Education Finance Watch: Technical Note

This note explains the methodology of the Education Finance Watch 2021 in its presentation of indicators and estimates. The discussion is divided into two parts. Section 1 describes the data used in the EFW brief, the sources that the team has used, and the steps undertaken to process and consolidate the various information to come up with the indicators presented in the brief. Section 2 explains the approach used to extend the series and impute missing values, which will allow computation of global estimates. Finally, Section 3 elaborates on the methodology in the analysis of the government budget documents in relation to COVID-19.

Section 1: EFW Database

The EFW, which is a collaborative effort between the World Bank and UNESCO's Global Education Monitoring Report team, has put together a database containing relevant information on education financing. To maximize data coverage, multiple sources of data have been used. These include the UNESCO Institute for Statistics (UIS), the World Bank World Development Indicators (WDI) and Public Expenditure Reviews (PERs), the International Monetary Fund (IMF) Government Finance Statistics (GFS) and World Economic Outlook (WEO), and the International Development Statistics (IDS) of the Organisation for Economic Co-operation and Development. Table 1 contains a summary of the information taken from these various sources. Each of the datasets are discussed in sub-section 1.1.

Table 1. Summary of data sources and key information in the database

Organization	Dataset	Available Information
United Nations Educational, Scientific and Cultural Organization	UNESCO Institute for Statistics (UIS)	Public and private education expenditure Economic data (GDP, GDP per capita) Enrollment and school-age population
World Bank	Public Expenditure Reviews	Public education expenditure
	World Development Indicators (WDI)	Economic data (GDP, GDP per capita, Official Exchange Rate)
International Monetary Fund	Government Finance Statistics (GFS)	Public education expenditure
	Fiscal Monitor and World Economic Outlook (WEO)	Fiscal data (total government expenditures) Economic data (GDP, GDP per capita)
Organisation for Economic Co-operation and Development	International Development Statistics database	Aid to education, by donor and recipient
United Nations Department of Economic and Social Affairs	World Population Prospects	Population, by age

GDP = gross domestic product

The EFW database contains the actual data downloaded from the respective websites or imported through Stata using the Statistical Data and Metadata Exchange (SDMX) platform.[1] These various sources

[1] UIS data has been accessed using the SDMX platform, using the code -getdata-. For more information, refer to: Sébastien Fontenay, 2016. "SDMXUSE: Stata module to import data from statistical agencies using the SDMX standard," Statistical Software Components S458231, Boston College Department of Economics, revised 30 Sep 2018.

of information have been consolidated to come up with key indicators, especially the public education expenditure and total public expenditure, which are of particular interest in analyzing government investments in education. The preference is to use UIS data where possible, and supplement missing information using data from World Bank and IMF.[2]

In this technical note, it is important to distinguish three sets of data. One set comprises the actual data reported by the different institutions. The second set contains a new set of indicators that combine these actual data following cleaning and processing procedures that are detailed in Section 1. These indicators are referred to as "boosted" indicators. The third set contains another set of indicators, which extends the series of boosted indicators by filling missing values through a systematic method of imputation explained in Section 2. These new indicators are referred to as "extended" indicators.

The rest of Section 1 is organized as follows. Sub-section 1.1 briefly discusses the various sources of information used. Next, the processes of consolidation and cleaning public education spending are discussed in sub-section 1.2. Sub-section 1.3 explains the steps undertaken to process the aid data to calculate total aid to education. Meanwhile, sub-section 1.4 provides information in consolidating the other relevant information such as the economic and fiscal data. Finally, sub-section 1.5 describes additional indicators that have been calculated using the "boosted" indicators, which are presented in the EFW brief.

1.1 Sources of Education Financing Data Included in the EFW Database

The following sub-section describes the data sources, including the method of data collection.

UNESCO. Most of the data in the database are from the UIS database. These include education expenditure data expressed in current and constant US dollars, constant PPP international dollars, as a percentage of GDP, and as a percentage of government expenditures. The data used to fill the questionnaire on education expenditure may come from annual financial reports by the Ministry of Finance and/or the Ministry of Education, and/or national accounts reports by the National Statistical Office. Education expenditures cover government spending as well as household spending. Apart from expenditures, other key education statistics such as enrollment and school-age population, as reported by the education agencies, are included. Non-education data has been collected as well, which UIS has used in their estimates. These include economic data such as GDP (current US dollars, and current and constant 2011 PPP international dollars) and GDP per capita (current US dollars, and current and constant 2011 PPP international dollars). According to UIS metadata, the economic indicators are primarily from the World Bank's World Development Indicators (WDI) and IMF World Economic Outlook (WEO). Demographic information, which has been pulled from various statistical databases such as the UN World Population Prospects and WDI, among others, are included as well. All indicators from the UIS database have been downloaded through SDMX in December 2020—except the data on household expenditure as percentage of GDP, which has been downloaded from the UIS archives. This is because UIS has stopped reporting this indicator since the September 2020 update.

World Bank. Data on education expenditure can be sourced from World Bank Public Expenditure Review (PER) documents, which are in-depth country assessments using data provided by governments. These include indicators on government education spending as a percentage of GDP as well as a percentage of

[2] Several steps have been made to analyze differences and minimize inconsistencies through the removal of influential data points and outliers. See Section 1.2 below on the steps and Appendix A for the diagnostics results.

total government expenditure. Meanwhile, the economic data pulled from the World Development Indicators (WDI) are sourced from its database of national accounts and OECD national accounts data files, and have been estimated consistent with the United Nations System of National Accounts. These economic data have varying currency values, which include current and constant 2010 US dollars, current and constant local currencies (countries have different reference years), and current and constant 2017 PPP international dollars. The data was downloaded from the World Bank website in November 2020.

International Monetary Fund. The IMF Government Finance Statistics contains information on government expenditures, including disaggregated expenditure on education by level. The data is obtained primarily by means of a detailed questionnaire distributed to government finance statistics correspondents, usually located in the ministries of finance or central banks of reporting countries.[3] The countries report government financial information in the framework of the Government Finance Statistics Manual, 2014 (GFSM 2014). The expenditure data in the database are expressed in terms of local currency as well as a percentage of GDP. Furthermore, it includes expenditures from budgetary central government, central government and general government.[4] Meanwhile, the World Economic Outlook (WEO) database contains selected macroeconomic indicators such as GDP (current and constant local currency, current US dollars, and current international dollars), GDP per capita (current and constant local currency, current US dollars, and current and constant PPP 2017 international dollars), GDP deflator (based on the difference between current and constant local currency GDP data), total government expenditure as proportion of GDP (%), and population (in millions). According to the October 2020 WEO metadata, GDP data is from official statistics on national accounts, and the valuation follows the UN System of National Accounts or the European System of Accounts.[5] It should also be noted that the IMF team has done some estimations to avoid breaks in the data, as well as forecasts up to 2025. The dataset contains a variable which indicates when the forecasts start. Lastly, the IMF Fiscal Monitor was obtained for information relating to public finance. The Fiscal Monitor contains actual and projected data relating to total government expenditure. [6] The GFS, WEO, and Fiscal Monitor datasets were downloaded from their respective websites in November 2020.

Organisation for Economic Co-operation and Development. The OECD maintains the International Development Statistics (IDS) databases, which record information of aid flows provided annually by 30 member countries of the OECD Development Assistance Committee (DAC), a forum of major bilateral donors established to promote aid and its effectiveness, as well as a growing number of donors that are not members of the committee such as the United Arab Emirates, Kuwait and Iceland. The database contains project-level and activity-level aid flows of official development assistance (ODA). ODA is public funds provided to developing countries to promote their economic and social development. It is concessional; that is, it takes the form of either a grant or a loan carrying a lower rate of interest than is

[3] A correspondent is a data reporter from a member country who is nominated or approved by his/her respective country coordinator to submit data electronically via the Integrated Collection System. http://datareportinghelp.imf.org/knowledgebase/articles/904623-correspondent-coordinator-definitions

[4] Based on the IMF GFS Manual (GSFM 2014), Budgetary Central Government expenditures form part of Central Government expenditures. In turn, Central Government expenditures are a sub-set of General Government expenditures. However, reporting protocols among countries in the database have differed, resulting in some reporting data in Budgetary Central Government expenditures but reporting nothing under Central or General Government expenditures. In some cases, there is information on Central Government expenditures, but no General Government expenditures.

[5] The methodology varies per country.

[6] The data on total government expenditure as a percentage of GDP in the Fiscal Monitor is the same as the data in the WEO, as the former is the source is the source of the latter. Nonetheless, both are included in the database.

available on the market and, usually, a longer repayment period. Developing countries are those in the DAC List of ODA Recipients. Bilateral donors are countries that provide development assistance directly to recipient countries. Bilateral donors 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 significant part of their activities in favor of developing countries. They include multilateral development banks (e.g., the World Bank and Inter-American Development Bank), United Nations agencies and regional organizations (e.g., the European Commission). The development banks also make non-concessional loans to several 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 by level—all expressed in constant 2018 U.S. dollars. The data were downloaded in January 2021, and recipient countries/territories are the unit of observations. [7]

United Nations Department of Economic and Social Affairs. The World Population Prospects 2019 contains information on the estimated total population disaggregated by age (0, 1, 2, ..., 100+) per country. 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-age population, assumed to be ages 5 to 24 years old (to cover pre-primary to tertiary education), were included in the database. The dataset was downloaded in November 2020.

1.2 Public education spending

The primary indicator of interest is public education spending. The UIS data on public education spending are expressed in many forms and units. These include public education spending as a percentage of GDP (%), public education spending as a percentage of total government expenditure (%), and government spending per student relative to GDP per capita (PPP international dollars), by education level. The indicator with the widest coverage is the public education spending as a percentage of GDP (%), which makes it a viable key indicator on public education spending. Therefore, to maximize coverage in terms of country and time information, the series is boosted by available data from the World Bank and the IMF.

New variables were generated to consolidate information from the various sources. Furthermore, diagnostics and consistency checks across the various sources of information were carried out in order to remove problematic outliers and influential data (see Appendix A). The following steps were undertaken.

Generate consolidated variable on Public Education Spending, % of GDP using UIS as based. A new variable is created using the UIS as the base. Values exceeding the 8.97326 (99^{th} percentile), suspected as outliers, are set to missing. This is used as a cut-off point when using data from other sources to boost the series.

Boosting the series using supplementary sources. IMF GFS data on general government expenditure on education was used as the first booster, provided that values are below the cut-off point and not equal to zero.[8] The third booster comes from the World Bank database of Public Expenditure Review. Table 2

[7] The downloaded data include aggregates such as "Developing Countries, Total" and "Developing Countries, Unspecified". The values of "Developing Countries, Total" includes territories/countries that are not part of the EFW database. The data was downloaded from this link: https://stats.oecd.org/Index.aspx?DataSetCode=CRS1

[8] Zero values are problematic because (1) it is hard to know for certain if this is a coding error or a missing value, and (2) they could potentially influence the data, especially in tabulating averages and in running the regression analysis.

shows the distribution of the data by source and some statistical summaries (mean, minimum and maximum).

Table 2. Distribution of data on public education spending, by source

	Source 1 UIS: Government education expenditure, % of GDP				**Source 2** IMF GFS: General government education expenditure, % of GDP				**Source 3** World Bank PER: Government education expenditure, % of GDP			
Year	Number of obs	mean	min	max	Number of obs	mean	min	max	Number of obs	mean	min	max
2009	113	4.52	1.25	8.97	11	4.55	3.6	6.75	16	4.52	0.7	9.4
2010	120	4.4	1.12	8.56	14	4.38	2.8	7.4	10	3.19	0.9	6.8
2011	117	4.23	0.79	8.48	12	4.25	1.63	6.52	13	3.97	1.7	6.8
2012	110	4.31	1.5	7.92	16	4.59	1.31	7.91	15	4.09	1.6	9.5
2013	120	4.44	1.02	8.49	12	4.18	1.45	8.02	10	4.08	1.7	6.5
2014	115	4.51	1.02	7.7	15	3.83	1.54	8.11	6	4.01	2	6.5
2015	108	4.54	1.47	8.43	18	4.3	1.71	7.01	5	2.77	2.1	3.37
2016	109	4.5	1.37	8.05	16	4.25	1.74	6.82	4	4.24	2.77	6.8
2017	117	4.45	1.45	7.96	10	3.22	0.13	6.05	4	4.3	2.2	6.7
2018	63	4.15	1.92	7.56	51	4.66	0.23	7.37	0			
2019	16	4.11	1.33	7.7	19	4.35	1.53	7.38	0			
TOTAL	**1,108**				**194**				**83**			

1.3 Aid to education

OECD classifies aid to education into four levels: Basic education, secondary education, post-secondary education and education 'unspecified'. Basic education covers primary education, basic life skills for youth and adults, and early childhood education. Secondary education includes both general secondary education and vocational training. Post-secondary education includes advanced technical and managerial training. Education, 'level 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 programme support is often reported within this subcategory. The sum of these aid expenditures is equivalent to the reported total aid to education in the IDS database. In addition, a part of general budget support which is generally unassigned to any particular sector, benefits the education sector. To take this into account, 20 percent of general budget support is added onto the indicator on total aid to education. The assumption is that general budget support is intended to support governments in maximizing their spending in education, based on the recommended share of education in the budget of 15-20%.[9]

1.4 Other relevant variables in the EFW database

Total government expenditure as a percentage of GDP. Total government expenditure is an important indicator, especially in analyzing its relationship with public education spending. To maximize data coverage, various sources have been used. In the database, a new variable is first created using the information on total government expenditures as a percentage of GDP (%) from IMF WEO as the base. It is preferred because it has the highest number of countries with available information on an annual basis. Second, diagnostics were performed to remove outliers. Looking at the distribution of the data, the 99th percentile value was seen too high at 103%, while the 95th percentile value of 54.7% might be too strict. Thus, the middle point between the 95th and 99th = 79.047 was used as the cut-off point. Table 3 shows the distribution of the data by source and the statistical summaries (mean, minimum and maximum).

Table 3. Distribution of data on total government expenditure

Total general government expenditure, % of GDP				
Year	Number of obs	mean	min	max
2009	190	32.67	11.76	79.04
2010	189	31.66	10.23	74.00
2011	189	30.94	9.66	65.28
2012	189	31.34	9.83	65.14
2013	188	31.38	10.91	63.31
2014	189	31.40	12.56	57.71
2015	188	31.47	11.07	56.80
2016	188	30.91	9.97	61.68
2017	188	30.44	8.39	63.95
2018	188	30.35	11.14	65.99
2019	188	30.46	12.62	65.95
TOTAL	**2,074**			

Source: IMF WEO.

[9] See UNESCO (2012) EFA Global Monitoring Report (Box 2.1, page 145) for a discussion of these assumptions

GDP. Economic data are collected because some of the education financing indicators are expressed as a share of either GDP or GDP per capita. Furthermore, these variables are needed in the imputation of missing values to calculate global estimates (see Section 2). To maximize data coverage, two sources of information were used, World Bank WDI and IMF WEO. The base data was World Bank WDI, boosted by the IMF WEO. In the computation of key indicators presented in the EFW brief, the GDP values are especially significant.

The GDP data series is boosted using the combined information from WDI and WEO series. Furthermore, they are rebased to reflect 2018 constant prices (US$). The rebasing steps are as follows.[10] For the WDI GDP data, which was in constant 2010 US dollars, the first step was to create an index by dividing the GDP (in constant values) in a given year by the 2018 GDP constant value. This would yield a series where the 2018 values will equal to one. For example, if the real GDP of the Philippines in 2009 was valued at $194 billion (constant 2010 US$), then dividing it by the real GDP in 2018 at $340 billion (constant 2010 US$), then the index in 2009 would equal to 0.57. The second step was to multiply the index by the current value of the GDP in 2018. Again, using the Philippines case, the index in 2009 at 0.57 will be multiplied by the nominal GDP in 2018 at $346.8 billion (current US$). This will result in the rebased value of $197.9 billion (constant 2018 US$) for 2009. For the WEO GDP data, a slightly different approach was used because the database did not contain GDP in constant US$, but rather in constant local currency. In the absence of real GDP in constant dollars, the index was computed using the real GDP in local currency. After this, the index was multiplied by the 2018 current US$ GDP value. Checks were performed for countries that have available data for both WEO and WDI to determine discrepancies in the rebased series. The discrepancies were close to zero, possibly due to rounding off as the WEO GDP data are presented in billion dollars.

1.5 Indicators presented in the EFW

The EFW brief presents various indicators combining the data from these different sources. Table 4 lists these indicators along with adefinition, data sources and method of calculations. The subsequent text provides a more detailed explanation.

[10] 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

Table 4. Summary of indicators presented in the 2021 EFW

Figures	Indicator (unit)	Definition	Data sources	Method of calculation
1 to 3	Global education spending (constant 2018 US dollars)	Measures the aggregated expenditure on education services by governments, households and donor. It also shows the contribution of each of the sources of education financing and the trends across the years.	UIS, World Bank, IMF, OECD	Calculated by summing up the extended series on spending of government, households and donors. Spending by source was estimated separately using the boosted series (see Section 1) and imputation methods (Section 2) to extend the data and fill gaps.
4	Government education spending per capita (constant 2017 PPP dollars)	Measures how much governments spend on education services given their target beneficiaries, i.e., all children of school-age from pre-primary to tertiary education.	UIS, World Bank, IMF, UNDESA	Boosted series on government education spending from UIS, WDI and GFS divided by total population of school-age (pre-primary to tertiary education) from UIS and UNDESA
5	Education spending as a percentage of government budget (%)	Measures the share of total government spending directed to education. Combined with the information on total government spending as a percentage of GDP, it shows the available fiscal space between the current state and the prescribed benchmarks of education as a percentage of GDP.	UIS, World Bank, IMF	Boosted series on government education spending from UIS, WDI, and GFS divided by boosted series on total government expenditures from GFS and WDI
6 to 7	Aid for education (constant 2018 US dollars)	Measures the amount of aid spent by donor countries, by level of education, and where they were spent in terms of regional and income classification.	OECD	Actual data on aid for education + 20 percent of general budget support
8	Household spending on education as a percentage of GDP (%)	Measures the magnitude of household spending in relation to the country's total income. The bigger the share, the larger the contribution of household spending.	UIS, World Bank, IMF	Data on household spending derived from UIS variables on household spending and enrollment divided by GDP data from WDI and WEO.
10	Expenditure per child (constant 2017 PPP dollars)	Measures the amount spent by government on basic education given the target beneficiaries of all children in primary and secondary education school-age population. Combined with the information on learning adjusted years of schooling, which is an indicator on learning outcomes, the indicator is used as a measure of efficiency.	UIS	Data on government spending on primary and secondary education derived from UIS variables on government spending divided by population of children in primary and secondary levels of education from UIS.

Global education spending. The indicator presented in Figures 1 to 3 combines the expenditure on education services of the main sources of education financing—government, households, and donors. The amounts are expressed in constant 2018 US$. To ensure comparability across the years and maximize the sample of countries in the aggregation, missing values were imputed for government and household spending, especially for recent years where gaps are more prevalent. The approaches used to systematically impute values are detailed in sections 2.1 and 2.2, respectively. Meanwhile, total aid to education has complete information from 2009 to 2019. Aid to education can be identified by recipient country so it is straightforward to identify aid spending for each income group. However, the destination of some aid to developing countries is not specified at the country level. This was re-distributed to the income groups: 60 percent to low-income countries, 20 percent to lower-middle income countries, and 20 percent to upper-middle income countries.[11] Furthermore, recognizing that a portion of aid is already reflected in government spending, 60 percent of total aid to education was deducted in each recipient country's government spending on education before aggregating for income group and global estimates.[12]

Education spending as a percentage of government budget. This indicator is presented alongside the data on total government expenditure as share of GDP. Calculating this indicator used the following boosted series: (1) government education spending as a percentage of GDP (see section 1.2), (2) total government expenditures as a percentage of GDP (see section 1.4), and (3) GDP in constant 2018 US dollars (see section 1.4). The first two variables were multiplied by GDP to yield expenditure values in 2018 constant US dollars. After this, the government education spending was divided by total government expenditure and multiplied by 100 percent.

Government education spending per capita. This is an indicator presented in Figure 4 to highlight how much governments spend given their target beneficiaries, who are children that should be receiving education, regardless of whether they are in enrolled or not. To compute for this indicator, the first step was to multiply the boosted series on public education spending as percentage of GDP by the series on GDP (2017 constant PPP$) using information from WDI, supplemented by IMF WEO where missing. The result was the value of total public education spending (constant 2017 PPP$). Next, this was divided by the school-age population (pre-primary to tertiary) as reported by countries to the UIS, and where unavailable, the data on school-age population (age 5 to 24 years old) from UNDESA. This yielded the government education spending per capita. Finally, to maximize information, averages for the periods 2013-2014 and 2017-2018 were calculated.

Aid for education. The estimates shown in Figures 6 and 7 are based on the OECD's ODA data for education and general budget support. As such, total education assistance is direct education plus 20 percent of general budget assistance. Total aid to basic education is a total of direct aid to basic education, 10 percent of general budget support, and 50 percent of education, 'level unspecified'. Meanwhile, total aid to secondary education is a total of direct aid to secondary education, 5 percent of general budget support and 25 percent of education, 'level unspecified'. Total aid to post-secondary education is estimated as a total of direct aid to post-secondary education, 5 percent of general budget support and 25 percent of education, 'level unspecified'. Presented in Figure 6 are values of "Developing Countries, Total", and in Figure 7 are the regional aggregates using World Bank regional classification.

[11] This assumption is based on the trends in the distribution of unspecified aid, upon review of the microdata for the years 2016 to 2019. Most of the unspecified aid went to GPE, UNICEF, UNESCO and publicly funded research institutions, whose allocation of aid to income groups closely reflect this assumption.

[12] See UNESCO (2012) EFA Global Monitoring Report (Box 2.1, page 145) for the discussion on these assumptions.

Household spending on education. Total household spending as percentage of GDP, which is presented in Figure 8, uses available information from UIS. Two indicators were used: one is the household spending per student as a percentage of GDP per capita (by level)[13], and the other is the total household spending as a percentage of GDP. It should be noted that the updated UIS database no longer reports the second indicator starting September 2020, hence the need to use the archived data, which could be downloaded from the UIS website. The first step is to convert the first indicator by multiplying the variable by the corresponding enrollment number in that level and the boosted GDP per capita (2010 constant US$). This yielded the value of household spending by level in 2010 constant 2010 US$. Only countries with complete information on all levels are kept. Next, the expenditures by level (i.e., primary, secondary and tertiary) are aggregated to yield total household spending. This is divided by the boosted GDP data (2018 constant US$), resulting in total household spending as a percentage of GDP. To consolidate, the data from the UIS archives was used as base and where missing, the computed values were taken.

Government primary and secondary expenditure per capita. Figure 10 uses the indicator expenditure per capita, which differs from the indicator in Figure 4. In this chart, expenditure per capita is computed as total public education spending on primary and secondary divided by primary and secondary education school-age population. Therefore, this indicator focuses on primary and secondary education spending. Three sets of variables are used to calculate this indicator: one, the government expenditure per student (primary and secondary) from UIS; and two, primary and secondary school-age population; and three, enrolment data in primary and secondary public schools. First, expenditure by level was calculated by multiplying the first and third variables. Second, these expenditures were combined using weights derived from the school-age population.

Section 2: Extending the data through imputation

Global education spending consists of government, household, and donor expenditures. The main objective is to provide a reasonable estimate of global education spending between the years 2009 and 2019 despite the lack of information for some countries. Consolidating various sources of information was not sufficient because coverage was still irregular, especially in the recent years. This resulted in highly uneven number of countries across the years in which some income groups are not represented, which would make global trends as well as income-level averages inconsistent and non-comparable.

To address these challenges, imputation of missing data was adopted. The imputation methods varied for the main sources of education spending (i.e., government and household).[14] For public education spending, regression analysis was used to take advantage of available data (economic and fiscal indicators) that were highly associated with this indicator. Meanwhile for household spending, the very low sample size did not provide enough statistical power to use regression analysis as a method of imputation.[15] As an alternative solution, interpolation was done using income-level averages. A more detailed account of the procedures undertaken are discussed below.

[13] In using the set of household spending variables expressed in proportion of GDP per capita, we are able to compute values that are in constant US$, which is the currency used for the global education spending estimates.
[14] As the IDS contains complete data from DAC donors from 2009 to 2019, no imputation was done on aid expenditures.
[15] The EFW team still attempted to use regression analysis. However, in the initial regression runs, available variables such as economic and fiscal data were not statistically significant.

2.1 Public Education Spending

Table 5 shows the number of countries that have information on public education spending per year given the combined sources (i.e., boosted series), and the equivalent coverage in terms of the world's school-age population. Coverage in terms of number of countries and contribution to the global school-age population have fluctuated. From 2009 to 2013, the number of countries with available information correspond to at least 90% of the world's school-age population. However, from 2014 onwards, the coverage has declined significantly. This is because different countries comprise the different years, with some years having less of countries that have bigger school-age population.[16] Due to lags in reporting, the coverage is remarkably less than half in the years 2018 and 2019, which present challenges in calculating global public spending that is as close as possible to global estimates.

Table 5. Number of countries with boosted government education spending data

Year	Number of countries reporting	Number of countries not reporting	Coverage, % of global school-age population	Coverage, % of global GDP
2009	141	77	90.0	96.6
2010	145	73	92.1	96.3
2011	143	75	90.8	94.5
2012	142	76	91.4	92.2
2013	143	75	91.7	92.5
2014	137	81	64.4	88.6
2015	132	86	65.1	88.7
2016	130	88	62.6	89.9
2017	132	86	60.1	88.0
2018	114	104	34.1	68.0
2019	36	182	9.8	10.6
Total	1,395	1,003		

Sources: EFW database; and UN World Population Prospects 2019, accessed November 2020.
Note: The total number of countries in the EFW database is 218. This number includes all countries with available data on the various sources, and has not been checked yet for outliers and influential data. The coverage per year is the sum of country shares to the global school-age population or global GDP that have available information on education spending. For the coverage using global school-age population, first extract the population of aged 5-24 years old in all countries by year from the UN population data. Second, all country values are summed up per year, yielding the annual global school-age population. Lastly, country values are divided by the global population and multiplied by 100% to yield the country shares to the global school-age population. Only the country shares of countries with education spending data (either IMF or UIS, or both) are aggregated. Meanwhile, for the coverage using global GDP, data on GDP in constant 2010 US$ (from WDI) are aggregated by year. Next, country values are divided by this aggregated value and multiplied by 100% to yield the country shares in percent.

Regression analysis using Ordinary Least Squares is used to predict missing values of education spending. The advantage in using this method is that it allows for a systematic approach, leaving less room for assumptions and discretionary judgment. It uses historical information on the statistical relationships of fiscal and economic data with education spending. The prediction of values will make use of these estimated relationships and the availability of information of these economic and fiscal data. The model also controls for unobservable country and time-specific variations, thus the inclusion of country and fixed effects.

[16] The differences are not just from the drop in the number of countries, but because of the composition as well. For example, there were 14 countries in 2013 (comprising 28% of the global school-age population) that were not in 2014. Meanwhile, there were 8 countries in 2014 (comprising 1.3% of the global school-age population) that were not in 2013.

The model constructed for this purpose is stated as Equation 1. Public education spending as share of GDP ($educshare$) is the dependent variable, and the predictors are total government expenditure as share of GDP ($govshare_{it}$) and GDP per capita ($gdppc_{it}$), with country β_{3i} and time fixed effects β_{4t} . The use of GDP as a denominator in both public education spending and total government expenditures minimizes the spread of the data, especially as education spending in absolute terms (even in constant 2018 US$) could vary between US$ 5.8 million to US$1.2 trillion.

Equation 1: $educshare_{it} = \beta_0 + \beta_{1it} govshare_{it} + \beta_{2it} gdppc_{it} + \beta_{3i} + \beta_{4t} + \varepsilon_{it}$

The data on public education spending as share of GDP uses the boosted series (see Section 1.2). Total government expenditure as a percentage of GDP also uses the boosted series (see Section 1.4). For the GDP per capita, information from WDI and WEO were also combined. WDI data was in 2010 constant US dollars and WEO data was rebased to 2010 to ensure conformity with WDI. The rebasing steps are as follows.[17] An index was computed by dividing the GDP per capita (in local currency) in a given year by its 2010 value (in local currency). The index value in 2010 should equal to zero. After this, the index is multiplied by the 2010 current GDP per capita (current US$) to yield the rebased series. In performing the checks, the average difference between the WDI and rebased series of WEO was calculated at -52.8—meaning on the average, WEO GDP per capita is larger by 53 US dollars. Plotting the scatterplot of these two series shows a very close linear relationship and a correlation of 99 percent (Appendix A).[18]

More diagnostics were conducted to detect outliers and influential data that were dropped from the sample. A simple linear regression was estimated using equation 1, but without the country and time fixed effects. Given the regression estimates, outlier tests such as studentized residuals, leverages, and Cook's D, as well as the DFBeta test for influential data were undertaken (See Appendix A for the results and full discussion). Data points that failed to pass both outlier and influential tests were dropped from the sample, meaning the values in all the variables (dependent and explanatory) were set to missing. Lastly, countries with only one year of complete information on all three variables are dropped from the sample, as they do not have enough information/variation. Thus, from 1,395 countries with information on government education spending as a percentage of GDP, this was reduced to 1,369 after the removal of the outliers and problematic information. Finally, for the regression analysis, only a sample of 1,335 countries were included as they have complete information on all other variables (GDP per capita and total government expenditure as a percentage of GDP). Table 6 shows the summary statistics (mean, min, max) of the final data included in the regression analysis, by year and income group.

[17] 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

[18] In the end, IMF data was not used as countries that used it were not included in the final sample for the regression analysis due to missing data in the other explanatory variables.

Table 6. Summary statistics: regression analysis (part 1)

	2009			2010			2011			2012		
	mean	min	max	mean	min	max	mean	min	max	mean	min	max
Variable: educshare												
Total	4.49	0.70	8.45	4.32	0.90	8.56	4.24	0.79	8.48	4.31	1.31	7.92
LIC	3.29	1.25	6.17	3.14	1.12	6.78	3.29	0.96	6.32	3.30	1.70	6.17
LMIC	4.21	0.70	7.60	4.17	0.90	7.59	4.18	0.79	8.14	4.26	1.50	7.92
UMIC	4.37	1.77	6.20	4.09	1.61	6.72	3.96	1.65	6.46	4.13	1.91	7.91
HIC	5.09	2.55	8.45	4.99	2.05	8.56	4.86	1.63	8.48	4.87	1.31	7.60
Variable: Total government expenditure, % of GDP												
Total	31.95	12.47	57.15	31.07	10.23	65.11	30.20	9.66	56.43	30.65	9.83	57.96
LIC	20.14	12.47	37.99	21.32	14.26	40.84	20.92	12.04	42.19	21.24	11.52	37.54
LMIC	24.25	12.67	48.62	24.88	12.70	49.15	24.50	13.20	45.70	25.03	14.22	48.97
UMIC	31.47	16.95	51.70	29.12	14.47	46.65	27.93	14.49	41.64	28.55	14.15	45.13
HIC	40.57	15.85	57.15	39.93	10.23	65.11	39.39	9.66	56.43	39.99	9.83	57.96
Variable: GDP per capita												
Total	15,621.82	230.19	101,939.62	15,519.04	234.24	104,965.30	15,116.59	235.99	105,264.75	14,614.68	238.82	102,404.61
LIC	526.09	230.19	837.33	574.92	234.24	892.57	615.26	235.99	1,416.43	598.93	238.82	909.31
LMIC	1,792.71	567.91	4,044.13	1,890.38	592.40	4,168.51	1,982.87	612.03	4,233.58	1,854.64	642.52	4,142.49
UMIC	6,246.49	2,919.40	10,594.99	6,411.25	2,898.94	11,286.24	6,529.40	2,965.15	11,678.13	6,610.55	2,999.42	12,039.30
HIC	34,724.81	7,683.13	101,939.62	35,995.39	8,000.38	104,965.30	35,657.47	8,313.27	105,264.75	35,238.38	8,580.09	102,404.61

Table 6. Summary statistics: regression analysis (part 3)

	2013			2014			2015			2016		
	mean	min	max	mean	min	max	mean	min	max	mean	min	max
Variable: educshare												
Total	4.41	1.02	8.49	4.47	1.54	8.11	4.47	1.47	7.59	4.49	1.54	8.03
LIC	3.30	1.02	6.12	3.56	1.93	6.78	3.64	1.47	6.37	3.40	2.02	5.87
LMIC	4.34	1.62	6.78	4.40	1.91	7.09	4.38	2.20	7.59	4.39	1.54	6.99
UMIC	4.24	2.43	8.02	4.27	2.00	8.11	4.22	2.10	7.08	4.34	2.76	7.28
HIC	5.04	1.45	8.49	4.95	1.54	7.70	5.00	1.71	7.57	4.98	1.74	8.03
Variable: Total government expenditure, % of GDP												
Total	30.89	10.91	60.27	31.45	12.56	57.28	31.51	12.58	56.80	30.79	12.57	56.66
LIC	22.26	12.42	34.70	22.25	12.56	32.32	23.19	13.03	34.76	21.96	13.55	31.10
LMIC	25.80	13.72	48.11	26.36	14.23	48.56	26.87	13.77	43.07	25.51	13.45	40.56
UMIC	28.66	13.99	43.74	29.04	13.60	44.81	29.34	12.58	42.69	28.52	12.57	41.88
HIC	39.88	10.91	60.27	39.69	12.57	57.28	39.22	14.45	56.80	38.02	15.19	56.66
Variable: GDP per capita												
Total	14,583.07	242.85	103,721.75	15,819.27	245.33	105,658.52	15,983.06	228.43	107,638.21	16,734.72	219.96	110,162.12
LIC	607.88	242.85	929.47	599.42	245.33	886.44	638.15	228.43	936.00	609.11	219.96	909.60
LMIC	1,968.18	670.84	4,220.39	2,151.45	711.30	4,578.29	2,204.51	732.00	4,308.42	2,134.44	729.66	4,310.57
UMIC	6,980.94	3,056.15	12,842.16	7,157.31	3,137.26	13,277.76	7,200.84	3,210.87	13,853.10	6,926.73	3,242.64	14,062.73
HIC	35,137.16	8,848.89	103,721.75	36,063.43	9,163.63	105,658.52	37,048.93	9,476.67	107,638.21	36,721.90	9,833.61	110,162.12

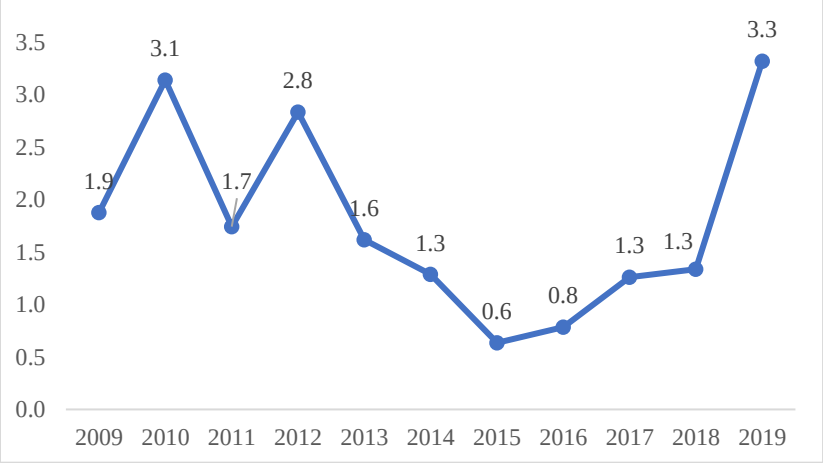
Table 6. Summary Statistics, Regression Analysis (part 3)

	2017			2018			2019			Total		
	mean	min	max	mean	min	max	mean	min	max	mean	min	max
Variable: educshare												
Total	4.42	1.45	7.91	4.44	1.51	7.56	4.27	1.33	7.70	4.40	0.70	8.56
LIC	3.48	1.45	5.60	3.91	2.13	6.99	7.70	7.70	7.70	3.44	0.96	7.70
LMIC	4.54	1.96	7.23	4.46	1.92	6.99	3.44	1.33	6.02	4.31	0.70	8.14
UMIC	4.39	2.47	7.51	4.14	2.26	7.56	4.25	2.56	7.05	4.22	1.61	8.11
HIC	4.71	1.64	7.91	4.77	1.51	7.37	4.45	1.53	7.38	4.92	1.31	8.56
Variable: Total government expenditure, % of GDP												
Total	30.46	10.39	56.48	31.01	13.23	55.66	29.05	13.52	45.71	30.94	9.66	65.11
LIC	22.29	10.39	33.03	23.05	14.26	32.61	20.29	20.29	20.29	21.83	10.39	42.19
LMIC	25.75	16.57	41.48	25.50	16.57	51.57	25.34	15.11	41.45	25.43	12.67	51.57
UMIC	28.28	12.78	41.12	27.91	13.23	40.72	29.62	13.52	45.71	28.94	12.57	51.70
HIC	37.55	13.66	56.48	38.43	14.01	55.66	31.04	14.34	41.84	39.10	9.66	65.11
Variable: GDP per capita												
Total	17,487.89	214.14	109,452.96	18,795.10	210.80	110,701.88	19,352.19	488.48	79,406.66	16,028.60	210.80	110,701.90
LIC	634.42	214.14	910.27	645.77	210.80	931.00	488.48	488.48	488.48	604.52	210.80	1,416.43
LMIC	2,241.75	779.20	4,343.44	2,038.79	817.78	4,210.86	2,527.59	1,287.82	3,715.77	2,026.70	567.91	4,578.29
UMIC	7,219.80	3,262.52	14,874.78	6,982.82	3,266.75	15,068.98	7,121.01	3,284.03	14,998.98	6,828.62	2,898.94	15,068.98
HIC	39,085.59	10,199.48	109,452.96	40,802.05	10,577.20	110,701.88	48,034.94	10,949.24	79,406.66	36,838.25	7,683.13	110,701.88

Based on the results of the linear regression (Appendix B), total government expenditure as a share of GDP as well as GDP per capita were statistically significant predictors of public education spending as a share of GDP. Looking at the goodness-of-fit measured by the R-squared, the model was able to explain the variation by 87%. Finally, the standard deviation of the residuals, measured by the Root Mean Squared Error (RMSE), is 0.57.

After retrieving the parameter estimates from the regression analysis, the data from IMF GFS, IMF WEO and World Bank WDI on total government spending as a percentage of GDP and GDP per capita were plugged in Equation 1 to compute the predicted values of public education spending. Plotting the average mean percentage error (Figure 1), which computes for the average percentage differences between the fitted and actual values, shows whether the predicted values are systematically smaller (negative) or bigger (positive) than the actual values. Based on Figure 1, it seems that the model's predicted values exceed the actual values by a range of 0.6% to 3.3%. For example, if the actual value was 4.5%, then it means that the predicted value could range from 4.58% to 4.7%. This means the error is quite small.

Figure 1. Average mean percentage error (%)



Note: Mean Percentage Error is computed by dividing the difference between fitted and actual values (fitted – actual) by the actual values, and then multiplying by 100%, per country. The average is calculated by obtaining the average of all countries in the sample per year.

Combining the actual values of public education spending and the predicted values where missing, referred to as the "Extended data", Table 6 shows the increase in the coverage brought by the imputation exercise.

Moreover, the imputation has not caused any significant deviation from the actual figures as most of the imputed data are in 2019. Using the combined actual and predicted values, total education spending in constant 2018 US$ was computed by multiplying them with GDP data from World Bank WDI and IMF WEO, where missing. Table 7 shows the aggregated values per year for both boosted and extended data.

Table 6. Average values of public education spending as a share of GDP, boosted and extended

Year	Boosted (n = 1,369)	Number of countries	Coverage, % of global school-age population	Extended data (n = 1,781)	Number of countries	Coverage, % of global school-age population
2009	4.5	131	87.4	4.4	160	94.0
2010	4.3	142	92.0	4.3	162	94.1
2011	4.2	142	90.8	4.3	162	94.2
2012	4.3	139	91.3	4.3	159	94.2
2013	4.4	142	91.7	4.4	163	94.2
2014	4.5	134	64.2	4.4	161	94.0
2015	4.5	131	65.1	4.4	165	94.2
2016	4.5	129	62.6	4.4	162	94.2
2017	4.4	129	60.1	4.2	163	94.2
2018	4.4	114	34.1	4.2	164	94.2
2019	4.1	36	9.8	4.4	160	94.2

Table 7. Government education expenditure, constant 2018 billion US$, boosted and extended

Year	Boosted (n = 1,369)	Number of countries	Coverage, % of global school-age population	Extended data (n = 1,781)	Number of countries	Coverage, % of global school-age population
2009	3,162.25	131	87.4	3,196.87	160	94.0
2010	3,294.24	142	92.0	3,320.48	162	94.1
2011	3,262.18	142	90.8	3,359.29	162	94.2
2012	3,292.89	139	91.3	3,456.89	159	94.2
2013	3,133.25	142	91.7	3,294.19	163	94.2
2014	3,101.40	134	64.2	3,358.31	161	94.0
2015	3,380.03	131	65.1	3,659.46	165	94.2
2016	3,508.61	129	62.6	3,744.20	162	94.2
2017	3,514.02	129	60.1	3,825.65	163	94.2
2018	2,732.54	114	34.1	3,839.28	164	94.2
2019	364.61	36	9.8	4,098.77	160	94.2

Notes:

1. Boosted data = Combined information from various sources following cleaning and processing procedures outlined in Section 1. Note that this will not equal the sum of countries in Table 5 because some observations were dropped following cleaning and processing.
2. Extended data = Combined actual and estimated data. This means boosted values are retained and the predicted values are used where boosted data is missing.
3. The coverage per year is the sum of country shares to the global school-age population that have available information on education spending. To compute, first extract the population of aged 5-24 years old in all countries by year from the UN population data. Second, all country values are summed up per year, yielding the annual global school-age population. Lastly, country values are divided by the global population and multiplied by 100% to yield the country shares to the global school-age population. Only the country shares of countries with education spending data are aggregated.
4. Twenty three (23) countries in the dataset do not have information on school-age population as they are not in the UN World Population Prospects. These are mostly small island states across the world (ex. American Samoa, British Virgin Islands, Isle of Man, Nauru, Palau, St. Kitts and Nevis, etc). Thus, when the coverage goes down, it is likely due to the lack of data from bigger countries. Case in point is 2017, which has 124 countries but having coverage of only 34% whereas in 2009, the coverage is wide at 87% despite only having 4 countries more.

2.2 Household education spending

The only source of data on household education spending was the UIS database. There are, however, a lot of gaps—85.2 percent are missing values in a panel of 218 countries across 11 years (2009–2019). To fill the gaps, interpolation was carried out.

Table 8. Summary statistics of household expenditure, % of GDP

year	Mean	Min	Max	Countries with data	No data
2009	1.23	0.05	8.40	44	174
2010	1.08	0.00	4.48	42	176
2011	1.33	0.00	6.08	42	176
2012	1.17	0.00	4.22	40	178
2013	1.31	0.00	4.99	41	177
2014	1.15	0.04	4.99	41	177
2015	0.99	0.02	4.64	39	179
2016	0.99	0.00	3.87	32	186
2017	0.88	0.00	3.12	29	189
2018	1.88	0.47	3.18	4	214
2019					218

First, for countries that have data but with missing values in between years, the gaps have been filled by the closest previous information. For example, Chad has information in 2009 and 2011 at 2.2 percent and 2.1 percent, respectively. In this first step of interpolation, the 2010 value will assume the 2009 value of 2.2 percent; and for 2012 to 2019, the values will assume the 2011 value of 2.1 percent. As an additional step, zero values are set to missing as this may be problematic—either they are errors or if not, too influential when computing averages.

Second, annual average household education spending as a percentage of GDP was computed by income group (Table 9). This average was used to replace the missing values for countries that do not have any information on household spending at all. However, this was done only for countries that have government education spending data. This was done to ensure a consistent sample in the aggregation of education spending in the global level. Therefore, the final variable on total household spending as a percentage of GDP includes both the actual and the interpolated values.

Table 9. Average household expenditure, % of GDP, by income group

year	Low income	Lower middle-income	Upper middle-income	High income
2009	3.13	1.66	0.97	0.58
2010	3.04	1.73	1.01	0.58
2011	2.95	1.70	0.99	0.60
2012	2.98	1.67	1.00	0.67
2013	2.92	1.72	1.01	0.67
2014	3.25	1.65	0.97	0.69
2015	3.25	1.66	0.99	0.68
2016	3.25	1.68	1.18	0.69
2017	3.25	1.60	1.15	0.68
2018	3.25	1.61	1.10	0.68
2019	3.25	1.61	1.10	0.68

Note: The averages were computed after the first step of interpolation.

Finally, total household spending by country was computed by getting the product of the household spending as a percentage of GDP and the GDP in constant 2018 US dollars (boosted and forecasts).[19]

Table 10. Average household expenditure, % of GDP

year	Mean	Countries with data (actual)	Mean	Countries with data (boosted + extended)
2009	1.2	44	1.3	156
2010	1.1	42	1.3	160
2011	1.3	42	1.3	160
2012	1.2	40	1.4	159
2013	1.3	41	1.4	162
2014	1.1	41	1.4	160
2015	1.0	39	1.4	160
2016	1.0	32	1.4	159
2017	0.9	29	1.4	160
2018	1.9	4	1.4	161
2019			1.4	158

Note: Boosted refers to data using first interpolation. Extended refers to data after the second interpolation.

Section 3: COVID-19 and Government Budgets

In order to begin to understand the short-term impact of Covid-19, we gathered information on education and total public budgets from 29 developing countries. The main sources were public budget laws, resources found at ministries of finance and Education Budget Briefs from UNICEF. The data collected was then shared with the country teams who reviewed it and provided their feedback. Initially, we focused on finding information for the top 5 countries with the highest proportion of school age population in their respective region. However, since budget information was not always available through online resources, adjustments had to be made to the original list to include countries with readily available information. Our final sample, which represents around 54% of the world´s school age population, consists of the following countries: Afghanistan, Argentina, Bangladesh, Brazil, Chile, Colombia, Egypt, Ethiopia, India, Indonesia, Jordan, Kazakhstan, Kenya, Kyrgyz Republic, Mexico, Morocco, Myanmar, Nepal, Nigeria, Pakistan, Panama, Peru, Philippines, Russia, Tanzania, Turkey, Uganda, Ukraine, Uzbekistan.

Because our main objective was to compare pre and post Covid trends in education budgets, we focused on gathering information for 3 fiscal years (two fiscal years before the pandemic to build the pre-Covid trend and the first fiscal year after the pandemic to build the post-Covid trend). In practice this meant looking at the two fiscal budgets that were approved before March 2020 (when the crisis started globally) and the first fiscal budget approved after this date. For countries with fiscal years ranging from January to December[20], we looked at budget information for 2019, 2020, and 2021. For countries with fiscal years

[19] The boosted data is the combined GDP data from WDI and WEO, rebased to 2018. It only includes actual data, and not forecasts. Forecasts come from the IMF WEO database. To differentiate actual and forecast data, WEO specifies which year forecasts begin.

[20] Afghanistan, Argentina, Brazil, Chile, Colombia, Indonesia, Jordan, Kazakhstan, Kyrgyz Republic, Mexico, Morocco, Nigeria, Panama, Peru, Philippines, Russia, Turkey, Ukraine, Uzbekistan

ranging from July-June[21] and October-September[22], this meant looking at budgets from 2018-2019, 2019-2020, and 2020-2021. Finally, for countries with fiscal years April-March,[23] we looked at budgets for the years 2019-2020, 2020-2021, and 2021-2022. To avoid confusion, we classified budgets as being pre or post Covid without using fiscal years or differentiating by the time of their approval. All budgets were expressed in 2021 prices using end of year inflation rates found at the IMF for the fiscal year in which budget was approved. [24]

One of the limitations of our data is that it only includes resources directly assigned to education at the central level, and not by local governments. [25] As a consequence, for countries where education budget comes mainly from unconditional fiscal transfers or from local revenue, our numbers underestimate the total amount that is spent on the sector. A fixed effect approach was used to address this problem by using changes instead of levels. This approach allowed us to focus not on the actual amount of resources that go to education but rather on the changes observed in those amounts. The caveat of this approach is that it only reflects the preferences of central governments, which in practice may differ from those of local governments, for which we have no information.

The second limitation is that the data represents planned instead of executed budget, which in some countries may differ significantly. However, because the impact of Covid is still very recent, this information does allow us to identify the signals in terms of budget allocation, sector prioritization, and overall expected trends in selected countries.

[21] Bangladesh, Egypt, Ethiopia, Kenya, Nepal, Pakistan, Tanzania, Uganda

[22] Myanmar

23 India

[24] https://www.imf.org/external/datamapper/PCPIEPCH@WEO/OEMDC/ADVEC/WEOWORLD

[25] Except for Ethiopia and Kazakhstan, for which expenditure at the local level was available for 2021. The countries that have education shares below 10% and therefore are likely to have other main financing sources besides budget assigned by the central government are: Argentina, Brazil, Egypt, India, Myanmar, Pakistan, and Russia.