secondary mathematics	2000	2019	92	91	91	84.2	72	-0.08	-0.09
Primary aged not out-of-school	2010	2019	92	92		77.1	99	0.5	
Primary completion	2010	2019	41	34		26.1	99	0.1	

Table 12: Population-focussed World Bank income group breakdown with end of primary details

	Year range		% of population with a usable trend	Average level within year range	Annual percentage point gain (continuous approach)	Annual percentage point gain (restricted approach)
Low income countries						
End of primary reading	2006	2019	32	9	-0.35	-0.16
End of primary mathematics	2003	2019	21	4	-0.02	0.03
Lower middle income countries						
End of primary reading	2001	2019	23	43	1.51	0.34
End of primary mathematics	2003	2019	16	20	0.81	0.10
Upper middle income countries						
End of primary reading	2001	2019	34	50	-0.01	-0.12
End of primary mathematics	2003	2019	36	48	-0.40	-0.40
High income countries						
End of primary reading	2001	2019	75	94	0.15	0.12
End of primary mathematics	2003	2019	74	74	0.35	0.23

Appendix 2: Trends in GPE countries

The next three tables provide details for the group of 76 countries within the Global Partnership for Education.

Table 13: Global Partnership for Education statistics

	Countries	Population aged 0 to 14 (thous.)	% of children covered by the largest quarter of countries	Year range for 4.1.1 statistics		% of population with any 4.1.1 data point	% of population with any 4.1.1 trend	% of population with any participation trend
GPE countries	76	694,297	77	2000	2019	92	56	100

Table 14: Global Partnership for Education with details by statistic

	Year range		% of population with any data point	% of population with any trend	% of population with a usable trend	Trend coverage (%)	Average level within year range	Annual percentage point gain (continuous approach)	Annual percentage point gain (restricted approach)
GPE countries									
Grades 2/3 reading	2006	2019	71	18	18	2.9	35	1.98	0.34
Grades 2/3 mathematics	2006	2019	71	18	18	3.2	48	1.61	0.36
End of primary reading	2001	2019	73	46	35	11.6	35	0.56	-0.27
End of primary mathematics	2003	2019	74	44	30	8.3	21	0.55	-0.17
Lower secondary reading	2000	2019	24	17	10	1.8	63	1.06	-0.03
Lower secondary mathematics	2000	2019	26	18	12	1.6	55	1.93	0.11
Primary aged not out-of-school	2010	2019	77	54		40.6	82	1.8	
Primary completion	2010	2019	100	100		96.7	62	2.6	

Table 15: Population-focussed Global Partnership for Education breakdown with end of primary details

	Year range		% of population with a usable trend	Average level within year range	Annual percentage point gain (continuous approach)	Annual percentage point gain (restricted approach)
GPE countries						
End of primary reading	2001	2019	31	25	0.69	-0.14
End of primary mathematics	2003	2019	27	16	0.55	-0.12

Appendix 3: A closer look at 2014 to 2019 PASEC trends

Ten PASEC countries had results for two levels of primary schooling and two subjects in both 2014 and 2019, which allowed for the determination of trends over time for these countries. This appendix draws from PASEC microdata and metadata made available to UIS by PASEC, the aim here being to confirm the reliability of the trends. The analysis presented here could be taken further, for instance by extending it to both subjects and both levels of the primary system. The current analysis focusses on reading at the end of primary.

UIS (2021a: 19), in comparing learner responses on electricity in the household with World Bank data on electricity, concluded that in at least one PASEC country the national sample in 2019 may not have been consistent with the 2014 sample, which in turn could have affected the reliability of the learning proficiency trend. Follow-up analysis into this matter concluded that the way electricity information was collected in the PASEC background questionnaires for the two years makes comparison difficult. However, this further analysis found that Burkina Faso and Chad do display levels of electricity access in 2019 in the PASEC data which are two to three times higher than what is found in the World Bank data, even after out-of-school children have been taken into account. *Importantly, other countries among the ten examined here do not display these discrepancies, suggesting that on the whole the PASEC samples are nationally representative.* Triangulation of data to examine the correctness of samples is important, given that in the past sampling issues have distorted trends in testing programmes other than PASEC. Yet it should be kept in mind that this type of analysis is more suggestive than conclusive. The ideal is some external quality assurance of the actual sampling process through examination of, in particular, the school and enrolment data used to determine the sampling frame, and publication of the results of this quality assurance step.

Figure 21 below illustrates item response theory (IRT) score distributions in 2014 and 2019[29], while Figure 22 presents distributions for classical percentage scores out of 100, based on just 41 common items used in both 2014 and 2019[30]. The patterns across the two figures can be considered highly consistent, and indicative of a correct scaling process when raw item responses were converted to IRT scores. For instance, Burundi's curve moved to the right and became steeper, regardless of which of the two figures is used. Niger displays a decline in a peak of worse performing learners, in both figures. This analysis confirms that the scaling process was clean, something already suggested by details provided in the PASEC technical documentation. One advantage with the kind of representation shown below is that it can help explain to teachers, who tend to be familiar with classical scores but not IRT scores, that distributions look relatively similar.

[29] The first of the five plausible values was used. Using the other five produced patterns which were virtually indistinguishable from those seen in Figure 21.

[30] A 42nd common item, f62 in 2019 named f63 in 2014, was excluded from the analysis as it displayed learner responses which were inconsistent across the two years, where the inconsistency was not easy to explain.

Figure 21: IRT end-of-primary reading trends in PASEC

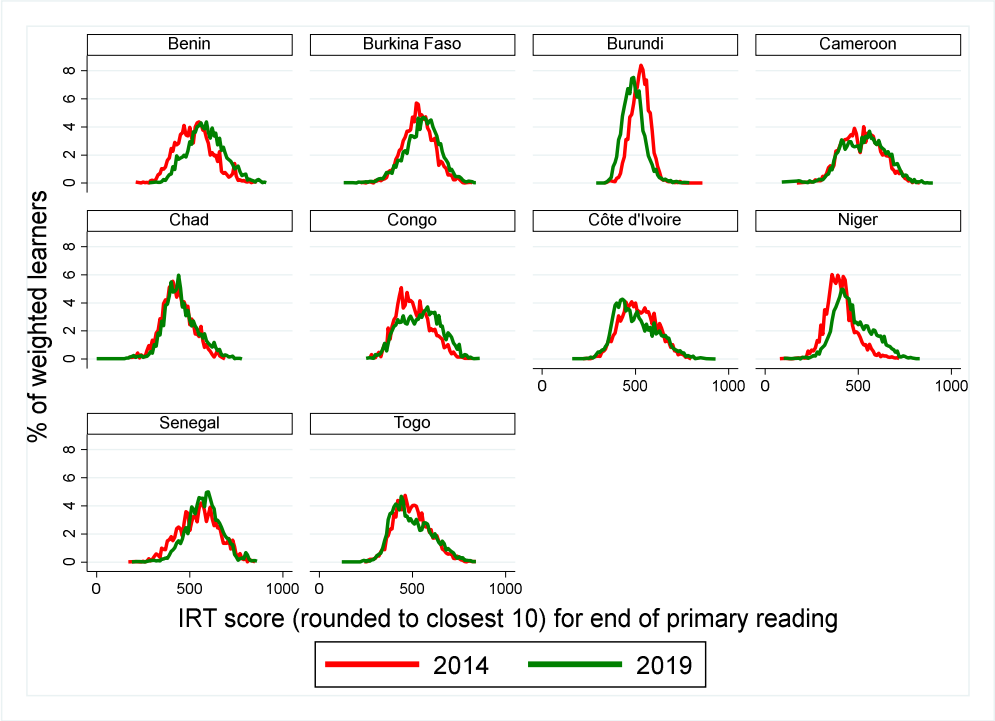
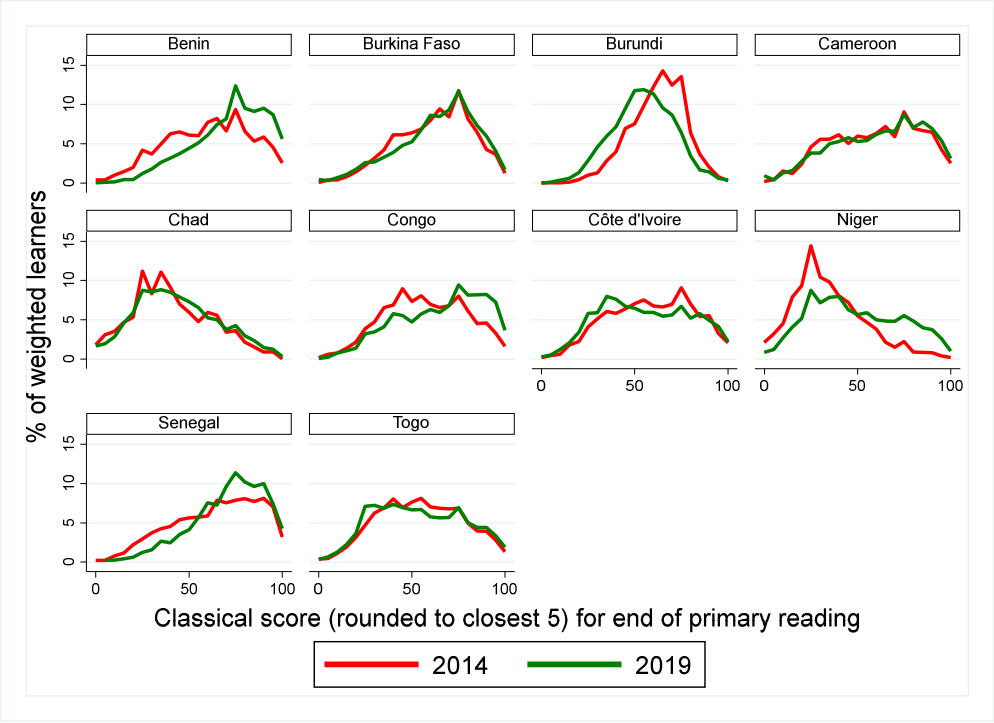


Figure 22: Classical end-of-primary reading trends based on anchor items in PASEC



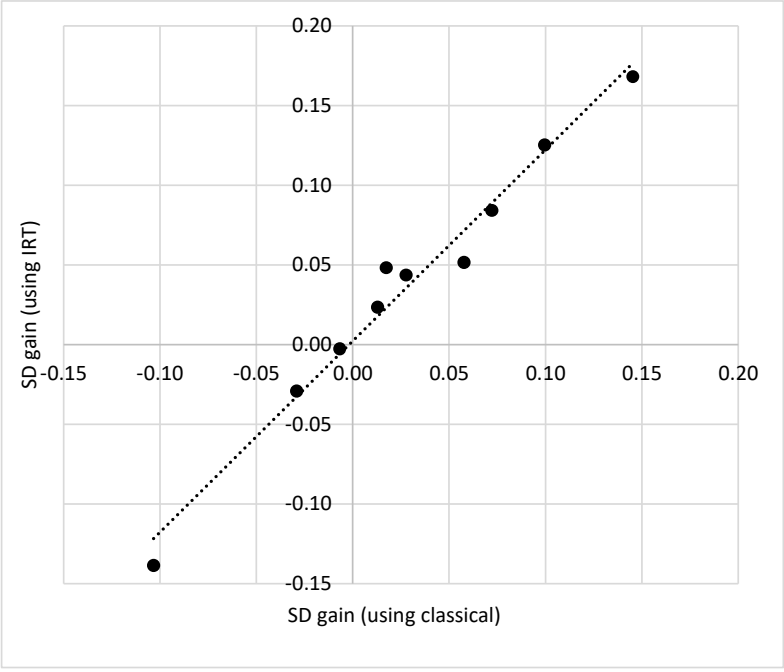
Though a similar analysis was not run for mathematics or for the lower primary level data, if the trends for end-of-primary reading appear consistent, it is very likely that other trends would also be consistent, as the methods and risks would be similar across the various PASEC trends.

Table 16 below presents *annual* gains in terms of standard deviations, using IRT scores and classical scores. IRT means are the ones published by PASEC (CONFEMEN, 2020), while classical means were calculated using the microdata. Standard deviations for 2014 are all from the microdata. The gain statistics based on IRT means are highly correlated with those based on classical scores, which again confirms the absence of any serious problem in the scaling process, of the kind seen in a few other measures of national trends. The correlation coefficient across the two columns of annual gain statistics is 0.99. This high correlation is represented graphically in Figure 23 below.

Table 16: Classical and IRT end-of-primary reading trends compared

	IRT				Classical			
	2014 mean	2014 std. dev.	2019 mean	Annual gain in std. dev.	2014 mean	2014 std. dev.	2019 mean	Annual gain in std. dev.
Benin	523	99	586	.13	60	23	71	.10
Burkina Faso	532	82	552	.05	62	20	64	.02
Burundi	525	51	490	-.14	62	15	55	-.10
Cameroon	518	104	530	.02	60	24	61	.01
Chad	433	83	451	.04	41	21	44	.03
Congo	503	92	542	.08	57	22	65	.07
Côte d'Ivoire	517	97	503	-.03	59	22	56	-.03
Niger	404	80	471	.17	35	19	49	.15
Senegal	548	106	576	.05	64	23	71	.06
Togo	497	94	496	.00	55	22	54	-.01

Figure 23: Classical and IRT end-of-primary reading trends compared



UIS (2019) has suggested that annual improvements greater than 0.6 standard deviations a year, using typical IRT measures, are rare. It moreover indicates that such high improvements can be expected in countries with a low point of departure in terms of learning proficiency levels. All the African countries of Table 16 can be considered to experience low levels of proficiency. The 2019 PASEC report warns that trends for reading in Chad and Niger over the

2014 to 2019 period should not be considered properly comparable, due to changes in the language used[31]. This leaves two countries from Table 16, Benin and Congo, with improvements that in the 'rare' category beyond a 0.06 annual standard deviation gain. These are two countries which according to the electricity access triangulation analysis referred to above display samples which are highly consistent across 2014 and 2019. Even an improvement as steep as Benin's 0.13 annual gain in standard deviations thus appears to be possible and plausible, though uncommon.

[31] CONFEMEN, 2020: 27, 220.