BENCHMARKS FOR SDG 4 INDICATORS:
A POLITICAL AND TECHNICAL BASIS FOR DISCUSSION

Table of Contents

List of tables, figures and boxes

Introduction

The Education 2030 Framework for Action had called on countries to establish "*appropriate intermediate benchmarks (e.g. for 2020 and 2025)*" for the SDG indicators, seeing them as "*indispensable for addressing the accountability deficit associated with longer-term targets*" (§28), a request that remains unrealized.

The extraordinary session Global Education Meeting in October 2020 reminded countries of this commitment. Its Declaration called on "*UNESCO and its partners, together with the SDG-Education 2030 Steering Committee, to … accelerate the progress and propose relevant and realistic benchmarks of key SDG 4 indicators for subsequent monitoring*" (§10).

Fulfilling this neglected commitment to set benchmarks would help renew emphasis on achieving SDG 4. Countries have started from different points and move at different speeds. Unless there is a clearer and shared understanding of where countries started from in 2015, what minimum levels they should achieve and how fast, there is a risk that lack of progress will go unnoticed. But to be effective, benchmarks must be designed to mobilize action and communicated in a transparent and informative way.

The effectiveness of the process to set, monitor and act on benchmarks rests on two factors:
- First, *political commitment is needed*. Setting benchmarks as requested by the Framework for Action cannot be done at global level, given the very large differences in starting points between countries. Benchmarks need to be feasible and based on national ownership. A global process may undermine these objectives. It is therefore proposed to define benchmarks at regional level. Countries within each region tend to have more challenges in common and more opportunities to enter into policy dialogue and learn from each other.
- Second, *technical challenges of measurement need to be overcome*. A set of indicators to benchmark was adopted by the Technical Cooperation Group (TCG) in August 2019. The proposal was based on a review of proposals by TCG members, which concluded that it would be possible to set benchmarks for 6 of the 43 SDG 4 indicators – plus the Framework for Action expenditure indicators – based on past trends, country coverage, frequency of data and policy relevance (**Table 1**).

The purpose of this document is to present options on addressing these two challenges, political and technical, as a basis for discussion at regional and global levels. Annex D (page 21) describes the concepts and methodology used in detail.

Political process for setting benchmarks at regional level

The starting point for a benchmarking setting process at the regional level should be the utilization of existing regional coordination mechanisms and the involvement of regional organizations with an education agenda. While UNESCO mobilizes its SDG 4 regional coordination mechanisms, the active participation of regional organizations is a necessary step for national ownership and essential to achieve alignment between global and regional education agendas and to avoid duplication.

Annex A presents definitions of regions used in SDG reporting and by the UN Statistical Division, mapping of corresponding regional and sub-regional organizations with an education agenda, and includes information on those agendas and whether there are systematic efforts to monitor the results of their implementation.

Table 1. Data availability for proposed benchmark indicators, weighted by school age population

Indicator	World		Africa (Sub-Saharan)		Africa (Northern) and Asia (Western)		Asia (Central and Southern)		Asia (Eastern and South-eastern)		Oceania		Latin America and the Caribbean		Northern America and Europe	
	Data point last 5 years	Trend	Data point last 5 years	Trend	Data point last 5 years	Trend	Data point last 5 years	Trend	Data point last 5 years	Trend	Data point last 5 years	Trend	Data point last 5 years	Trend	Data point last 5 years	Trend
1. Global indicator 4.1.1 Minimum learning proficiency in reading and mathematics	46%	3%	31%	0%	2%	2%	93%	0%	6%	0%	3%	0%	88%	0%	29%	29%
2. Global indicator 4.1.2 Completion rate	87%	82%	88%	84%	76%	62%	94%	94%	95%	89%	36%	0%	92%	92%	51%	34%
3. Thematic indicator 4.1.4 Out-of-school rate	62%	39%	54%	44%	68%	67%	76%	8%	31%	29%	94%	57%	65%	65%	100%	100%
4. Global indicator 4.2.2 Participation rate one year before primary	45%	43%	50%	42%	69%	68%	20%	17%	35%	35%	98%	60%	67%	67%	94%	94%
5. Global indicator 4.c.1 Percentage of trained teachers	53%	51%	54%	49%	50%	43%	98%	97%	22%	22%	4%	3%	42%	42%	5%	4%
6. Education 2030 benchmarks Education expenditure as share of budget and GDP	72%	67%	77%	77%	11%	11%	100%	98%	32%	30%	96%	96%	89%	89%	96%	56%
7. Global indicator 4.5.1 [Equity indicator]	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..

The TCG under the leadership of the UIS has begun and will continue consultations with UNESCO's regional coordination mechanisms and with regional organizations to communicate the decisions made in 2019 and to discuss possible next steps towards endorsement of benchmarks in the respective regions. As part of the process, a (sub-)regional organization will need to:

- **confirm** that it is willing to support a benchmarking process among its member states (with whatever adaptations they feel are needed)
- **coordinate** with other organizations where memberships overlap
- **identify** a timeline of consultation and other steps that will lead to benchmarks being approved
- **communicate** these steps to the TCG to develop a global roadmap
- **request** technical support, where necessary, to facilitate a regional benchmarking process

It is important to stress the following two aspects of flexibility in the proposed process:

- The definition of region is flexible. Any (sub-) regional organization that expresses an interest to lead the process of defining benchmarks for its member states may do so as long as it coordinates the process with other organizations with which it has members in common.
- Setting benchmarks for seven global indicators is consistent with the possibility that a regional organization may wish to set additional benchmarks for other indicators that are key to their agenda.

Technical process for setting regional benchmarks: two approaches

The setting of SDG 4 indicator benchmarks will serve three objectives:

- **Availability**: identify data gaps that prevent monitoring progress on key SDG 4 indicators;
- **Accountability**: assess progress relative to feasible, historically observed trends; and
- **Actionability**: lead to data collection and policy responses to fill gaps and accelerate progress.

Selecting a benchmarking method to achieve these objectives will be based on the following FERST principles:

- **Fairness**: Countries accept the value of benchmarks and that their values are set in a fair way taking SDG 4 aspirations, their initial conditions and feasible past progress into account.
- **Efficiency**: The data that need to support the benchmarks are available for the largest possible number of countries, on a regular basis and in a timely way.
- **Relevance**: The indicators are selected to correspond to national and regional agendas and the assessment of whether the benchmark has been met can be linked to policy responses.
- **Simplicity**: Benchmarks need to be understood by all countries, while striking a balance between the three objectives outlined above.
- **Transparency**: The process by which benchmarks were developed needs to be verifiable and, to the extent possible, systematic, while it needs to be communicated clearly.

The selection of the seven indicators largely meets the principles of efficiency and relevance. This document addresses two main ways to select benchmarks for the first five indicators (i.e. all except those related to financing and equity) to meet the principles of fairness, simplicity and transparency (**Table 2**).

The first approach is suitable for regions or sub-regions that are relatively homogeneous. A common, **regional minimum benchmark** is set as a minimum that all countries should achieve by 2030. Different ways can be used to set the minimum. For instance, at the lowest end, the regional benchmark could be equal to the minimum progress the country with the lowest indicator value in the region at baseline can achieve. A more ambitious regional benchmark could be equal to the minimum progress a country with an indicator value, say, at the bottom quarter, third or half of countries in the region can achieve.

The second approach assumes that a common regional benchmark is not realistic because countries differ too much even within a region or sub-region. Instead, every country has its own benchmark. When all the country-specific benchmarks are added up, an implicit regional target 'benchmark' emerges. In setting their own benchmarks, an important reference point is a **country-specific minimum benchmark** which reflects feasible progress observed historically for countries with a similar initial level of the indicator or starting point. **Box 1** provides definitions of the benchmarking terms used in this document.

Table 2. Regional benchmarking approaches considered

Description	Do all countries in a region have the same benchmark?	Is the benchmark feasible for all countries?	Does achieving the minimum benchmark (or higher) result in meaningful* progress?
Approach 1: Common regional minimum benchmark for all countries Each country in a region has the same benchmark, which is equal to the feasible progress an indicative country in the region is expected to make (e.g. the country furthest behind, the country in the bottom 25% etc.)	Yes	Depends on the level of the benchmark: higher implies less feasible	Depends on the level of the benchmark: higher implies more progress
Approach 2: Different benchmark for each country Each country sets its own benchmark. One approach is to use **country-specific minimum benchmark** based on its initial value and a rate of progress reflecting past observed progress (e.g.: the median progress for countries that have improved since 2000 as applied below).	No	Yes	Yes, but for countries with slow progress historically

* 'Meaningful progress' compares the projected value for the region if countries achieve the minimum benchmark or their projected value (whichever one is higher) to the projection for 2030 in absence of the benchmark.

Despite the fact that the two approaches differ in this important respect, regions could opt for a variation that includes both. For instance, under Approach 1, a region or sub-region may opt for a common benchmark for all countries. However, this benchmark will be too low for several countries in the region **(Table 3)**. Some of them may therefore select their own more ambitious benchmark**.**

Table 3. Approach 1: Regional minimum benchmark for indicator 4.1.1b

Region or country-income group	Baseline (2015 ±2 years)	Minimum regional benchmark	Countries achieving benchmark in baseline
Africa (Sub-Saharan)	20	29	24
Africa (Northern) and Asia (Western)	46	42	68
Asia (Central and Southern)	37	38	50
Asia (Eastern and South-eastern)	65	51	69
Oceania	58	34	36
Latin America and the Caribbean	50	45	70
Europe and Northern America	78	68	74
Low income	10	28	0
Lower middle income	35	34	33
Upper middle income	62	47	63
High income	80	67	84

Under Approach 2, countries may accept the **country-specific minimum benchmark** based on their initial value and a target feasible rate of progress or they may reject it and adopt instead a higher benchmark depending on their national ambitions and priorities. The need for countries to take an active role in setting their benchmarks is envisaged in the Framework for Action.

Proposed interim national benchmarks

For technical assistance in setting national benchmarks, a dataset will be provided three key reference points for each indicator: (1) minimum regional benchmarks defined as the average country-specific minimum benchmark for the lowest third of countries in each region, (2) the country-specific minimum benchmark which is a feasible benchmark for a country based on its latest indicator value and the median rate of progress for countries with a similar value that improved since 2000, and (3) a country-specific projection which reflects a country's own historical progress for the indicator since 2000. These three reference points are intended to provide countries a basis for setting their own national benchmarks.

Until countries select their own benchmark for each of the seven indicators, the following interim national benchmarks are proposed following approach 2. The method for assigning interim national benchmarks is to take the highest value of the three reference points provided: (1) the regional minimum benchmark, (2) the country-specific minimum benchmark and (3) the country-specific projection. The following example illustrates how the two approaches can yield different benchmarks for a country **(Table 4).**

- Country A, B and C are in the same region and have the same **regional minimum benchmark** (60 percent) for 2030.

- Countries A and B are projected to exceed the regional benchmark by 2030 at 75 percent and 68 percent, respectively. For these countries, the regional minimum benchmark is not ambitious. Because both countries have the same starting point, their **country-specific minimum benchmarks** for 2030 are the same.

- However, Countries A and B differ in projection. Country A was historically a high performer and is expected to achieve 75 percent by 2030 which exceeds its **country-specific minimum benchmarks**, while Country B is projected to be below its **country-specific minimum benchmark** by 2030. As a result, the proposed national benchmark for Country A is its **country-specific projection** given its rapid progress historically.

- For Country B, the proposed national benchmark is the **country-specific minimum benchmark** because achieving this level is expected to be feasible for Country B and offers a realistic goal.

- Finally, for Country C, its **country-specific projection** for 2030 and **country-specific minimum benchmark** is below the **regional minimum benchmark**; as a result, its proposed national benchmark is the **regional minimum benchmark**.

Table 4. How the proposed interim national benchmarks were defined

	Baseline	Regional minimum benchmark for 2030	Country-specific minimum benchmark for 2030	Country-specific projection for 2030	Proposed national benchmark for 2030	Nationally set benchmark
Country A	58	60	72	75	75	-
Country B	58	60	72	68	72	-
Country C	36	60	54	52	60	-

Annex B presents results from the application of these approaches. **Tables B1-B3** compare how benchmarks would look like for each region if each approach were applied. **Tables B4-B5** present the benchmarks of Approach 1, for four of the seven indicators, which uses the progress rate of the bottom third of countries and which is being proposed as a basis for discussion.

Box 1. Core concepts used in the proposed benchmark approaches

Country-specific projection for 2030[1]: This is the value that a country is expected to achieve by 2030 based on its historical trend from 2000 to the latest available year. For most indicators, countries with higher levels tend to have lower progress, and this is accounted for in the projection model by using the country's progress relative to other countries with the same starting point.

Country-specific minimum benchmark for 2030: This is the value that a country could feasibly achieve by 2030 given the progress made by other countries historically with the same starting point. Note that that the minimum benchmark for 2030 differs from the country's projection for 2030: the latter is what the country *is expected to achieve*, while the former is based on what *other countries with the same starting point have actually achieved*.

Regional minimum benchmark for 2030: This is the minimum acceptable level that all countries in a region should achieve. Because many regions have a diverse range of countries with both high and low performers, this benchmark may not apply to countries that are expected to or could feasibly achieve a higher level—in these cases countries are expected to set their own more ambitious benchmarks or use the feasible national benchmarks. For the examples presented in Annex 2, the minimum regional benchmark for 2030 is the average[2] of the country-specific minimum benchmarks for the lowest third of countries.

---

[1] These are described for 2030 but are applicable to any reference year.
[2] A school-aged weighted average was used.

Nationally set benchmarks for 2030: Countries are expected to set their own national benchmarks reflecting national priorities and goal. The above three reference points are provided to assist in deriving their own benchmarks.

Proposed interim national benchmark for 2030: This is the highest of (1) the regional minimum benchmark, (2) country-specific minimum benchmark, or (3) country-specific projection for 2030.

Other key methodological issues in setting benchmarks

The following potential methodological issues may be noted:

- While benchmarks intend to capture feasible progress achievable in a region, the lack of sufficient data by region means that global progress rates are inevitably applied as the basis for calculations.
- For simplicity, feasible progress is defined over levels achieved until 2018 and will not be updated to capture progress achieved since 2018.
- Simplicity is also behind the implicit decision to define benchmarks as a *level* to be achieved at any year between now and 2030, rather than a benchmark *trend* with values set for every year.

With respect to the latter point, while the benchmark is defined as a level, the rate of progress will also be monitored. **Annex C** proposes a dashboard that distinguishes:

- Whether the benchmark is reached or not is marked by a colour code; lack of data is marked by grey.
- Whether the value of the indicator is increasing (fast) or decreases (fast) is marked by arrows.

The following issues also need to be considered:

- Several indicators are available at different education **levels** (e.g. primary, lower secondary and upper secondary in indicators 4.1.1, 4.1.2, 4.1.4, 4.c.1) or **domains** (e.g. reading and mathematics in indicator 4.1.1). Although benchmarks can be set and monitored for all combinations of levels and domains, one level or domain for each indicator may be selected as the level and domain of focus (**Table 5**).
- The **baseline** year is 2015. The value is defined as the average of 2013 to 2017.
- The dashboard can be updated and monitored on an annual basis as new information is incorporated. But it may also be desirable to take stock of the benchmarked indicators in 2020, 2025 and 2030. Often data will not be available for these years of reference. For instance, indicator 4.1.1 is typically available every 3 to 5 years depending on the frequency of the assessment in which the country is participating. As with the baseline, it is proposed to report the latest value in the last 5-year period.
- Data for some indicators are potentially available from multiple **sources**. For instance, the completion rate and the out-of-school rate may be estimated through administrative data collected by the UIS survey or through household survey data. For the time being, it is proposed that a single source is used (**Table 6**). In the future, methodologies that ensure the efficient use of both sources may be considered.

- The choice of regional benchmarks and the fact that collectively they may not add up to the achievement of the target does not dilute the commitment of the international community to achieve the targets as spelled out in the Education 2030 Framework for Action.

Table 5. Additional methodological points on choice of benchmarks

Indicator	Indicator levels/domains (1)		Possible level/ domain (2)	Baseline estimate (3)		Sources (4)
	Levels	Domains		Year	Method	
Global indicator 4.1.1 Minimum learning proficiency in reading and mathematics	3 Grades 2/3; end of primary; end of lower secondary	2 Reading; mathematics	End of primary, reading	2015	According to protocol for reporting indicator	Learning assessments
Global indicator 4.1.2 Completion rate	3 Primary; lower secondary; upper secondary	–	Lower secondary	2015	Last value in 2011-2015	Household surveys and censuses
Thematic indicator 4.1.4 Out-of-school rate	4 Pre-primary; primary; lower secondary; upper secondary	–	Lower secondary	2015	Last value in 2011-2015	UIS education survey
Global indicator 4.2.2 - Participation rate one year before primary	1	–	–	2015	Last value in 2011-2015	UIS education survey
Global indicator 4.c.1 - Percentage of trained teachers	4 Pre-primary; primary; lower secondary; upper secondary	–	Primary	2015	Last value in 2011-2015	UIS education survey
Education expenditure as share of total expenditure	1	–	–	2015	Last value in 2011-2015	UIS education survey
Education expenditure as share of GDP	1	–	–	2015	Last value in 2011-2015	UIS education survey

References

UIS (2019a) *SDG 4 indicator benchmarking consultation*. Montreal: UNESCO Institute for Statistics.

UIS (2019b) *Sixth meeting of the Technical Cooperation Group on the Indicators for SDG 4 Education 2030 Post meeting consultation on indicator development and monitoring*. Montreal: UNESCO Institute for Statistics

Annex A. Political process

Table A.1. Regions and regional organizations

SDG regions	UNSD regions	Regional organizations
	Africa	AU
Africa (Sub-Saharan)	Sub-Saharan Africa	
	Eastern Africa	EAC IGAD
	Middle Africa	ECCAS
	Southern Africa	SADC
	Western Africa	ECOWAS
Africa (Northern)	Northern Africa	ALECSO
	Asia	
Asia (Western)	Western Asia	ALECSO
Asia (Central and Southern)	Central Asia	
	Southern Asia	SAARC
Asia (Eastern and South-eastern)	South-eastern Asia	ASEAN SEAMEO
	Eastern Asia	OECD
Oceania	**Oceania**	PIF SPC Forum Education Ministers (FEM)
	Australia and New Zealand	OECD
	Melanesia	
	Micronesia	
	Polynesia	
	Americas	OAS
Latin America and the Caribbean	Latin America and the Caribbean	OEI
	Caribbean	CARICOM OECS
	Central America	CECC
	South America	MERCOSUR ANDEAN
Northern America	Northern America	SPC OECD
Europe	**Europe**	COE EU OECD
	Eastern Europe	
	Northern Europe	
	Southern Europe	
	Western Europe	

Table A.2. Regional organizations' education agendas

	Education agenda / strategy	Monitoring framework	Monitoring report	Targets / benchmarks
Sub-Saharan Africa				
African Union (AU)	Continental Education Strategy for Africa 2016-2025	Yes	No	No
Northern Africa/Western Asia				
Arab League Educational Cultural and Scientific Organization (ALECSO)	–	–	–	–
Asia/Pacific				
Association of Southeast Asian Nations (ASEAN)	Work Plan on Education 2016-2020	No	Yes	No
Southeast Asian Ministers of Education Organization (SEAMEO)	Education Agenda 2035	No	No	No
South Asian Association for Regional Cooperation (SAARC)	SAARC Development Goals	In process	No	No
SPC Pacific Community	Education Quality Assessment Programme	Yes	Yes	No
Forum Education Ministers				
Latin America/Caribbean				
Caribbean Community (CARICOM)	Human Resource Development 2030 Strategy	Yes	Yes	No
Central American Educational and Cultural Corporation (CECC)	Central America Education Programme (PEC)	Yes	Yes	No
Organization of Eastern Caribbean States (OECS)	OECS Education Sector Strategy	Yes	Yes	No
Europe and Northern America				
Council of Europe (COE)	Operational Programme	No	No	No
European Union (EU)	Education and Training 2020	Yes	Yes	Yes
Other				
Organization for Economic Co-operation and Development (OECD)	–	Yes	Yes	No

Annex B. Technical alternatives

Table B.1. Comparison of benchmark values for indicator 4.1.1 (end of primary, reading)

Region	Baseline Regional Average Value 2015 (+/- 2 years)	Regional Minimum Benchmark (Approach 1)	Projected regional average (if no benchmarks are achieved – business as usual)	Projected regional average if interim national benchmarks are achieved (Approach 2)
SDG Regions				
SDG: Africa (Northern) and Asia (Western)	64	58	65	74
SDG: Africa (Sub-Saharan)	27	21	38	48
SDG: Asia (Central and Southern)	46	38	50	57
SDG: Asia (Eastern and South-eastern)	72	60	76	84
SDG: Europe and Northern America	91	97	92	98
SDG: Latin America and the Caribbean	50	48	50	65
SDG: Oceania	68	29	71	76
WB Regions				
WB: conflict-affected situations	27	18	36	45
WBG: high income	91	97	93	98
WBG: low income	10	17	23	32
WBG: lower middle income	43	34	49	57
WBG: upper middle income	71	55	73	83

Notes:

- Under Approach 1, the benchmark is equal to the average of the country-specific minimum benchmarks for the bottom third of countries in the region. Note that the average is a school-age population weighted.

- Under Approach 2, the regional values are not a regional benchmark per se but the regional average of the interim national benchmarks; this offers a goal for the region to achieve.

Table B.2. Regional benchmarks for SDG indicator 4.1.1 - Reading at the end of primary

Region	Average baseline value 2015±2 years	Regional Minimum benchmark 2030 (Approach 1)	Average of interim national benchmarks 2025 (Approach 2)	Average of interim national benchmarks 2030 (Approach 2)
SDG Regions				
SDG: Africa (Northern) and Asia (Western)	63.7	58.4	71	74
SDG: Africa (Sub-Saharan)	27.2	21.0	44	48
SDG: Asia (Central and Southern)	46.3	38.0	53	57
SDG: Asia (Eastern and South-eastern)	72.3	60.1	80	84
SDG: Europe and Northern America	91.2	97.3	97	98
SDG: Latin America and the Caribbean	50.3	47.6	61	65
SDG: Oceania	67.5	28.7	74	76
WB Regions				
WB: conflict-affected situations	26.9	17.8	41	45
WBG: high income	91.2	97.1	98	98
WBG: low income	10.5	17.4	28	32
WBG: lower middle income	43.4	33.5	53	57
WBG: upper middle income	71.1	55.2	79	83

* The average baseline is weighted by the total school age population in each country.

* Average of countries' interim national benchmarks for 2025 and 2030 is the regional average if the interim national benchmarks are achieved.

Table B.3. Regional benchmarks for SDG indicator 4.1.2- End of Primary

Region	Average baseline value 2015±2 years	Regional minimum benchmark 2030 (Approach 1)	Average of interim national benchmarks 2025 (Approach 2)	Average of interim national benchmarks 2030 (Approach 2)
SDG Regions				
SDG: Africa (Northern) and Asia (Western)	88	98	94	96
SDG: Africa (Sub-Saharan)	65	64	81	86
SDG: Asia (Central and Southern)	89	86	95	97
SDG: Asia (Eastern and South-eastern)	96	98	99	99
SDG: Europe and Northern America	99	100	100	100
SDG: Latin America and the Caribbean	93	98	97	98
SDG: Oceania	89	84	90	92
WB Regions				
WB: conflict-affected situations	66	61	81	85
WBG: high income	99	100	100	100
WBG: low income	53	60	74	80
WBG: lower middle income	84	82	92	96
WBG: upper middle income	96	98	99	100

Table B.4. Regional benchmarks for SDG indicator 4.1.4 OOSCI End LS

Region	Average baseline value 2015±2 years	Regional minimum benchmark 2030 (Approach 1)	Average of interim national benchmarks 2025 (Approach 2)	Average of interim national benchmarks 2030 (Approach 2)
SDG Regions				
SDG: Africa (Northern) and Asia (Western)	12	15	10	8
SDG: Africa (Sub-Saharan)	29	33	23	21
SDG: Asia (Central and Southern)	14	11	10	9
SDG: Asia (Eastern and South-eastern)	8	11	7	6
SDG: Europe and Northern America	2	6	2	3
SDG: Latin America and the Caribbean	7	14	6	5
SDG: Oceania	6	16	7	7
WB Regions				
WB: conflict-affected situations	29	38	24	22
WBG: high income	3	9	3	3
WBG: low income	42	35	35	32
WBG: lower middle income	15	20	11	10
WBG: upper middle income	7	10	6	5

Table B.5. Regional benchmarks for SDG indicators 4.2.2 and 4.c.1 lower-secondary

I	4.2.2 Participation rate one year before primary		4.c.1 Trained teachers (lower-secondary)	
	Baseline 2015 (+/- 2 years)	Regional minimum benchmark	Baseline 2015 (+/- 2 years)	Regional minimum benchmark
SDG Regions				
SDG: Africa (Northern) and Asia (Western)	54	59	83	98
SDG: Africa (Sub-Saharan)	54	45	68	67
SDG: Asia (Central and Southern)	77	56	74	83
SDG: Asia (Eastern and South-eastern)	92	80	86	99
SDG: Europe and Northern America	94	88	87	99
SDG: Latin America and the Caribbean	90	86	84	77
SDG: Oceania	80	68	82	78
WB Regions				
WB: conflict-affected situations	50	41	71	65
WBG: high income	92	90	89	98
WBG: low income	30	42	63	68
WBG: lower middle income	74	56	76	80
WBG: upper middle income	91	70	84	81

Table B.6. Regional benchmarks for public expenditure indicators

	1.a.2 percent of government expenditure on education		Government expenditure on education (% of GDP)	
	Baseline 2015 (+/- 2 years)	Regional minimum benchmark	Baseline 2015 (+/- 2 years)	Regional minimum benchmark
SDG Regions				
SDG: Africa (Northern) and Asia (Western)	14	15	4	4
SDG: Africa (Sub-Saharan)	17	15	4	4
SDG: Asia (Central and Southern)	14	15	4	4
SDG: Asia (Eastern and South-eastern)	15	15	4	4
SDG: Europe and Northern America	12	15	5	4
SDG: Latin America and the Caribbean	17	15	5	4
SDG: Oceania	13	15	4	4
WB Regions				
WB: conflict-affected situations	15	15	4	4
WBG: high income	12	15	5	4
WBG: low income	17	15	3	4
WBG: lower middle income	15	15	4	4
WBG: upper middle income	16	15	4	4

Annex C. Summary benchmark table

Legend

■ (dark grey)	No data
■ (light grey)	No data for trend
■ (dark green)	Exceeds benchmark by much
■ (light green)	Exceeds benchmark by a little
■ (light red)	Misses benchmark by a little
■ (dark red)	Misses benchmark by much
↑	increases slowly
⇔	no change
↓	decreases slowly
⬇	decreases quickly
⬆	increases quickly

Annex D. Benchmarking SDG 4 Methodology Annex

Country-specific Projections for 2030

For each indicator and for each country, projections were calculated based on countries' historical trends in three steps: (1) estimating the historical relationship between indicator level and annual progress for the indicator, (2) estimating each country's relative level of progress historically compared to other countries with the same level of the indicator, and (3) projecting future levels of the indicator, recursively, for each country based on (1) and (2). This approach accounts for the fact that for many indicators, progress tends to be lower as the indicator reaches 100 percent.

1. Estimating the relationship between level and progress for each indicator: To estimate the relationship between progress and level historically for a given indicator, data on the average measure of progress for each country since year 2000 and the average level of the indicator for the same period were used. The definition of progress varied by indicator and was specified either as the percentage point difference or as proportionate change (see **Table D.1** for specific methodologies for each indicator). The definition was chosen in order to maximize the fit of the model used to estimate the relationship between progress and the indicator level. For many indicators, countries with lower levels[3] of an indicator made more rapid progress across time historically; for other indicators, progress appeared to not to depend on current level. As a result, two approaches were applied to estimate the relationship between annual progress and current level. For the first approach, where there was an apparent relationship between level and progress, percentiles of progress conditional on level were estimated using quantile regressions. Quantile regressions were estimated in 5 percentage point intervals between the 10th and 90th percentiles. These provided fitted percentiles conditional on level. In **Figure D.1**, progress in terms of average annual percentage point increase in SDG 4.1.1 (minimum learning proficiency in reading at the secondary level) and average level of the indicator for each country is plotted, along with the conditional 25th, 50th and 75th percentiles. The negative slope of these three conditional percentile lines reflects lower progress being exhibited by countries with higher levels of the indicator. Albania's historical annual increase and value is shown as an example (denoted as a red circle) and falls at the 75th percentile given its average level. In other words, compared to other countries at the same level, Albania's progress exceeded 75 percent of countries. For the second approach, when progress does not apparently decrease (or increase) with level, an alternative approach is used and described below.

---

[3] Note that most indicators, a higher level is considered better (e.g.: percent achieving minimum learning proficiency), but in some cases a lower level is better (e.g.: the percent of out-of-school children); for this methodological note, the terms used refer to an indicator in which higher is better.

Figure D.1. SDG 4.1.1 (lower secondary, reading) historical indicator progress and level by country
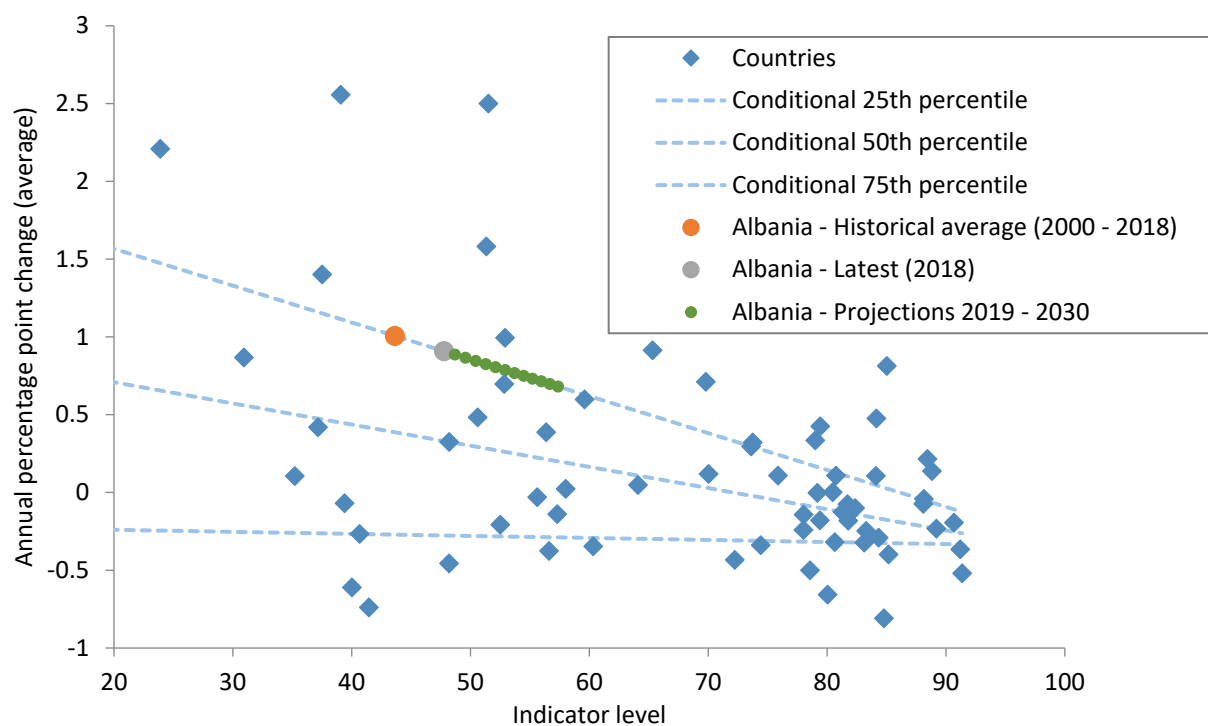

2. Defining a country's _relative_ progress: To estimate a country's historical relative performance, the country's conditional percentile was identified using the same data as in Step 1 and the estimated conditional percentiles. This was achieved by identifying which two conditional percentile lines estimated in Step 1 that a country's data point fell between; a weighted average of the two lines' percentiles was then used to specify its relative progress. For example, if, given a country's average level of the indicator, its progress fell between the 75th and 80th quantile regression lines, then the estimated percentile for a county is the weighted average between these two percentiles. The weights were defined by the distance between the two lines. If the 75th percentile for that level of the indicator was 5 percentage points per year and the 80th percentile for that level of the indicator was 6 percentage points per year, and if the country's historical annualized progress was 5.2 percentage points, then the country's assigned percentile would be the 76th percentile. For countries whose annualized historical progress fell outside the conditional 10th or 90th percentiles lines, they were assigned to either the 10th or 90th percentile, respectively, to eliminate extreme and unlikely levels of progress. In the case of Albania, its progress fell on the conditional 75th percentile line; as a result, its progress was defined as being at the conditional 75th percentile.

3. Recursive projections: Step 1 provides an estimate of the relationship between progress and level for an indicator historically in terms of conditional percentiles, and Step 2 provides a definition of a country's relative progress. The future projections for a country's indicator are calculated for each year in two sub-steps for each year. The first sub-step is to take the latest indicator value (starting

with the actual latest value for the first year or the latest projected value for subsequent years), and compute the annual progress defined in Step 1 for its relative progress defined in Step 2. For the example depicted in **Figure D.1**, the latest value for Albania was 47.8 percent (the X-coordinate of the green circle in **Figure D.1**). Because Albania's historical performance was defined to be at the conditional 75th percentile in Step 2, the 75th percentile of progress for a country with a level of 47.8, estimated in Step 1, was an increase of 0.9 percentage points (the Y-coordinate of the green circle in **Figure D.1**). For the second sub-step, the subsequent year's value is projected by adding the annual percentage point increase. In the example depicted in **Figure D.1**, an increase of 0.9 percentage points from the indicator value of 47.8 in 2018 results in a projected value of 48.8 percent for 2019 (the X-coordinate of the furthest left orange circle in **Figure D.1**). These two sub-steps are repeated using the projected value for 2019 to project a value for 2020 and so on until 2030. In the figure, the orange circles depict the projected values (X-coordinate) and projected increase (Y-coordinate) for each from 2019 to 2030. Albania was essentially a higher performer historically. Its future progress is expected to "follow" the conditional 75th percentile line; in this sense, a decline in progress is expected to for Albania because countries with higher levels of the indicator have made less progress historically. However, Albania is still expected to be high performer relative to its level. Future projections are presented in **Figure D.2** for Albania. Note that the number of years between a country's latest value and 2030 varies by country so this procedure may be repeated a different number of times for each country. Finally, for the second approach, when there is no apparent relationship between progress and level for an indicator, then projected progress is the average historical trend which is constant for each year of projection; this too is bound by the 10th and 90th percentiles.

Figure D.2. Albania's historical level and projected levels for SDG 4.1.1 (lower-secondary reading)
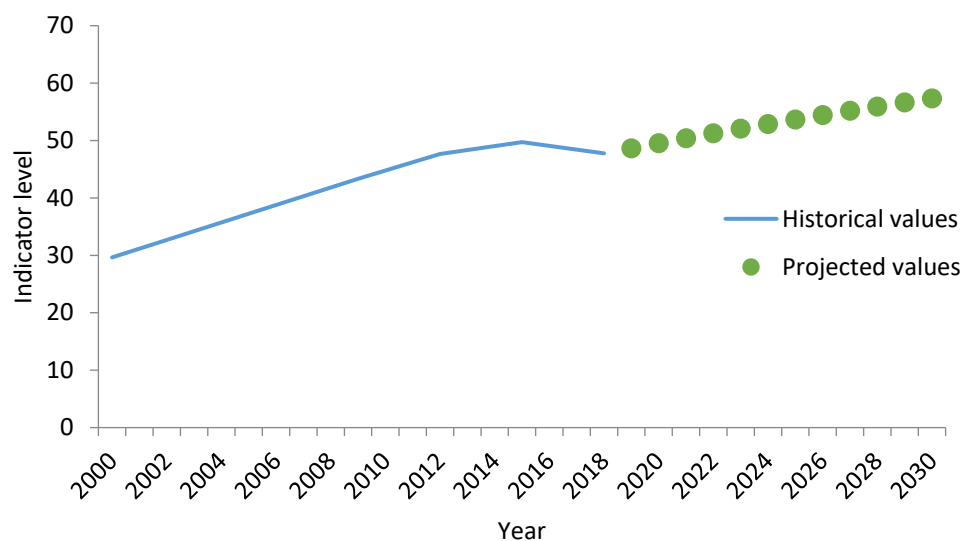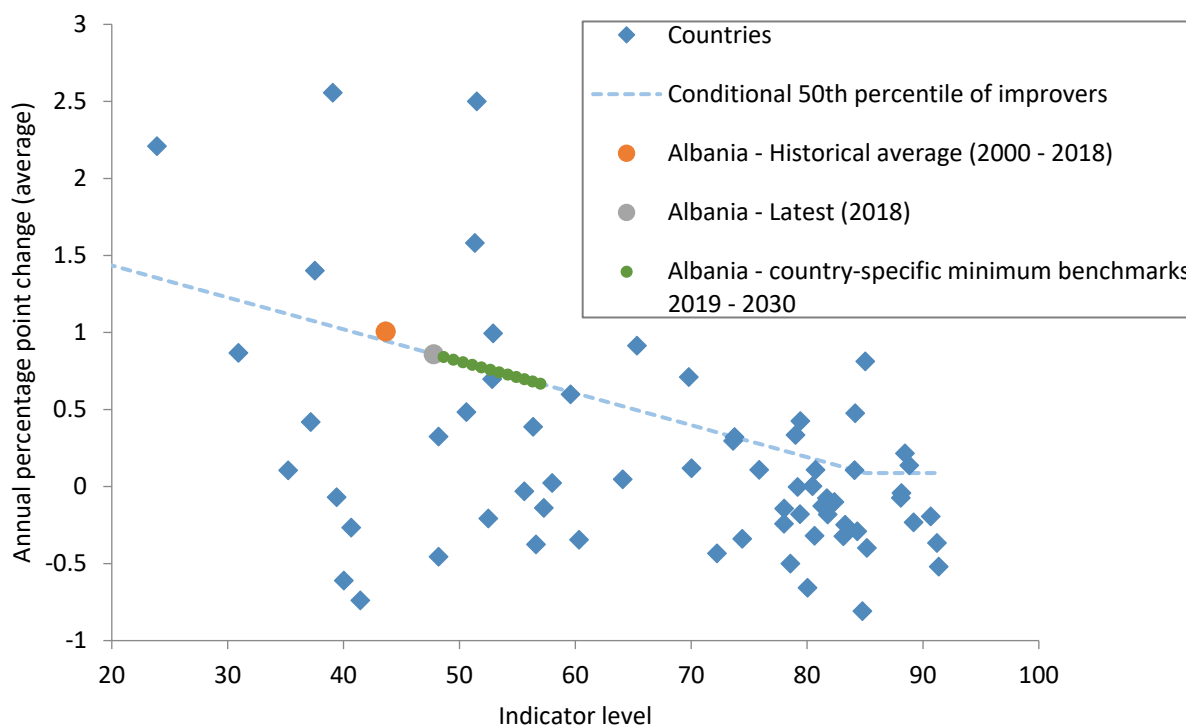
Country-specific minimum benchmarks for 2030

Country-specific minimum benchmarks were defined to be feasible conditional on a country's starting point. For indicator and each country, a minimum benchmark was defined for 2030 reflecting progress conditional on a country's starting point that has been historically achieved by half of improvers. The choice of half of improvers, in other words, the 50th percentile of improvers, was to reflect a level of progress that improving countries were historically equally likely to be above or below. A higher level of progress would result in countries, historically, being less likely than not to achieve and, vice versa, for a lower level of progress. The methodology to construct the minimum benchmark for 2030 was essentially the same as that for projections, with some key differences in Steps 1 and Step 2. The following describes these differences; all other aspects of the computation are the same as for the projections.

1. Estimating the conditional 50th percentile of improvers: The same data in Step 1 for projections were used except the data was limited only to countries that improved, and only one quantile regression was estimated, the conditional 50th percentile.

2. Estimating a country's feasible progress: For the projections, each country was assigned a relative level of progress. For the country-specific minimum benchmarks, the feasible level of progress was defined as the conditional 50th percentile except for high achieving countries. High achievers were defined as having a latest indicator level higher than a threshold that varied by indicator (see **Table D.1**). For these countries, the feasible level of progress was truncated at that for the high achiever threshold. The reason for truncating the feasible minimum progress for high achieving countries is that the conditional 50th percentile can be virtually zero or negative given the low or negative progress of countries with a high level of the indicator. For SDG 4.1.1 (secondary reading) high achievers were defined as those having a level above 85 percent, and **Figure D.3** (blue line) plots the feasible level of progress conditional on indicator level.

3. Recursive minimum benchmarks for 2030: The process for defining benchmarks for each year after the year with the latest value is analogous to that of the projections. Based on the latest year with an indicator value for a country, the feasible level of progress is calculated based on Step 2 (e.g.: the blue line depicted in **Figure D.3**). This feasible level of progress is added to the latest indicator value to provide the minimum benchmark for the first year, and this process is repeated. For example, for Albania, the feasible level of progress for its latest value of the indicator, 47.8 percent in 2018, was 0.86 percentage points (slightly lower than its projected progress of 0.91 percentage points). The minimum benchmark for 2019 is therefore 48.7 percent. This process is repeated from 2019 to 2030 and depicted by the orange circles in **Figure D.3**. In this sense, feasible progress "follows" the feasible progress line in **Figure D.3** (blue line). The result is that for Albania, the minimum benchmark for 2030 is 57.0 percent compared to the projected value for 2030 of 57.4 percent. In this case, both the projected value and the minimum benchmark for 2030 are very close. This is because the conditional 50th percentile of improvers is nearly the same as the 75th percentile of countries overall; there are a number of countries in this indicator for which the indicator has declined. This is somewhat unique to learning outcomes because learning outcomes vary from year to year, and may decline when inclusion in the education systems expands to more disadvantaged or at-risk children.

Figure D.3. SDG 4.1.1 (lower secondary, reading) historical indicator progress and level by country


Imputation

Two data points are required to generate projections for a country while only one data point is required to generate the country-specific minimum benchmarks. This is because the projection methodology requires a historical measure of relative progress. For countries that had only one data point or whose data points were less than three years apart, projected values were estimated using the conditional 50th percentile (median) level of progress rather than a country-specific, historical level of progress. For countries with no data points, an imputation method was used to estimate a 2015 baseline value from which projections and country-specific minimum benchmarks could then be generated. The purpose of the imputed values was to provide data for defining regional benchmarks which are described below. Imputed values were not used in estimating the conditional percentile lines used in the first steps of either the projection method or the minimum benchmark method described above.

The imputation method for countries without data involved estimating the relationship between countries' GDP per capita and indicator level. Imputed values were then estimated value based on a country's GDP per capita. For countries without GDP per capita data, no imputations were calculated. To estimate this relationship, the relationship was modeled using a linear regression with the level of the indicator converted to logits (log odds) as the dependent variable and log GDP per capita as the

independent variable. This model appeared to provide the best fit for the data. For example, **Figure D.4** depicts SDG 4.1.1 (lower secondary reading) and log GDP per capita as well as the fitted line based on the estimated regression model. Countries without values for the indicator were assigned fitted baseline values given their log GDP per capita (blue line, **Figure D.4**).

Figure D.4. Baseline value and log GDP per capita (SDG 4.1.1 secondary reading)
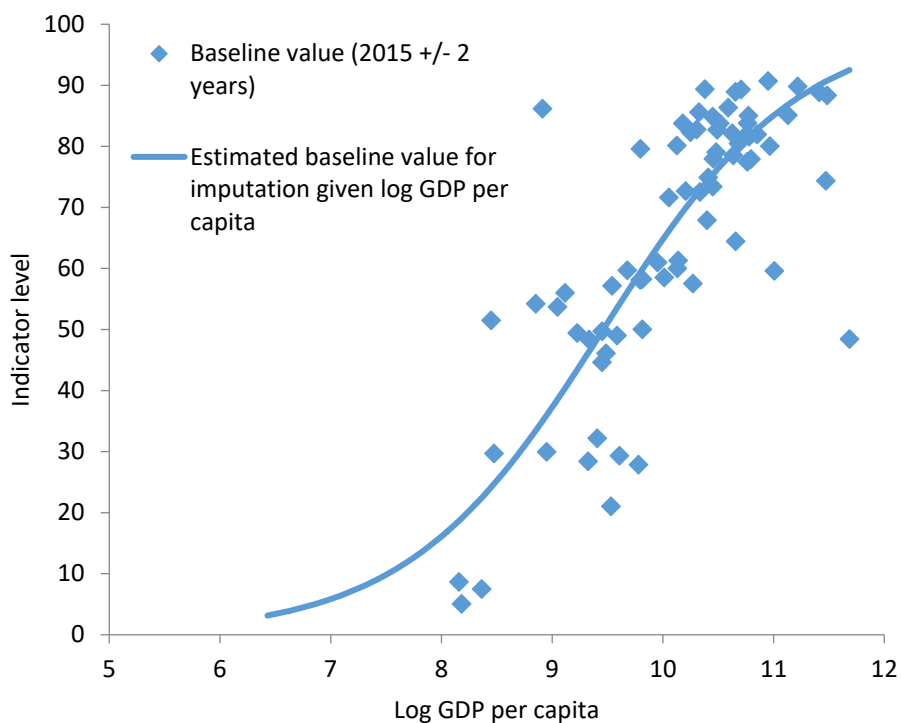
Regional benchmarks

Applying benchmarks to all countries within a region requires a measure that balances feasibility for low performing countries and relevance for high performing countries. This is particularly challenging in regions with a wide range of outcomes. The proposed approach was to define a minimum regional benchmark for 2030 for all countries in a particular region should achieve. To do this, the school-aged population-weighted average of the country-specific minimum benchmarks for the lowest tercile (third) of countries was used, including country-specific minimum benchmarks based on imputed baselines described previously. This acts as a minimum acceptable level for countries within a region.

Table D.1. Methodology specific to each indicator

Indicator	Definition of progress	Method of modeling the historical relationship between level and progress	Definition of high achievers
4.1.1. reading grades 2/3	percentage point change	percentile conditional on level	90
4.1.1. reading primary	percentage point change	absolute percentiles (unconditional)	90
4.1.1. reading lower sec.	percentage point change	percentile conditional on level	85
4.1.1. math grades 2/3	percentage point change	absolute percentiles (unconditional)	90
4.1.1. math primary	percentage point change	absolute percentiles (unconditional)	90
4.1.1. math lower sec.	percentage point change	percentile conditional on level	90
4.1.4 completion rate primary	growth rate of indicator	percentile conditional on level	90
4.1.4 completion rate lower secondary	growth rate of indicator	percentile conditional on level	90
4.1.4 completion rate upper secondary	growth rate of indicator	percentile conditional on level	90
4.1.4 completion rate primary (GEMR estimated data)	growth rate of indicator	percentile conditional on level	95
4.1.4 completion rate lower secondary (GEMR estimated data)	growth rate of indicator	percentile conditional on level	95
4.1.4 completion rate upper secondary (GEMR estimated data)	growth rate of indicator	percentile conditional on level	90
4.1.5 out-of-school rate primary	percentage point change	percentile conditional on level	10
4.1.5 out-of-school rate lower secondary	percentage point change	percentile conditional on level	10
4.1.5 out-of-school rate upper secondary	percentage point change	percentile conditional on level	10
4.2.2 pre-primary participation rate	percentage point change	percentile conditional on level	90
4.c.1 percent of teachers with minimum qualification pre-primary	growth rate of indicator	percentile conditional on level	90
4.c.1 percent of teachers with minimum qualification primary	growth rate of indicator	percentile conditional on level	90
4.c.1 percent of teachers with minimum qualification lower secondary	growth rate of indicator	percentile conditional on level	90
4.c.1 percent of teachers with minimum qualification upper secondary	growth rate of indicator	percentile conditional on level	90
1.a.2 percent of government expenditure on education	percentage point change	absolute percentiles (unconditional)	90
Government expenditure on education (% of GDP)	percentage point change	absolute percentiles (unconditional)	90