Brief Methodological Overview of Skills Clock

Table of Contents

Overview

The World Skills Clock projects learning trends of both secondary-level literacy/numeracy skills and digital skills for the global youth population of ages 15 to 24, disaggregated by country.

1. Population projections

The population projections[1] used for these analyses come from the Wittgenstein Centre Human Capital Data Explorer using SSP2 scenario.[2] Missing data was interpolated using the standard exponential growth function based on the growth rate between data points, which are provided in five-year intervals.

The annual growth rate for a given interval is calculated as $r = \ln\left(\frac{Population_{y+5}}{Population_y}\right) * \frac{1}{5}$, and the corresponding interpolations are calculated as $Population_{y+1} = Population_y * e^{rt}$, where t=1.

2. Secondary-level literacy and numeracy skills

Secondary-level literacy and numeracy skills represent traditional reading and math skills typically associated with in-school learning. To assess the proportion of youth with secondary-level literacy and numeracy skills, data were primarily taken from the UNESCO Institute for Statistics (UIS) database, which reports data (disaggregated by gender) for the monitoring of Sustainable Development Goal (SDG) Indicator 4.4.1 on the proportion of children and young

---

[1] https://tntcat.iiasa.ac.at/SspDb; The file used is SspDb_country_data_2013-06-12.csv.

[2] Detlef P. van Vuuren, Elke Stehfest, David E.H.J. Gernaat, Jonathan C. Doelman, Maarten van den Berg, Mathijs Harmsen, Harmen Sytze de Boer, Lex F. Bouwman, Vassilis Daioglou, Oreane Y. Edelenbosch, Bastien Girod, Tom Kram, Luis Lassaletta, Paul L. Lucas, Hans van Meijl, Christoph Müller, Bas J. van Ruijven, Sietske van der Sluis, Andrzej Tabeau, **Energy, land-use and greenhouse gas emissions trajectories under a green growth paradigm**, Global Environmental Change, Volume 42, 2017, Pages 237-250,
ISSN 0959-3780, DOI:10.1016/j.gloenvcha.2016.05.008. All data are available here: http://dataexplorer.wittgensteincentre.org/wcde-v2/.

people (a) in grades 2/3; (b) at the end of primary; and (c) at the end of lower secondary achieving at least a minimum proficiency level in reading and math.

For the secondary-level learning benchmarks, the minimum levels of proficiency are set at Level 2 of the Programme for International Student Assessment (PISA), Intermediate Benchmark of the Trends in International Mathematics and Science Study (TIMSS) Grade 8, and other equivalent levels set for regional and national learning assessments.[3] For countries that had both literacy and numeracy data in a given year, the average value of these two measures was used. For countries that had either literacy or numeracy data in a given year, this value was used. Because there were many countries that had not participated in either of these two secondary assessments, we estimated learning for countries without data with a statistical predictive model that builds on earlier models for identifying the determinants of learning outcomes[4].

We tested a range of variables suggested by this earlier research to be correlated with learning outcomes, and the primary-level learning dataset. Whenever available, we used gender-specific estimates of these variables. These were compiled into a hierarchical set of predictive learning production models using multivariate regression.[5] These regressions use the most recent available year of data for each country with at least one data point. The two models used are shown below.

Table 1. Regression results for the learning production functions.

	(1)	**(2)**
Log GDP per capita	0.06**	0.09***
Gender: Male	-0.09*	-0.09*
Percent reaching primary learning benchmark	0.02***	
Region: Europe and Central Asia		0.22***
Region: Latin America and the Caribbean		-0.05
Region: Middle East and North Africa		-0.14
Region: North America		0.09
Region: South Asia		0.17
Region: Sub-Saharan Africa		-0.02
Harmonized Learning Outcomes		0.01***
Constant	2.16***	0.73***
N	262	402
R-squared	0.80	0.79

[3] For more information on minimum proficiency levels set by the Global Alliance to Monitor Learning, see this table.
[4] Some of this literature was summarized in the UNESCO 2008 Global Monitoring Report on learning. In 2011, the Education Policy and Data Center (EPDC) proposed a production function to predict learning outcomes using a multi-stage regression analysis (Education Policy and Data Center. 2011. "Estimating Education output in Developing Countries". EPDC Working paper 2011-1. FHI-360, Washington, DC.). Hanushek and Krueger also analyzed school-level inputs as determinants of learning outcomes in various publications (Hanushek, Eric A. 1995, "Interpreting Recent Research on Schooling in Developing Countries." World Bank Research Observer, 10(2):227-246; Hanushek, Eric A. 1997, "Assessing the Effects of School Resources on Student Performance: An Update", Educational Evaluation and Policy Analysis 19(2), Summer 1997, pp. 141-164.; Hanushek, Eric A. and J.A. Luque, 2003. "Efficiency and equity in schools around the world." Economics of Education Review 22 (2003) 481–502; Krueger 2003.)
[5] The predicted learning outcome for each country is: $L = c + \sum_f \beta_f * F_f$. In the equation, F denotes the factors of learning, β_f the regression coefficient, and c the constant.

Note: *significant at p<.05, **at p<.01, ***at p<.001. The region 'East Asia and Pacific' is taken as the reference category in (2).

Model (1) represents the best fit using GDP per capita, primary learning assessments, and a dummy gender variable. When primary learning data was available, this model was used to predict secondary numeracy and literacy learning. Model (2) represents the second-best fit using socio-economic and regional variables. When primary learning data was unavailable in any year for a given country, this model was used to predict secondary numeracy and literacy learning. The resulting data were used to calculate growth rates for each country and gender where two or more historical or imputed data points are available. The average growth rate was then calculated for each income group, and the remaining countries without growth rates were assigned their income group average. The floor of the growth rates is set to 0. To predict secondary numeracy and literacy learning for both genders combined, we used the population-weighted mean of males and females.

After filling in missing data from years 2010 to 2019, linear trend forecasting was used to predict country values for the year 2020. The predictions for the year 2020 are multiplied by a factor representing the effect of learning loss due to COVID-19. The calculation of this multiplier is detailed in Section 4 of this document. These estimates were then combined with population projections, using the medium SSP2 scenario. Population data are available by 5-year age groups, gender, and six educational attainment categories: no education, incomplete primary education, primary education, lower secondary education, upper secondary education, and post-secondary education. For everyone who completed at least lower secondary education, we assumed that the share of people with basic skills is identical to the estimated share from the empirical data/statistical modelling. An additional adjustment is made for the out-of-school youth: the share is estimated to be half of the calculated share for the population who only finished primary school, 21.5% for those who had completed some years of primary school, and 8.5% for those who never went to school. These adjustments are based on PISA for Development (PISA-D) and Multiple Indicator Cluster Surveys (MICS) Foundational Learning Skills data which include separate assessments for the out-of-school youth.

We assume the growth rates calculated for the time period before 2020 are unchanged for the years 2021 to 2050. Thus, once the adjusted value for 2020 is set for each country, linear trend forecasting was used to predict country values for the years 2021 to 2050.

3. Digital skills

Digital skills support the development of skills to use and understand technology, search for and manage information, communicate, build knowledge and solve problems. To assess the proportion of youth with digital skills, data were primarily taken from the MICS, available on the UNICEF Global Database, and other household surveys as reported in the UNESCO Institute for Statistics (UIS) database, for the monitoring of SDG Indicator 4.4.1 on the proportion of youth and adults with ICT skills, by type of skill. When both MICS and UIS data are available for a country, MICS data are used. Data for China[6] and India[7] were extracted from national

[6] *Blue Book of New Media 2021: Annual Report on Development of New Media in China No. 12 (2021)*, Social Sciences Academic Press (China), 2021.

reports on the proportion of adults with basic ICT skills. Data for the United States[8] were extracted from the Programme for International Assessment of Adult Competencies (PIAAC) on the proportion of youth with digital literacy (i.e., proportion of 16- to 24-year-olds who participated in the 'problem solving in a technology rich environment' assessment).

Available gender-disaggregated data were extracted for the following skills: copying or moving a file or folder, using copy and paste tools to duplicate or move information within a document, sending e-mails with attached files, and transferring files between a computer and other devices.[9] For each country, the average among these figures is taken as the proportion of youth and adults with basic digital skills for each country.

Based on PIAAC[10] results suggesting youth aged 16–24 years are 24 per cent more likely to have basic ICT skills than the total 16- to 65-year-old population, gender-disaggregated data on youth and adults (i.e., data from UIS and national reports on proportion of youth and adults with basic ICT skills) were narrowed to youth by multiplying each value by 1.24, capping the final country values at 100 per cent.

For countries with missing data on digital skills, regression-based imputation was conducted using the natural log of gross national income per capita (calculated using the World Bank Atlas method), the percentage of individuals using the Internet, the percentage of individuals with access to electricity, and a regional dummy variable. Data on the first three variables were retrieved from the World Bank Open Data, while countries' regions were based on UNICEF's categorization of regions aligned with the SDG regions. For countries with missing data on GNI per capita, internet usage, or access to electricity, the average value of their respective country income groups was used. Table 2 below shows the regression results using the latest available data for each country.

Table 2. Regression results for the digital learning production functions.

Ln GNI per capita	10.37***
Internet usage (% of population)	0.45***
Access to electricity (% of population)	-.43***
Region: Eastern Europe and Central Asia	-0.21
Region: Eastern and Southern Africa	-22.10**
Region: Latin America and Caribbean	-12.57*
Region: Middle East and North Africa	-5.00
Region: North America	14.35
Region: South Asia	-4.92

[7] Government of India Ministry of Statistics & Programme Implementation, *Household Social Consumption on Education in India: NSS 75th Round*, July 2017-June 2018.
[8] Mamedova, Saida, and Emily Patakowski, 'A Description of U.S. Adults Who Are Not Digitally Literate', Statistics in Brief NCES 2018-161, U.S. Department of Education National Center for Education Statistics, May 2018.
[9] These four computer-related activities are used by ITU for the reporting of basic ICT skills (e.g., see: International Telecommunication Union, *Measuring the Information Society Report 2018, Volume 1*, ITU, Geneva, 2018).
[10] Organisation for Economic Co-operation and Development, *Skills Matter: Additional Results from the Survey of Adult Skills*, OECD Skills Studies, OECD Publishing, Paris, 2019.

Region: West and Central Africa	-20.96**
Region: Western Europe	1.41
Constant	-33.72
N	112
R-squared	0.83

Note: *significant at p<.05, **at p<.01, ***at p<.001. The region 'East Asia and Pacific' is taken as the reference category.

The model was applied to available gender-disaggregated data for each year from 2014 to 2019 to fill in missing values for each year. For countries resulting with imputed values of '0%' for 2019, the highest figure from the last 10 years is imputed, or, if no data is available for previous years, a value of '1%' is imputed. Average growth rates were then calculated for each country; countries with declining or flat trends were assigned the average growth rates of their respective country income groups. These adjustments are based on the assumption that digital skills have been accelerated by the COVID-19 pandemic starting in 2020. Then, linear trend forecasting was used to predict country values for the year 2020, which is the base year for projections.

We assume the growth rates calculated for the time period before 2020 are unchanged for the years 2021 to 2050. Thus, once the adjusted value for 2020 is set for each country, linear trend forecasting was used to predict country values for the years 2021 to 2050.

After completing the imputations and projections separately for the male dataset and female dataset, the simple average of the two was computed to provide overall figures for each country. These estimates were then combined with the population data as outlined in Section 2.[11]

4. Modeling for the COVID-19 Shock

We apply the equation below to estimate the COVID dampening effect on skill attainment for youth and in and out of education.

pre-COVID youth skill rate X (1 – in-person school time lost / pre-COVID length of country school life X (1 – pre-COVID Internet access rate))

In this equation, the *in-person school time lost / pre-COVID length of country school life* is used to proxy learning loss if no mitigation. We use data from the UNESCO Global Monitoring of School Closures[12] to estimate the loss of in-person schooling year as of 28th February 2022, which covers 743 days or about 2.04 years with traceable records of school closures. Each partially open day is counted as a 0.5 full-closure day. For each country, we multiply the 2.04 year with "(full-closure days + partial closure days/2) divided by total non-academic-break

[11] Unlike the secondary-level skills data which are taken from school-based assessments, the digital skills data are taken from household surveys. For this reason, we do not apply the same adjustment for out-of-school youth to the digital skills data as has been done for the secondary-level skills data.
[12] https://en.unesco.org/covid19/educationresponse

days" to estimate in-person schooling lost in year unit. We then divide lost year by the school life expectancy of primary and secondary education, which is available on the UIS.

The *pre-COVID internet access rate* from ITU is used to proxy a country's capacity to mitigate the learning loss. It is true that online learning is just one mitigation measure. We therefore run a correlation test between the pre-COVID internet access rate and the UNICEF remote learning readiness index[13] (RLRI) covering 67 countries. The result shows 0.63 correlation, implying a moderate-to-strong correlation. RLRI is composed of three domains: households, a government's policy response capacity, and the emergency preparedness of the national education sector calculated from multiple data sources of household surveys, government surveys and UNICEF's Strategic Monitoring Questions.

Together, *1 – in-person school time lost / pre-COVID length of country school life X (1 – pre-COVID Internet access rate)* signals what percentage of pre-COVID skill achievement can be preserved given the school-closure situation as of 28th February 2022. Note that in this modelling pre-COVID means data for 2019 or the latest available pre-2020. For simplification, we also assume a same dampening effect to both youth in and out of education, and that the additional COVID-related dropout is offset by the pre-COVID decline trend in out-of-school.

[13] https://data.unicef.org/resources/remote-learning-readiness-index/