Education Finance Watch 2023: Technical Note

Introduction

This note, which explains the methodology of Education Finance Watch (EFW) 2023 in its presentation of indicators and estimates for **total global education spending,** introduces the methodology used for EFW2023 sections 1 (global education spending trends), 2 (government spending), and 7 (data spotlight: monitoring education spending). The note is partitioned into three primary sections for clarity and thoroughness. Section 1 describes the data used in the EFW report, the sources that we used, and the steps undertaken to process and consolidate the information used to develop the indicators presented in EFW report. Section 2 explains the methodology used to augment the data series by combining additional sources and imputing missing values, thereby facilitating the computation of global estimates. The methodology adopted for EFW2023 aligns with that used in EFW2022, although there are discrepancies between the data in EFW2022 and those in EFW2023. The reasons for these discrepancies are explored in Section 3. Appendixes A and B provide details about the database before and after we fill in missing data. Appendix C defines key terms used in the technical notes. At last, Appendix D details the methodology used in EFW section 6 (spotlight: demographic shifts).

Section 1: EFW Database

The EFW2023 database is a product of a collaborative effort of the World Bank; the Global Education Monitoring (GEM) Report; and the United Nations Educational, Scientific and Cultural Organization (UNESCO) Institute for Statistics (UIS). To maximize data coverage for the estimates of global education spending, the EFW2023 used multiple data sources, including the UIS database, World Bank BOOST database, World Bank World Development Indicators (WDI), International Monetary Fund (IMF) Government Finance Statistics (GFS) and World Economic Outlook (WEO), and International Development Statistics (IDS) of the Organisation for Economic Co-operation and Development (OECD). Table 1 summarizes the information taken from these various sources. Each of the datasets is discussed in Subsection 1.1.

Table 1. Summary of Data Sources and Key Information in the EFW2023 Database

Organization	Dataset	Available information	Method of access
United Nations Educational, Scientific, and Cultural Organization	Institute for Statistics	Government education expenditures Enrollment and school-aged population	Downloaded from website bulked data in April 2023: http://data.uis.unesco.org/
	GEM Report (compiled from household budget survey reports)	Household education spending	Obtained from GEM Report
World Bank	BOOST	Public education budget and expenditure	Shared by BOOST team
	World Development Indicators	Economic data (GDP, GDP per capita, official exchange rate)	Accessed World Bank databases using Stata module: -wbopendata-
	Human Capital Index database	Learning-adjusted years of school, expected years of	Downloaded from https://databank.worldbank.org/home

		schooling, harmonized learning outcomes	
International Monetary Fund	Government Finance Statistics	Public education expenditure, total government expenditures	Downloaded from https://data.imf.org/?sk=5804c5e1-0502-4672-bdcd-671bcdc565a9
	World Economic Outlook	Economic data (GDP, GDP per capita, GDP deflators)	Downloaded from https://www.imf.org/en/Publications/WEO/weo-database/2023/April
Organisation for Economic Co-operation and Development	International Development Statistics database	Aid to education according to donor and recipient	Downloaded from https://stats.oecd.org/
United Nations Department of Economic and Social Affairs	World Population Prospects	Population according to age	Downloaded from https://population.un.org/wpp/

Note: The data source for household education spending was the same as in 2022. Household education spending data was provided by Global Education Monitoring (GEM) Report team based on national household budget survey reports and UIS and OECD data.

Data in the EFW database were accessed using various platforms. Key education finance indicators such as government educational expenditures as a share of gross domestic product (GDP) and as a proportion of total government expenditures were all accessed through UIS official website under the Bulk Data Download Service button.[1] The World Bank BOOST team provided the expenditure data from its database[2] for use in the analysis. The World Bank WDI were accessed using the Stata module.[3] Fiscal and economic information from the IMF WEO and GFS was downloaded from the IMF website. Data on aid to education were retrieved from the OECD Creditor Reporting System.

In this technical note, it is important to distinguish two sets of data. One set comprises the actual data reported by the different institutions, including UIS, World Bank BOOST, and IMF GFS. The second set contains a new set of indicators comprising the actual data—as well as imputed values (where missing) using growth rate trends from World Bank BOOST and IMF GFS. This new set of indicators is referred to as "EFW estimates."[4] Section 2 is a detailed discussion of the process of combination and imputation, including analyses and other data checks.

The rest of Section 1 is organized as follows. Subsection 1.1 briefly discusses the sources of information used. Subsection 1.2 describes the primary variable of interest: public education spending data. Subsection 1.3 explains the steps taken to process the aid data to calculate total aid to education. Subsection 1.4 provides information on consolidation and rebasing of GDP data, which are used to estimate government and household education spending in constant 2021 U.S. dollars. Subsection 1.5 explains other key indicators presented as tables and figures in the EFW2023.

[1] To enrich the dataset for key variables, UIS data were sourced directly from UIS official website, unlike in 2022, when data were obtained from the SDXM Stata package. This is the website: https://apiportal.uis.unesco.org/bdds

[2] The BOOST database contains information on disaggregated and aggregated spending on various sectors, not just education. For the purpose of the EFW2023 analysis, only education-related spending data were used.

[3] The -wbopendata- module allows users to retrieve data from the World Bank database using Stata. For more information, see https://blogs.worldbank.org/opendata/new-release-wbopendata-stata-module.

[4] In using the EFW2023 database, refer to the accompanying file "EFW2023 Metadata and Variables List," which distinguishes the two sets of information and their respective data sources.

1.1 Sources of Education Financing Data Included in the EFW Database

This subsection describes the data sources, including the method of data collection.

UNESCO. Key education finance indicators such as government education expenditures as a share of GDP and as a proportion of total government expenditures, government education expenditure data expressed in current and constant U.S. dollars, constant purchasing power parity (PPP) international dollars; household spending, and other key education statistics such as student enrollment and school-age population were downloaded from the UIS website. The data were taken from annual financial reports of the Ministry of Finance and the Ministry of Education and national accounts reports from the National Statistical Office.

World Bank. Data on education budget and expenditures were obtained from the BOOST database and include indicators on government education spending as a percentage of GDP and as a percentage of total government expenditures. The BOOST program is a World Bank collaborative effort launched in 2010 to provide access to micro fiscal data and to make well-structured and highly disaggregated budget data available to policy makers, researchers, and the general public to improve budgetary decision making and analysis and increase transparency and accountability. Since its launch, the program has been used in approximately 100 countries and a large variety of corporate analyses, including public expenditure reviews, fiscal monitoring activities, and cross-country benchmarking. The economic data pulled from the WDI were sourced from its database of national accounts and OECD national accounts data files and have been estimated to be consistent with the United Nations (UN) System of National Accounts (SNA). These economic data have varying currency values, including current and constant U.S. dollars, local currencies (countries have different reference years), and 2017 PPP international dollars. Data were accessed using Stata module -wbopendata- in May 2023. Data from the Human Capital Project, which included learning outcomes indicators, were downloaded in May 2023.

International Monetary Fund. The IMF GFS contains information on government expenditures, including disaggregated expenditures on education according to educational level. The data are primarily obtained from responses to a detailed questionnaire distributed to government finance statistics correspondents, usually in the ministries of finance or central banks of reporting countries.[5] The countries report government financial information in the framework of the 2014 Government Finance Statistics Manual (GFSM 2014). BOOST and IMF GFS data were used for imputation. Expenditure data in the database are expressed in terms of local currency and percentage of GDP and include budgetary central government, central government, and general government expenditures. [6] General government expenditures, not budgetary central government or central government expenditures, were used for data analysis.

The WEO database contains selected macroeconomic indicators such as GDP (current and constant local currency, current U.S. dollars, current international dollars), GDP per capita (current and constant local currency, current U.S. dollars, current and constant PPP 2017 international dollars), GDP deflator (based on the difference between current and constant local currency GDP data), total government expenditures as a proportion of GDP (percentages), and population (in millions).

[5] Correspondents are data reporters from member countries who their country coordinator nominates or approves to submit data electronically via the Integrated Collection System (http://datareportinghelp.imf.org/knowledgebase/articles/904623-correspondent-coordinator-definitions).

[6] Based on the IMF GFS Manual (GSFM 2014), budgetary central government expenditures form part of central government expenditures, which are a subset of general government expenditures, although reporting protocols of countries in the database have differed, with some countries reporting data on budgetary central government expenditures but reporting nothing on central or general government expenditures. In some cases, there is information on central government expenditures but not general government expenditures.

The GFS and WEO datasets were downloaded from their respective websites in May 2023. According to the May 2023 WEO metadata, GDP data are from official statistics on national accounts, and the valuation follows the UN SNA or the European System of Accounts.[7]

The IMF team made certain estimations to prevent disruptions in data continuity and has projected data up to 2025. The dataset includes a variable called *estimatesstartafter* to indicate commencement of these forecasts. In EFW2023, estimated data starting from 2021 were selectively used to ensure that most of the GDP data consisted of actual values.

OECD. The OECD maintains the IDS databases, which record information on aid flows the OECD Development Assistance Committee (DAC) member countries, and non-DAC member countries provide annually. The database contains project- and activity-level aid flows of official development assistance (ODA)—public funds provided to developing countries to promote economic and social development. It is concessional, taking the form of grants or loans with lower interest rates than are available on the market and, usually, longer repayment periods. The database also contains other official flows (OOF), defined as official sector transactions that are excluded from ODA. OOF are grants to developing countries for representational or essentially commercial purposes or official bilateral transactions intended to promote development but having a grant element of less than 25 percent. Developing countries are those on the DAC list of ODA recipients. Bilateral donors are countries that provide development assistance directly to recipient countries. They also contribute substantially to the financing of multilateral donors through contributions recorded as multilateral ODA. Multilateral donors are international institutions with government membership that conduct all or a substantial part of their activities in favor of developing countries. They include multilateral development banks (e.g., World Bank, Inter-American Development Bank), UN agencies, and regional organizations (e.g., European Commission). Development banks also make non-concessional loans to middle- and high-income countries; these are not counted as part of ODA. The indicators downloaded from the IDS are total aid, total budgetary support, total aid for education, and aid for education according to level—all expressed in constant 2020 U.S. dollars, which were later converted to constant 2021 U.S. dollars. The data were downloaded in May 2023, and recipient countries and territories are the unit of observation.[8]

United Nations Department of Economic and Social Affairs. The World Population Prospects contains information on estimated total population per country, disaggregated according to age, and covers 2012 to 2021. The estimates are calculated using demographic variables such as fertility; child, adult, and overall mortality; and international migration using country-specific official statistics. Only the school-aged population, assumed to be aged 5 to 24 (to cover pre-primary through tertiary education), was included in the database. The dataset was downloaded in May 2023.

1.2 Public Education Spending

The primary indicator of interest is public education spending. UIS data on public education spending are expressed in many forms and units, including public education spending as a percentage of GDP, public education spending as a percentage of total government expenditures, and government spending per student relative to GDP per capita (in 2021 constant dollars). The indicator with the widest coverage is public education spending as a percentage of GDP, which makes it the best indicator to use in estimating global public education spending. Conditioned on the availability of GDP data, we can estimate spending in levels and not just as a ratio. As such, GDP data availability must be maximized using two sources, as detailed in Subsection 1.4.

[7] The methodology varies according to country.

[8] The downloaded data include aggregates such as "developing countries, total" and "developing countries, unspecified." The country list of "developing countries, total" include countries and territories that are not part of the EFW database. The data were downloaded from https://stats.oecd.org/Index.aspx?DataSetCode=CRS1.

The EFW2023 has benefitted from the participation of the UIS, the custodian agency of data on Sustainable Development Goal 4. In its efforts to bridge the gaps on education expenditure data, the UIS conducted a data-mapping and -validation exercise, the output of which is a dataset on the key indicators: government education spending as a share of GDP and share of education in total government expenditures that combines other available sources, including the World Bank BOOST initiative and IMF GFS. The specific methodology to substitute data from IMF GFS and BOOST for missing UIS data is detailed in Section 2. The data were also validated to ensure that the combined series were consistent with each other.

1.3 Aid to Education

OECD classifies aid to education into four levels: basic, secondary, post-secondary, and unspecified. Basic education covers early childhood, primary, and lower secondary education; basic life skills; school meals; and primary education equivalent for adults. Secondary education includes upper secondary education and vocational training. Post-secondary education includes advanced technical and managerial training. Unspecified refers to any activity that cannot be attributed solely to the development of a particular level of education, such as education research and teacher training. General education program support is often reported within this subcategory. The sum of these aid expenditures is equivalent to reported total direct aid to education in the IDS database.

1.4 Gross Domestic Product

Because the key education financing indicator is expressed as a share of GDP, GDP data are needed to calculate education spending. To maximize data coverage, two sources of information were used: World Bank WDI and IMF WEO. The base data were from the World Bank WDI, with data from the IMF WEO substituted for missing WDI data. .

The GDP data series combines information from WDI and WEO series rebased to reflect 2020 constant U.S dollars. The rebasing steps are as follows.[9] For the WDI GDP data, which were in constant 2015 U.S. dollars, the first step was to create an index by dividing GDP (in constant values) in a given year by the 2021 GDP constant value. Therefore, the index value for 2021 equals 1. For example, if the real GDP of Aruba in 2020 was valued at $26,100,038,937.564 (2015 US$), then dividing it by the real GDP in 2021 at $3,126,019,399.065 (2015 US$) the index in 2020 would be equal to 0.84. The second step was to multiply the index by the current value of the GDP in 2021. Again, using the Aruba case, the index in 2020 (0.84) will be multiplied by the nominal GDP in 2021 of $3,126,019,399.065 (current US$), which will result in the rebased value of $3,126,019,399.065 (constant 2021 US$) for 2009. For the WEO GDP data, a slightly different approach was used because the database contained GDP in constant local currency rather than constant U.S. dollars. In the absence of real GDP in constant U.S. dollars, the index was computed using real GDP in local currency. After this, the index was multiplied by the 2021 current U.S. dollar GDP value. Checks were performed for countries with data available for both WEO and WDI to determine whether there were any discrepancies in the rebased series. The discrepancies were close to zero, possibly because of rounding because the WEO GDP data are presented in billion dollars.

1.5 Indicators Presented in the EFW2023

The EFW report draws from different data sources.

[9] The method of rebasing was based on the World Bank methodology of rebasing GDP series. This method ensures that the growth rates exhibited in the constant local price series are preserved. Source: What is your constant U.S. dollar methodology? World Bank Data Help Desk

Global education spending. Figures 2a, 2b, and 3 in the EFW2023 report present total expenditures on education services from the main sources of education financing—governments, households, donors—expressed in constant 2021 U.S. dollars. Each country's annual government educational expenditures are computed by multiplying the percentage of government expenditures on education by GDP and subtracting development aid.

Government education spending per capita. This indicator is presented in figures 5a, 6b, and 7a in the EFW2023 report. It highlights government spending in relation to the number of beneficiaries: children who should receive education, regardless of their enrollment status. To calculate for this indicator, the first step was to multiply the estimates of public education spending as a percentage of GDP by the corresponding GDP (2021 US$) of that country using information from the WDI supplemented by IMF WEO where missing. The result was the value of total public education spending (2021 US$). Next, this was divided by the school-aged population. The number of children enrolled in pre-primary through tertiary school from UIS was used as the school-aged population. When this information was unavailable, data on the school-aged population (aged 5–24) from the United Nations Department of Economic and Social Affairs was used to fill the gap. Using information on the school-aged population and GDP, government spending per capita on education was calculated.

Government education spending per capita as a share of GDP per capita. This is an indicator presented in figures 5b and 7b in the EFW2023 report to describe the proportion of overall economic resources dedicated to educational spending. The source of this indicator is the combination of initial government funding per pre-primary, primary, secondary, and tertiary student as a percentage of GDP per capita from the UIS database. A substantial number of these four indicators have missing values, leading to noticeable data gaps in figures 5b and 7b in the EFW2023.

Prioritization of education in government expenditures. Figure 4 of the EFW2023 report plots the indicator on government expenditures as a share of GDP and the indicator showing the proportion of total government expenditure dedicated to education, as of 2021. Benchmarks are provided to indicate which countries are within targets. The commonly used benchmarks for education spending are 4 percent to 6 percent of GDP and 15 percent to 20 percent of total government expenditures.[10]

As previously mentioned, information on public education spending as a percentage of total government expenditures is initially extracted from UIS data. When UIS data is unavailable or needs verification, either IMF GFS or BOOST data is used. This mirrors the process used to calculate public education spending as a percentage of GDP. Detailed discussion on the process of combination and imputation is provided in Section 2.

Aid for education. The estimates shown in the EFW2023 report are based on OECD ODA data for education and general budget support. Total aid to education includes direct aid to education plus an imputation of 20 percent of aid devoted to general budget support (aid provided to governments not earmarked for specific projects or sectors). Direct aid to education is earmarked for education spending.

The approaches used to impute values are detailed in Section 2. Meanwhile, information on total aid to education is complete for 2012 to 2021. Aid to education can be identified according to recipient country, so it is straightforward to identify aid spending for each income group, although some aid values are not specified at the country level. To address this, each of the unspecified allocations per year was distributed to the income groups using a percentage provided in the GEM Report.[11] Because a portion of aid is already

[10] See the recommendations of the Education 2030 Framework for Action (https://unesdoc.unesco.org/ark:/48223/pf0000245656.locale=en).

[11] The values are based on trends in distribution of unspecified aid upon review of the microdata for 2016 to 2020. Most unspecified aid went to the Global Partnership for Education, the United Nations Children's Fund, UNESCO, and publicly funded research institutions, whose allocation of aid to income groups closely reflect this assumption.

reflected in government spending, 60 percent of total aid to education was deducted from each recipient country's government spending on education before aggregating for income group and global estimates.[12]

The process of converting current aid values to constant values uses a different methodology than that used for GDP conversion. The exchange rates used are annual averages of daily spot rates against the dollar, as calculated by the OECD Economics Directorate. The inflation figures used are the OECD's series of GDP implicit price deflators for each currency, with par set at the base year. To summarize, the DAC deflators convert dollar-denominated data for any year to dollars with the purchasing power they had in a specified base year (in this case, 2006). Converting a flow expressed in current dollars from country X in year Y into 2006 dollars requires dividing it by the 2006-base deflator for country X and year Y and multiplying by 100.

It was not possible to rebase the values for unspecified aid because of the lack of a reference country for rebasing, so current currency values were used for calculations.

Total government expenditures as a percentage of GDP. Indicators from the IMF WEO data were initially used to calculate total government expenditures as a percentage of GDP. When IMF WEO data is not available, GFS and combination of UIS and WDI data is used. Government expenditures are calculated based on GDP and total government expenditures as a percentage of GDP.

Learning outcomes and changes in education spending. In figure 16 of the EFW2023 report, real changes in average government education spending between 2017 and 2021 are plotted against learning outcomes. The indicator for learning outcomes is learning-adjusted years of schooling, downloaded from the World Bank Human Capital Project, which is a metric that combines components of quantity (expected years of schooling) and quality (harmonized learning outcomes). Expected years of schooling measures the number of years of school a child born today can expect to complete by age 18. It is based on age-specific enrollment rates between ages 4 and 17 and has a maximum value of 14. Harmonized learning outcomes are calculated using a conversion factor. For more details on the methodology, see Filmer and colleagues (2020), Kraay (2019), and Patrinos and Angrist (2018).

Household spending on education. The EFW2023 report relies on the same methods to estimate global household education spending as EFW2022. The EFW2022 report used information from the GEM Report from various published reports. Except for countries in the OECD and UIS databases, where the indicator was readily available, household education expenditures as a share of GDP was estimated as the product of two indicators: reported share of education in total household consumption expenditures (from national budget survey reports) and share of household consumption in GDP (from WDI). Most countries had one estimate in the 2010s. If multiple data points were available, the latest value (in countries for which there were annual data) or an average of available data points (in countries with sparse data) was used. Household education expenditures as a share of GDP was multiplied by absolute GDP in U.S. dollars to be used in calculations of share of households spending in total education spending. Because data availability was limited, one estimate per country refers to the entire decade. The average of the respective income group was imputed for countries without data. Because updated household spending data for 2021 were unavailable, it was assumed that household spending would increase proportionally to global GDP. Using this ratio, household spending data for 2021 were estimated.

[12] See UNESCO (2012) EFA Global Monitoring Report (Box 2.1, page 145) for a discussion of these assumptions. https://unesdoc.unesco.org/ark:/48223/pf0000218003

Section 2: Extending the Data Using Combination and Imputation

To estimate aggregated global education spending each year from 2012 to 2021, it is important to ensure that as many countries as possible are included and the composition of countries is consistent throughout. Data availability varies from year to year, so values have been imputed to fill gaps.

The EFW2023 has benefitted from the participation of the UIS, the custodian agency of data on Sustainable Development Goal 4. The UIS, in its efforts to fill the gaps on education expenditure data, conducted a data mapping and validation exercise, the output of which is a dataset on the key indicators: government education spending as a share of GDP and share of education in total government expenditures that combines available sources, such as the World Bank BOOST initiative and IMF GFS. The data were also validated to ensure that the combined series are consistent with each other.

For EFW2023, the same list of 218 countries used in EFW2022 was used. The most current World Bank country list, as of May 2023, was used as the key reference (https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups). Four countries (Anguilla, Cook Islands, the Holy See, Montserrat), which were included in the UIS but not in the 218 country and territory list, were excluded. Other countries which were not present in the World Bank database but found in others, such as Saint Barthélemy and Niue, were also removed.

With the availability of more historical data after imputation, the EFW2023 report includes data on key indicator educational expenditures as a percentage of GDP for 189 countries and territories. A detailed description of the 10-year data is in Table 2.

Table 2. Summary of UIS Data on Government Education Spending as a Percentage of GDP

	Percentiles	Smallest		
1%	1.05227	0		
5%	1.737109	5.67e-06		
10%	2.14504	8.02e-06	Obs	4,746
25%	3.04541	8.09e-06	Sum of wgt.	4,746
50%	4.18938		Mean	4.349447
		Largest	Std. dev.	1.977375
75%	5.38056	15.07053		
90%	6.48108	15.73922	Variance	3.91001
95%	7.42208	22.32221	Skewness	2.850042
99%	10.93203	44.33398	Kurtosis	41.82471

The data series that the UIS provided consolidates information from various sources, but much information is missing if the 218 countries and territories included in the EFW database are used. The proportion of countries with no available data ranges from 23 percent to 60 percent (Table 3) for key indicator educational expenditures as a percentage of GDP. The highest percentage of missing values is for 2021.

Table 3. Distribution of Data on Government Education Spending as a Percentage of GDP

Year	Number of countries with data	Missing (%)
2012	154	29
2013	157	28
2014	160	27
2015	167	23
2016	165	24
2017	168	23
2018	164	25
2019	155	29
2020	106	51

2021	88	60
Average		32

Note: Proportion of missing values is computed as 218 minus the number of countries with data divided by 218.

We observe abnormal data points from Cuba and Somalia, which are so small that they could be very influential in the results. They were also compared with data from BOOST and IMF, and a large difference was found, so it was decided to drop data points from these two countries, which is discussed further in Appendix A.

Table 4 indicates the number of countries and territories with no data for the whole of 2012 to 2021, with 6 to 9 years of missing values between 2012 and 2021 (more than 50 percent missing), and with 3 to 5 missing years (25 percent to 50 percent missing). Some of those with 25 percent to 50 percent missing may make significant contribution to global public spending, such as China.

Table 4. Number of Countries with Missing Data for 2012–2021 According to Number of Years for Which Data Are Missing

Number (percentage) of years	Number of countries
10 (100%)	0
6–9 (>50)	16
3–5 (25–50)	33

The validation exercises that UIS conducted meant that some data points from the other sources were not combined with the UIS data. Therefore, information from other sources, specifically from IMF and BOOST, can be used to expand coverage of the key indicator and increase the robustness of the estimates. To ensure consistency and not negate the previous validation effort of the UIS, only trends or growth rates were used (not actual data on government education spending as a percentage of GDP in the IMF or BOOST database) to interpolate or extrapolate missing information. In summary, the following steps were undertaken to combine the UIS data with the BOOST or IMF data.

1. **Fill in missing values using country growth trend.** Growth rates from BOOST or IMF (whichever was closer to the UIS data on average between 2012 and 2021) were applied to the UIS data series to calculate missing values.
2. **Fill in missing values using income level growth trends.** Average growth rates for each income level were calculated and applied to the series to fill the remaining gaps.

In 2023, UIS has already combined part of IMF data with their UIS data published on the official website, so fewer added data points are expected from IMF than from BOOST.

The following subsections discuss in more detail these two steps, including analyses and checks to ensure consistency.

2.1 Fill in missing values using country growth trend

Because trends or growth rates will be used, coverage of the dataset should be from 2011 to 2021 so that the growth rate between 2012 and 2021 for those observations with missing values for 2021 can be applied.

Before data from the other sources are combined with UIS data, each country is categorized into one of these three types:

- Type 1: IMF and BOOST data are available.
- Type 2: IMF or BOOST data are available.
- Type 3: No UIS data are available, but at least one data point is available from IMF or BOOST.

For Type 1, where both IMF and BOOST data are available, the data with a growth rate closer to that of the UIS was used. To determine whether IMF or BOOST is closer, the average of the UIS series from 2012 to 2021 was taken and compared with the average value of the BOOST or IMF series from the same period. To compare, the absolute difference (regardless of whether the difference is positive or negative) between the series was calculated.

For example, Afghanistan's average education spending as a percentage of GDP between 2012 and 2021 was 3.0 percent according to the UIS, 4.5 percent according to the IMF, and 3.5 percent according to BOOST. The differences (UIS minus IMF; UIS minus BOOST) are 1.5 and 0.5 percentage points, respectively, indicating that the gap between and UIS and BOOST is narrower than that between UIS and IMF. As such, growth rates from BOOST will be applied to calculate missing values in Afghanistan.

Applying the growth rates from the other sources for Type 1 countries, 30 additional data points were added (Table 5)—15 from IMF and 15 from BOOST.

Table 5. Number of Additional Data Points According to Type 1

International Monetary Fund				BOOST		
year	Freq.	Percent	Cum.	Freq.	Percent	Cum.
2018	1	2.70	62.16	1	6.67	6.67
2019	2	5.41	67.57	1	6.67	13.33
2020	6	16.22	83.78	9	60.00	73.33
2021	6	16.22	100.00	4	26.67	100.00

For Type 2, for which only one of the other data series (IMF or BOOST) is available, the trend is used when available. This is straightforward because the growth rates from the available alternative source are simply applied. Applying the growth rates from that alternative source, 143 additional data points were added (Table 6)—53 from IMF and 90 from BOOST.

Table 6. Number of Additional Data Points According to Type 2

International Monetary Fund				BOOST			
year	Freq.	Percent	Cum.	year	Freq.	Percent	Cum.
2020	26	20.31	78.91	2012	12	9.23	40.00
2021	27	21.09	100.00	2013	8	6.15	46.15
				2014	8	6.15	52.31
				2015	7	5.38	57.69
				2016	7	5.38	63.08
				2017	7	5.38	68.46
				2018	5	3.85	72.31
				2019	8	6.15	78.46
				2020	12	9.23	87.69
				2021	16	12.31	100.00

For Type 3, countries with no UIS data but at least one data point from IMF or BOOST, the series with at least one data point or the series with the most data points is used. Average values for 2012 through 2021 are presented in Table 7[13]. Only Kosovo has information from IMF and BOOST, and the series with more data points (BOOST) was used.[14] After using data from IMF or BOOST when available, 10 observations can be added (Table 8).

[13] Iso3 in the table refers to the abbreviation of the country's name.
[14] The BOOST database has data for Kosovo from 2012 to 2021, whereas the IMF database has data only from 2016 to 2021, so it makes more sense to use the former.

Table 7. Countries with no United Nations Educational, Scientific and Cultural Organization Institute for Statistics Data but at Least One Data Point from International Monetary Fund or BOOST

iso3	countr~e	meanuis	meanimf	meanbo~t
XKX	Kosovo	.	4.428911	4.349429

Note: meanuis = Mean of values from UIS database; meanimf = Mean of values from IMF GFS database; meanbo~t = Mean of values from BOOST database.

Table 8. Number of Additional Data Points According to Type 3

YEAR	International Monetary Fund	BOOST	TOTAL
2012	0	1	1
2013	0	1	1
2014	0	1	1
2015	0	1	1
2016	0	1	1
2017	0	1	1
2018	0	1	1
2019	0	1	1
2020	0	1	1
2021	0	1	1
TOTAL	0	10	10

Combining all three sources has increased the number of countries and territories with data.[15] Table 9 shows the distribution of data points from each source and total data points after combining the data sources, although gaps remain. Twenty-four percent of data is missing given our benchmark panel data of 218 countries and territories from 2012 to 2021, which is lower than 32 percent missing data when using the UIS data series (Table 3). The most (76 countries) and highest percentage (35 percent) of countries with missing values is in 2021.

Table 9. Number of Additional Data Points According to Source

Year	UIS[a]	International Monetary Fund	BOOST	Total[b]	Ideal[c]	No DATA[d]	MISSING (%)[e]
2012	154	0	13	167	218	51	23
2013	157	0	9	166	218	52	24
2014	160	0	9	169	218	49	22
2015	167	0	8	175	218	43	20
2016	165	0	8	173	218	45	21
2017	168	0	8	176	218	42	19
2018	164	1	7	172	218	46	21
2019	155	2	10	167	218	51	23
2020	106	32	22	160	218	58	27
2021	88	33	21	142	218	76	35
Average							24

a. Consolidated data from mapping and validation exercise done by UIS.
b. All observations with available data.
c. Total number of observations ideally covered in Education Finance Watch database.
d. Calculated by subtracting "Total" from "Ideal."
e. Calculated by dividing "No Data" by "Ideal" number of observations.

[15] See Appendix A for the more detailed explanation on imputation process and diagnostic results.

2.2 Fill in Missing Values Using Income Level Trends

To increase coverage further, missing values must be imputed through interpolation or extrapolation. Imputation is conducted by using the average growth rate of countries within a specific income level to estimate missing values for other countries of the same income level. For example, government education spending as a percentage of GDP for low-income countries increased by 3.7 percent from 2020 to 2021, so 3.7 percent is used to calculate 2021 values for low-income countries with missing information.

Please note that our assumption may not always hold true and should be interpreted with caution. The assumption is that a country follows the same trend as the other countries in its income level, but there might be differences in the composition of countries and territories in income categories from year to year, which might lead to an under- or over-estimation of the actual trend considering the country with available data may not be representative for all countries in this income level. In addition, income level averages are computed based on UIS data and on imputed values based on trends from IMF or BOOST, as explained in Subsection 2.1.

First, average government education spending is estimated as a percentage of GDP in each income group (Table 10), and then year-on-year growth rates are generated according to income group (Table 11).

Table 10. Estimated Average Education Spending as a Percentage of GDP According to Income Group

YEAR	Low income	Lower middle income	Upper middle income	High income
		%		
2012	3.2	4.9	4.5	4.9
2013	3.2	4.9	4.6	5.0
2014	3.4	4.8	4.7	4.8
2015	3.4	4.7	4.8	4.8
2016	3.1	4.8	4.9	4.8
2017	3.2	4.9	4.5	4.5
2018	3.1	4.6	4.5	4.6
2019	3.1	4.8	4.4	4.6
2020	3.3	4.8	4.7	5.0
2021	3.4	4.5	4.7	4.7

Table 11. Estimated Rate of Growth of Education Spending as a Percentage of GDP According to Income Group

YEAR	Low income	Lower middle income	Upper middle income	High income
		%		
2012	-6.6	2.5	2.3	3.8
2013	0.2	0.1	3.9	2.2
2014	6.1	-0.9	2.4	-3.5
2015	0.7	-2.3	1.3	0.2
2016	-8.0	1.2	1.1	-1.7
2017	1.5	2.8	-6.6	-5.2
2018	-2.5	-6.4	-1.4	2.6
2019	1.4	4.8	-2.0	-1.3
2020	3.9	0.0	7.2	9.8
2021	3.7	-6.2	0.1	-6.0

Note: In calculating year-on-year growth rates, average values according to income group were used (see Table 10).

An additional 343 observations were added for 2012 to 2021 through this interpolation exercise, increasing the number of total observations to 2,020 (Table 12)[16] and reducing the proportion of missing from 24

[16] Country cases and diagnostics are discussed in Appendix B.

percent to 7 percent after the first imputation step (through combination) and quite significantly from 32 percent missing using the UIS data series only.

Table 12. Summary of Data Points Added After Step 1 and 2 Imputations

Year	Observations, n	UIS data only		Plus Step 1: Combination		Plus Step 2: Imputation	
		With data, n	Missing, %	With data, n	Missing, %	With data, n	Missing, %
2012	218	154	29	167	23	202	7
2013	218	157	28	166	24	202	7
2014	218	160	27	169	22	202	7
2015	218	167	23	175	20	202	7
2016	218	165	24	173	21	202	7
2017	218	168	23	176	19	202	7
2018	218	164	25	172	21	202	7
2019	218	155	29	167	23	202	7
2020	218	106	51	160	27	202	7
2021	218	88	60	142	35	202	7
Average			32		24		7

Note: UIS data series are from UIS bulked data.

Section 3 Comparing EFW2022 and EFW2023 Databases

Although the methodologies used in EFW2023 align with those from EFW2022, a difference in absolute numbers of total real education spending can be seen when comparing Figure 2 in EFW2023 with Figure 6 in EFW2022. This discrepancy can be attributed to three primary factors. First, eight countries and territories that lacked educational expenditure data in EFW2022 had data in EFW2023 because new data were available from UIS, IMF, and BOOST, or data for 2019 were updated, enabling data to be imputed for 2020 (as shown in Table 13).

Table 13. Comparison of EFW2023 and EFW2022 2020 Data for Type 1 Countries

Country	Education spending as percentage of GDP (UIS actual)	Education spending as percentage of GDP (after imputation)	GDP (2020 current US$ billion)	GDP (2020 constant US$ billion)	Educational expenditures (UIS actual, 2020 constant US$ billion)	Educational expenditures after imputation (2020 constant US$ billion)
2020 data in EFW2022						
American Samoa		13.8	0.7	0.7		0.1
Cayman Islands		2.3	5.6	5.8		0.1
Equatorial Guinea	0.4	0.4	10.1	12.4	<0.1	<0.1
Iraq		4.7	184.3	202.3		9.4
Libya		2.3	50.4	32.6		0.7
Nigeria		3.4	432.2	425.3		14.4
Norway	8.4	8.4	362.2	464.2	38.	38.9
Palau		7.7	0.3	0.3		<0.1
2020 data in EFW2023						
American Samoa			0.7	0.7		
Cayman Islands			5.6	5.6		
Equatorial Guinea			10.0	10.0		
Iraq			166.8	166.8		
Libya			25.4	25.4		
Nigeria			432.3	432.3		
Norway			362.2	362.2		
Palau			0.3	0.3		

Second, two countries with data in EFW2022 were absent from the EFW2023 data set (Kuwait, Yemen), and although Cuba's data were available for EFW2023, there was a substantial discrepancy from its EFW2022, so it was marked as missing to maintain data integrity (as shown in Table 14).

Table 14. Comparison of EFW2023 and EFW2022 2020 Data for Type 2 Countries

Country	Education spending as percentage of GDP (UIS actual)	Education spending as percentage of GDP (after imputation)	GDP (2020 current US$ billion)	GDP (2020 constant US$ billion)	Educational expenditures (UIS actual, 2020 constant US$ billion)	Educational expenditures after imputation (2020 constant US$ billion)
2020 data in EFW2022						
Cuba	0.0	0.0	107.4			
Kuwait						
Yemen, Rep.		5.6				
2020 data in EFW2023						
Cuba	12.7		107.4	107.4		13.7
Kuwait					6.9	6.9
Yemen, Rep.	6.1		18.8	18.8		1.1

Note: EFW, Education Finance Watch; GDP, gross domestic product; UIS, United Nations Educational, Scientific and Cultural Organization Institute for Statistics

Third, many countries and territories had data in EFW2022 and EFW2023, but the figures had been updated. For instance, there was a difference of more than $10 billion in the data for six countries. The reasons for these differences include replacement of imputed data with new data, replacement of actual data with imputed data, differing actual values, and the rebasing from 2020 to 2021. These changes can be seen in the alterations in current GDP, UIS's actual educational expenditures as a percentage of GDP, and the imputed share (as shown in Table 15).

Table 15. Comparison of EFW2023 and EFW2022 2020 Data for Type 3 Countries

Country	Education spending as percentage of GDP (UIS actual)	Education spending as percentage of GDP (after imputation)	GDP (2020 current US$ billion)	GDP (2020 constant US$ billion)	Educational expenditures (UIS actual, 2020 constant US$ billion)	Educational expenditures after imputation (2020 constant US$ billion)
2020 data in EFW2022						
China	3.6	3.6	14,687.7	16,403.8	586.1	586.1
Germany	5.4	5.4	3,889.7	4,150.9	225.2	225.2
United Kingdom	5.9	5.9	2,704.6	2,912.2	173.2	173.2
Indonesia	3.5	3.5	1,058.7	1,143.9	39.9	40.0
Korea, Rep		5.1	1,644.3	1,738.9		89.5
United States	5.0	5.0	21,060.5	22,006.7	1,110.9	1,111.0
2020 data in EFW2023						
China		3.5	14,722.7	14,722.7		521.6
Germany	4.7	4.7	3,846.4	3,846.4	179.2	179.2
United Kingdom	5.5	5.5	2,759.8	2,759.8	152.6	152.6
Indonesia	2.8	2.8	1,058.4	1,058.4	29.8	29.8
Korea, Rep.		4.4	1,637.9	1,637.9		72.9
United States		4.9	20,953.0	20,953.0		1,027,.0

Appendix A: Outlier Detection After Step 1 of Imputation

After combining IMF and BOOST data with UIS data to yield the variable ***educshare*** (indicating the educational expenditures as a percentage of GDP) following the steps and conditions in section 2, further analyses were conducted to check the outliers.

- **Checking for influential observations.** Checking the distribution of the combined data series, the 99th largest value is 12.08 and the 1st smallest value is 1.07 (Table A.1). We will then look at the data points below the 1st percentile and above the 99th percentile to check if there are abnormal outliers.

Table A.1. Description of the Indicator Government Education Spending as a Percentage of GDP (Combined)

Integrated: Gov't education spending, % of GDP

	Percentiles	Smallest		
1%	1.06738	0		
5%	1.74292	5.67e-06		
10%	2.15058	8.02e-06	Obs	5,069
25%	3.071543	8.09e-06	Sum of wgt.	5,069
50%	4.23424		Mean	4.412404
		Largest	Std. dev.	2.104539
75%	5.41381	22.32221		
90%	6.520573	22.77333	Variance	4.429084
95%	7.50288	27.0257	Skewness	3.076177
99%	12.07966	44.33398	Kurtosis	37.3147

- **Investigating the observations that are below the first percentile, six countries are observed.** Among them, Somalia and Cuba have abnormal data points (as shown in Table A.2.).

Table A.2. Data Points Below the 1st Percentile Distribution for the Indicator: Government Education Spending as a Percentage of GDP (Combined)

iso3	countryname	year	gov_educgdp	GF09__gdp_gg	boost_edugdp	educshare
ARE	United Arab Emirates	2012	1.3051802	.	1.2711788	1.30518
SSD	South Sudan	2012	1.2455282	.	.	1.245528
SSD	South Sudan	2013	1.0245	.	.	1.0245
MCO	Monaco	2014	1.02195	.	.	1.02195
SSD	South Sudan	2014	1.23246	.	.	1.23246
BMU	Bermuda	2017	1.32496	.	.	1.32496
GNQ	Equatorial Guinea	2017	.40774684	.	.	.4077469
SOM	Somalia	2017	.12717403	.12717403	4.426e-06	.127174
CUB	Cuba	2017	8.444e-06	.	.	8.44e-06
CUB	Cuba	2018	8.025e-06	.	.	8.02e-06
GNQ	Equatorial Guinea	2018	.3731156	.	.	.3731156
SOM	Somalia	2018	5.674e-06	.23445096	7.774e-06	5.67e-06
CUB	Cuba	2019	9.047e-06	.	.	9.05e-06
MCO	Monaco	2019	1.1589	.	.	1.1589
GNQ	Equatorial Guinea	2019	.60447216	.	.	.6044722
SOM	Somalia	2019	8.338e-06	.27066174	.	8.34e-06
SOM	Somalia	2020	8.088e-06	.	.	8.09e-06
GNQ	Equatorial Guinea	2020	.4469	.	.	.4469
CUB	Cuba	2020	.00001152	.	.	.0000115
SOM	Somalia	2021	9.111e-06	.	.	9.11e-06
GNQ	Equatorial Guinea	2021	.25290446	.	.	.2529045

- **Somalia, 2018-2021:** The numbers from the UIS data differ significantly from the IMF data. Therefore, values have been changed to 'missing data' to prevent undue influence on the final results (Table A.3.).

Table A.3. Somalia's Government Education Spending as a Percentage of GDP (Combined)

iso3	countryname	year	gov_educgdp	GF09__gdp_gg	boost_edugdp	educshare
SOM	Somalia	2017	.12717403	.12717403	4.426e-06	.127174
SOM	Somalia	2018	5.674e-06	.23445096	7.774e-06	5.67e-06
SOM	Somalia	2019	8.338e-06	.27066174	.	8.34e-06
SOM	Somalia	2020	8.088e-06	.	.	8.09e-06
SOM	Somalia	2021	9.111e-06	.	.	9.11e-06

- **Cuba, 2017-2019:** The numbers from 2017 to 2019 are too small and too different from those from previous years, so values have been changed to 'missing data' to prevent undue influence on the final results (Table A.4.).

Table A.4. Cuba's Government Education Spending as a Percentage of GDP (Combined)

iso3	countryname	year	gov_educgdp	GF09__gdp_gg	boost_edugdp	educshare
CUB	Cuba	2005	10.55778	.	.	10.55778
CUB	Cuba	2006	9.0571699	.	.	9.05717
CUB	Cuba	2007	11.86821	.	.	11.86821
CUB	Cuba	2008	14.05908	.	.	14.05908
CUB	Cuba	2009	13.12486	.	.	13.12486
CUB	Cuba	2010	12.83731	.	.	12.83731
CUB	Cuba	2011	.	.	.	.
CUB	Cuba	2012	.	.	.	.
CUB	Cuba	2013	.	.	.	.
CUB	Cuba	2014	.	.	.	.
CUB	Cuba	2015	.	.	.	.
CUB	Cuba	2016	.	.	.	.
CUB	Cuba	2017	8.444e-06	.	.	8.44e-06
CUB	Cuba	2018	8.025e-06	.	.	8.02e-06
CUB	Cuba	2019	9.047e-06	.	.	9.05e-06
CUB	Cuba	2020	.00001152	.	.	.0000115
CUB	Cuba	2021	.	.	.	.

Note: All variables are general government education spending as a percentage of gross domestic product. Sources are as follows: **gov_educgdp** = UIS; **GF09_gdp_gg** = IMF GFS; **boost_edugdp** = BOOST; **educshare** = combining UIS with IMF/BOOST following step 1 of imputation.

- **Investigating the observations that are above the 99th percentile, we find three countries.** As it is shown in table A.5., observations from BOOST data are not significantly different from the rest of the data in other years, so they were retained (Table A.5.).

Table A.5. Data Points Above the 99th Percentile Distribution for the Indicator: Government Education Spending as a Percentage of GDP (Combined)

iso3	countryname	year	gov_educgdp	GF09__gdp_gg	boost_edugdp	educshare
KIR	Kiribati	2012	13.384385	13.427666	13.512659	13.38438
MHL	Marshall Islands	2012	.	.	14.029198	14.50774
MHL	Marshall Islands	2013	.	.	14.626206	15.12511
MHL	Marshall Islands	2014	.	.	15.466999	15.99458
MHL	Marshall Islands	2015	.	.	15.692577	16.22786
KIR	Kiribati	2016	13.483116	10.083363	10.083363	13.48312
MHL	Marshall Islands	2016	15.05292	.	14.556396	15.05292
KIR	Kiribati	2017	13.600176	10.81031	10.81029	13.60018
MHL	Marshall Islands	2017	15.070533	.	15.654461	15.07053
KIR	Kiribati	2018	.	11.999222	11.999243	15.09597
MHL	Marshall Islands	2018	15.00368	.	15.470738	15.00368
SLB	Solomon Islands	2019	.	.	10.404278	13.48145
KIR	Kiribati	2019	.	13.574123	13.574129	17.0773
KIR	Kiribati	2020	.	12.352543	12.35255	15.54046
MHL	Marshall Islands	2021	15.739225	.	17.596504	15.73922

iso3	countryname	year	gov_educgdp	GF09__gdp_gg	boost_edugdp	educshare
KIR	Kiribati	2012	13.384385	13.427666	13.512659	13.38438
KIR	Kiribati	2013	12.26	12.063234	12.123603	12.26
KIR	Kiribati	2014	12.953124	11.348642	11.348645	12.95312
KIR	Kiribati	2015	12.114028	9.9108503	9.9108482	12.11403
KIR	Kiribati	2016	13.483116	10.083363	10.083363	13.48312
KIR	Kiribati	2017	13.600176	10.81031	10.81029	13.60018
KIR	Kiribati	2018	.	11.999222	11.999243	15.09597
KIR	Kiribati	2019	.	13.574123	13.574129	17.0773
KIR	Kiribati	2020	.	12.352543	12.35255	15.54046
KIR	Kiribati	2021	.	15.585125	.	.

iso3	countryname	year	GF09__gdp_gg	gov_educgdp	boost_edugdp	educshare
MHL	Marshall Islands	2012	.	.	14.029198	14.50774
MHL	Marshall Islands	2013	.	.	14.626206	15.12511
MHL	Marshall Islands	2014	.	.	15.466999	15.99458
MHL	Marshall Islands	2015	.	.	15.692577	16.22786
MHL	Marshall Islands	2016	.	15.05292	14.556396	15.05292
MHL	Marshall Islands	2017	.	15.070533	15.654461	15.07053
MHL	Marshall Islands	2018	.	15.00368	15.470738	15.00368
MHL	Marshall Islands	2019	.	9.9164801	17.527353	9.91648
MHL	Marshall Islands	2020	.	.	15.767972	8.921072
MHL	Marshall Islands	2021	.	15.739225	17.596504	15.73922

iso3	countryname	year	GF09__gdp_gg	gov_educgdp	boost_edugdp	educshare
SLB	Solomon Islands	2012	.	.	7.259079	9.406025
SLB	Solomon Islands	2013	.	.	8.9395329	11.58349
SLB	Solomon Islands	2014	.	.	8.6123831	11.15958
SLB	Solomon Islands	2015	.	.	8.9475581	11.59389
SLB	Solomon Islands	2016	.	.	9.2119755	11.93651
SLB	Solomon Islands	2017	.	.	10.210336	13.23014
SLB	Solomon Islands	2018	.	.	10.241141	13.27006
SLB	Solomon Islands	2019	.	.	10.404278	13.48145
SLB	Solomon Islands	2020	.	.	10.241274	13.27023
SLB	Solomon Islands	2021	.	.	8.9318972	11.57359

- **Checking rates of change in values for each country to detect unusual fluctuations, such as excessively high spikes or notably low dips.** We calculate growth rate of public education spending as percentage of GDP and check the distribution. We investigate countries for which the growth rates are below the 1st or above the 99th percentile cut-offs, because they are likely to be

outliers. We analyze the growth rate trend within each available data source (UIS, IMF, BOOST) to determine their believability and consistency (as shown in Table A.6.).

Table A.6. Description of the Growth Rate of Government Education Spending as a Percentage of GDP (Combined)

gr_educshare

	Percentiles	Smallest		
1%	-31.18136	-99.99554		
5%	-18.09018	-78.42928		
10%	-12.43102	-74.6967	Obs	4,307
25%	-4.978083	-69.05685	Sum of wgt.	4,307
50%	.0572698		Mean	1.367448
		Largest	Std. dev.	14.79196
75%	6.094066	118.3676		
90%	15.55512	155.5169	Variance	218.8022
95%	23.28199	167.0347	Skewness	2.442918
99%	48.21889	228.8621	Kurtosis	29.40395

Table A.7. Data Points Below the 1st Percentile Distribution for the Growth Rate of Government Education Spending as a Percentage of GDP (Combined)

iso3	countryname	year	GF09__gdp_gg	gov_educgdp	boost_edugdp	educshare	gr_educshare
COD	Congo, Dem Rep	2017	.	1.45495	1.5121341	1.45495	-31.32785
COG	Congo, Rep	2014	.	3.1837101	.	3.18371	-36.50208
COG	Congo, Rep	2018	.	2.9954801	.	2.99548	-32.69371
CZE	Czech Republic	2017	4.0647906	3.80616	4.0647906	3.80616	-31.4666
GHA	Ghana	2013	.	4.5764699	.	4.57647	-42.21001
GNQ	Equatorial Guinea	2021	.	.25290446	.	.2529045	-43.40916
KHM	Cambodia	2021	.	1.6705739	.	1.670574	-44.35779
KNA	St Kitts and Nevis	2015	.	2.54987	.	2.54987	-31.18136
MAC	Macao SAR, China	2013	1.9487294	2.0504899	.	2.05049	-38.3341
MHL	Marshall Islands	2019	.	9.9164801	17.527353	9.91648	-33.90635
NER	Niger	2016	.	2.9584301	3.6938608	2.95843	-34.06208
SOM	Somalia	2018	.23445096	5.674e-06	7.774e-06	5.67e-06	-99.99554
STP	Sao Tome and Principe	2014	.	3.7490599	.	3.74906	-36.88122
ZWE	Zimbabwe	2018	.	2.0504899	.	2.05049	-64.76083

All data are from the UIS as shown in Table A.7., so we accept the data.

Table A.8. Data Points Above the 99th Distribution for the Growth Rate of Government Education Spending as a Percentage of GDP (Combined)

	iso3	countryname	year	gov_educgdp	GF09__gdp_gg	boost_edugdp	educshare	gr_educshare
369	ARE	United Arab Emirates	2019	3.8602099	.	1.5420234	3.86021	155.5169
1974	CAF	Central African Republic	2012	1.9803472	.	.	1.980347	78.7446
2533	COG	Congo, Rep	2013	5.0138814	.	.	5.013881	83.03976
2536	COG	Congo, Rep	2016	5.2105161	.	.	5.210516	58.73234
3280	DMA	Dominica	2016	5.1297195	.	.	5.12972	50.49433
4331	GEO	Georgia	2013	3.4732218	3.4732218	2.8870602	3.473222	82.21613
4645	GNB	Guinea-Bissau	2017	.	.	2.1959611	2.571994	56.4564
4709	GNQ	Equatorial Guinea	2019	.60447216	.	.	.6044722	62.00667
6198	KAZ	Kazakhstan	2020	4.4462038	4.8142574	4.8098285	4.446204	55.62654
6504	KNA	St Kitts and Nevis	2016	3.9979768	.	.	3.997977	56.7914
6687	LAO	Lao PDR	2013	3.2338099	.	7.5547525	3.23381	77.53847
6750	LBN	Lebanon	2014	.	.	4.7996258	4.456492	83.09562
6944	LCA	St Lucia	2022	4.8989314	.	.	4.898932	61.33386
7062	LKA	Sri Lanka	2016	3.22929	.	.	3.22929	53.04836
7376	MAC	Macao SAR, China	2020	6.3603201	5.9892371	.	6.36032	107.9467
7740	MDV	Maldives	2012	4.968268	.	.	4.968268	61.26918
7873	MHL	Marshall Islands	2021	15.739225	.	17.596504	15.73922	76.4275
8112	MMR	Myanmar	2012	1.5864288	1.5864288	.70169187	1.586429	86.56844
10353	RWA	Rwanda	2021	5.5843801	.	4.1645637	5.58438	67.65702

Note: All these countries data sources are from UIS data except Guinea Bissau and Lebanon.

Two countries for which the pertinent observations did not originate from UIS data (Guinea Bissau and Lebanon) were more closely examined. The data for these countries were sourced from BOOST. In previous years, BOOST data have not varied much from UIS data. In addition, BOOST is the only source of these data, so these entries were accepted.

Table A.9. Guinea-Bissau's and Lebanon's Government Education Spending as a Percentage of GDP (Combined)

iso3	countryname	year	GF09__gdp_gg	gov_educgdp	boost_edugdp	educshare	gr_educshare
GNB	Guinea-Bissau	2012	.	2.00106	1.9173647	2.00106	7.707795
GNB	Guinea-Bissau	2013	.	2.1324899	2.0760103	2.13249	6.568015
GNB	Guinea-Bissau	2014	.	2.6467723	2.9335087	2.646772	24.11653
GNB	Guinea-Bissau	2015	.	2.3843326	2.9119318	2.384333	-9.915462
GNB	Guinea-Bissau	2016	.	.	2.0076638	1.643905	-31.05389
GNB	Guinea-Bissau	2017	.	.	2.1959611	2.571994	56.4564
GNB	Guinea-Bissau	2018	.	2.0530073	1.7528518	2.053007	-20.17838
GNB	Guinea-Bissau	2019	.	2.2656409	2.2451687	2.265641	10.35718
GNB	Guinea-Bissau	2020	.	2.7792223	2.996576	2.779222	22.66825
GNB	Guinea-Bissau	2021	.	2.5975792	2.9034163	2.597579	-6.53575

iso3	countryname	year	GF09__gdp_gg	gov_educgdp	boost_edugdp	educshare	gr_educshare
LBN	Lebanon	2012	.	2.14903	2.5072166	2.14903	30.04563
LBN	Lebanon	2013	.	2.43397	2.6213765	2.43397	13.25901
LBN	Lebanon	2014	.	.	4.7996258	4.456492	83.09562
LBN	Lebanon	2015	.	.	4.9539114	4.599748	3.21454
LBN	Lebanon	2016	.	.	5.0242141	4.665025	1.419135
LBN	Lebanon	2017	.	.	5.3568662	4.973895	6.620979
LBN	Lebanon	2018	.	.	6.028789	5.597781	12.5432
LBN	Lebanon	2019	.	.	.	.	.
LBN	Lebanon	2020	.	.	.	.	.
LBN	Lebanon	2021	.	.	.	.	.

- **Combining all three sources has increased the number of observations with values.** Table A.10. shows the distribution of observations from each source and the total observations after combining the data sources.
    - **Gaps remain.** An average of 24 percent of data is missing given our panel of 218 countries and territories from 2012 to 2021; 2021 has the most missing data.

Table A.10. Distribution of Data on Government Education Spending as a Percentage of GDP

Year	UIS[a]	International Monetary Fund	BOOST	Total[b]	Ideal[c]	No data[d]	Missing, %[e]
2012	154	0	13	167	218	51	23
2013	157	0	9	166	218	52	24
2014	160	0	9	169	218	49	22
2015	167	0	8	175	218	43	20
2016	165	0	8	173	218	45	21
2017	168	0	8	176	218	42	19
2018	164	1	7	172	218	46	21
2019	155	2	10	167	218	51	23
2020	106	32	22	160	218	58	27
2021	88	33	21	142	218	76	35
Average							24

a. Consolidated data from the mapping and validation exercise done by UIS.
b. All observations with available data.
c. Total number of observations ideally covered in EFW database.
d. Calculated by subtracting Total from Ideal.
e. Calculated by dividing No Data by Ideal number of observations.

Appendix B: Outlier Detection After Step 2 of Imputation

A. Country Cases

In this appendix, we will look for the outliers after imputation. To illustrate, some country examples are provided here according to income group. Note the following variable names:

- **educshare**: Education spending as a percentage of GDP, combining all sources of data (Step 1 outcome)
- **mean_educs~c:** Average education spending as a percentage of GDP in a given income level, using *educshare* data (derived from Step 1 outcome)
- **gr_mean_ed~c:** Calculated growth rate of *mean_educs~c,* which is the growth rate or trend of average education spending as a percentage of GDP in a given income level
- **educshare~nt:** Values imputed using the growth rate of income-level averages to complete the *educshare* series

Table B.1. Description of Key indicators After Imputation
Aruba as High-Income Country Example

iso3	countryname	WBincgrp	year	educshare	mean_educs~c	gr_mean_ed~c	educshare~nt
ABW	Aruba	HIC	2012	6.34759	4.896451	3.833631	.
ABW	Aruba	HIC	2013	6.44277	5.001834	2.15224	.
ABW	Aruba	HIC	2014	5.85084	4.828811	-3.459205	.
ABW	Aruba	HIC	2015	5.88783	4.838036	.1910485	.
ABW	Aruba	HIC	2016	5.49106	4.756759	-1.679966	.
ABW	Aruba	HIC	2017	.	4.508447	-5.220183	5.204417
ABW	Aruba	HIC	2018	.	4.626945	2.628352	5.341207
ABW	Aruba	HIC	2019	.	4.568654	-1.259826	5.273917
ABW	Aruba	HIC	2020	.	5.017233	9.818633	5.791744
ABW	Aruba	HIC	2021	.	4.715712	-6.009713	5.443677
ABW	Aruba	HIC	2022	.	4.389088	-6.926281	5.066633

Table B.2. Description of Key indicators After Imputation
Turkmenistan as Upper Middle-Income Country Example

iso3	countryname	WBincgrp	year	educshare	mean_educs~c	gr_mean_ed~c	educshare~nt
TKM	Turkmenistan	UMC	2012	3.04925	4.452812	2.298069	.
TKM	Turkmenistan	UMC	2013	.	4.625211	3.871678	3.167307
TKM	Turkmenistan	UMC	2014	.	4.735779	2.390552	3.243023
TKM	Turkmenistan	UMC	2015	.	4.799669	1.349101	3.286775
TKM	Turkmenistan	UMC	2016	.	4.854779	1.148194	3.45973
TKM	Turkmenistan	UMC	2017	.	4.533322	-6.621453	3.230646
TKM	Turkmenistan	UMC	2018	.	4.467728	-1.446924	3.183901
TKM	Turkmenistan	UMC	2019	3.12065	4.378973	-1.986583	.
TKM	Turkmenistan	UMC	2020	.	4.694835	7.213146	3.345747
TKM	Turkmenistan	UMC	2021	.	4.69745	.0557091	3.347611
TKM	Turkmenistan	UMC	2022	.	6.549719	39.43137	4.66762

Table B.3. Description of Key indicators After Imputation
Vanuatu as Lower Middle-Income Country Example

iso3	countryname	WBincgrp	year	educshare	mean_educs~c	gr_mean_ed~c	educshare~nt
VUT	Vanuatu	LMC	2012	.	4.886937	2.523716	5.222811
VUT	Vanuatu	LMC	2013	.	4.892641	.1167081	5.228907
VUT	Vanuatu	LMC	2014	5.18195	4.848704	-.8980167	.
VUT	Vanuatu	LMC	2015	5.60338	4.73536	-2.337618	.
VUT	Vanuatu	LMC	2016	.	4.791497	1.185487	5.669807
VUT	Vanuatu	LMC	2017	4.53343	4.923523	2.755437	.
VUT	Vanuatu	LMC	2018	.	4.608872	-6.390779	4.243709
VUT	Vanuatu	LMC	2019	1.77407	4.830977	4.819086	.
VUT	Vanuatu	LMC	2020	2.20399	4.82873	-.0465291	.
VUT	Vanuatu	LMC	2021	.	4.531748	-6.15031	2.068438

Table B.4. Description of Key indicators After Imputation
Yemen, Rep as Low-Income Country Example

iso3	countryname	WBincgrp	year	educshare	mean_educs~c	gr_mean_ed~c	educshare~nt
YEM	Yemen, Rep	LIC	2012	5.46479	3.181033	-6.594532	.
YEM	Yemen, Rep	LIC	2013	.	3.187893	.2156313	5.476573
YEM	Yemen, Rep	LIC	2014	.	3.382576	6.106945	5.811025
YEM	Yemen, Rep	LIC	2015	.	3.407242	.7292238	5.853401
YEM	Yemen, Rep	LIC	2016	.	3.136282	-7.952462	5.387911
YEM	Yemen, Rep	LIC	2017	.	3.183849	1.51665	5.469627
YEM	Yemen, Rep	LIC	2018	.	3.105131	-2.472415	5.334395
YEM	Yemen, Rep	LIC	2019	.	3.14981	1.438874	5.411151
YEM	Yemen, Rep	LIC	2020	.	3.272342	3.890144	5.621653
YEM	Yemen, Rep	LIC	2021	.	3.393961	3.716573	5.830585
YEM	Yemen, Rep	LIC	2022	.	3.235067	-4.68167	5.557617

B. Results and Diagnostics

- After the variable ***educshare*** was combined with the interpolated values from **educshare_int**, further analyses were conducted to check the outliers.

    - **Checking for influential observations.** Review of the distribution of the data series revealed that the first percentile of the distribution after imputation is at 0.83 and the 99th percentile is at 12.59.

With estimates: Public education expenditure, % of GDP

	Percentiles	Smallest		
1%	.8316338	0		
5%	1.49373	0		
10%	1.872517	0	Obs	10,812
25%	2.753169	0	Sum of wgt.	10,812
50%	4.011544		Mean	4.323007
		Largest	Std. dev.	2.356039
75%	5.41299	27.72204		
90%	6.84873	28.70728	Variance	5.550922
95%	8.068079	28.86719	Skewness	2.440599
99%	12.59481	44.33398	Kurtosis	20.89202

Appendix C: Glossary of Key Terms Used

Bilateral aid: Money that a donor country provides directly to a recipient country[17]

Budget execution: Phase during which public resources are used to implement policies included in the budget[18]

Concessional loans: Loans that are more generous than market loans, as achieved through lower interest rates, longer grace periods, or a combination of both[19,20]

Constant: Monetary values are expressed in terms of their real growth, which means they are adjusted for inflation using a base year[21]

Current: Monetary values, which are influenced by inflation,[22] expressed in the value of the currency for the particular year in which they are reported

Development Assistance Committee list of ODA recipients: All low- and middle-income countries that are eligible, based on gross national income per capita, to receive ODA[23]

Gross domestic product: Value of goods and services produced in a country during a given period of time and bought by end users[24]

Household education spending: Expenditures on education of one or more household members[25]

Multilateral aid: Money that donor countries provide to recipient countries through an international development organization[26]

Non-concessional loans: Money lent based on standard market interest rates and terms, which are therefore less generous than concessional loans[27]

Official Development Assistance: Financial support (grants or concessional loans from OECD DAC member countries to developing countries to advance development in areas such as health, sanitation, education, infrastructure, and strengthening tax systems and administrative capacity[28]

Other official flows: Transactions from the official sector that do not meet the criteria for ODA assistance, including grants for representational or essentially commercial purposes; official bilateral transactions that have less than 25 percent of a grant element; and official bilateral transactions aimed primarily at facilitating exports, regardless of the percentage of the grant element[29]

---

[17] https://www.oecd.org/dac/dac-glossary.htm#Total_Receipts
[18] http://www1.worldbank.org/publicsector/pe/befa05/ADBHandbookchap6.pdf
[19] https://stats.oecd.org/glossary/detail.asp?ID=5901
[20] For purposes of ODA, loans were considered concessional if their grant element was at least 25 percent (calculated at a discount rate of 10 percent) until 2018. Since 2019, the percentage of the grant element and the discount rate for loans to be considered as concessional vary based on the income group of the recipient country. https://www.oecd.org/dac/financing-sustainable-development/development-finance-standards/What-is-ODA.pdf
[21] https://datahelpdesk.worldbank.org/knowledgebase/articles/114942-what-is-the-difference-between-current-and-constan
[22] https://datahelpdesk.worldbank.org/knowledgebase/articles/114942-what-is-the-difference-between-current-and-constan
[23] https://www.oecd.org/dac/financing-sustainable-development/development-finance-standards/daclist.htm
[24] https://www.imf.org/external/pubs/ft/fandd/basics/gdp.htm
[25] http://uis.unesco.org/sites/default/files/documents/rws-bangkok2016-uis_household_survey_on_education_expenditure.pdf
[26] https://www.oecd.org/dac/dac-glossary.htm#Total_Receipts
[27] https://www.oecd.org/dac/financing-sustainable-development/development-finance-standards/What-is-ODA.pdf
[28] https://www.oecd.org/dac/financing-sustainable-development/development-finance-standards/
[29] https://data.oecd.org/drf/other-official-flows-oof.htm

Appendix D. Methodology for the Estimates: Total Government Education Expenditure Growth from 2020 to 2030 in EFW Section 6

The EFW2023 shed light on changes in the school-aged population and projects its fiscal implications for the selected lower-income and lower-middle-income countries from 2020 to 2030. This section explored how the demographic shift interacts with the evolution of total government education expenditures, leading to an increase or decrease in per capita public education expenditures. Two indicators were used in the analysis: number of individuals aged 5 to 24 according to year and country and total government expenditures on education expressed in 2021 U.S. constant dollars according to year and country.

Estimate of school-aged population from 2020 to 2030

The EFW2023 estimated the population aged 5 to 24 in 2020 based on age-specific data that the UN-DESA collected[30] and adopted the projection of the population within the same age range in 2030 from the World Bank databank.[31]

Estimate of growth in total government education expenditures from 2020 to 2030

The EFW2023 assumed that total government education expenditures will continue to grow from 2022 to 2030 at the rate observed from 2009 to 2019. Based on this assumption, EFW2023 predicted total government education expenditures in 2030 using the compound annual growth rate from 2009 to 2019. Total government education expenditures for 2020 to 2021 were excluded to ensure that the disruptive impact of the COVID-19 global pandemic on education finance did not influence the analysis. The specific models are detailed in model 1 and 2.

$$Compound\ annual\ growth\ rate = \left(\frac{Expenditures_{2019}}{Expenditure_{s2010}}\right)^{\frac{1}{2019-2010}} - 1 \text{ (Model 1)}$$

$$Expenditures_{2030} = Expenditure_{s20} \times (1 + compound\ annual\ growth\ rate)^{(2030-2019)} \text{ (Model 2)}$$

[30] https://population.un.org/wpp/Download/Standard/Population/
[31] https://databank.worldbank.org/source/population-estimates-and-projections