

Methodology Section

Given the need to harmonize national and international data sources and definitions, we describe below three major elements of our methodology in detail: (i) combining national level labor market information with absolute population and educational attainment data, (ii) harmonizing subnational level labor market and educational attainment information with national level estimates, and (iii) forecasting disaggregated numbers up to 2030.

(i) Achieving a complete set of disaggregated, national-level estimates on Youth Employment

Our objective in this exercise is to obtain a complete set of disaggregated numbers for each country, in each year, and grouped by labor force status (employed, unemployed, [fully] inactive¹, student/in training), and disaggregated by age, educational attainment, and sector of employment (if employed).

In this first step, we combine three types of information:

- a. national level youth labor force *shares* disaggregated by age groups (15 - 24, and 25+), gender (male and female), and education attainment levels (less than primary, primary, secondary and tertiary),
- b. Youth employment *shares* by sector (14 subsectors, which in a simplified view can be aggregated to Agriculture, Industry, and Services), and
- c. national level *absolute* population numbers. Both a. and b. were sourced from ILO's official statistics database, whereas c. was sourced from the International Institute for Applied Systems Analysis (IIASA).²

We first interpolate IIASA's 5-year estimates using sprague (osculatory) multipliers to receive single-age population for each year disaggregated by gender and education level attainment. We then apply ILO's unemployment, NEET (Not in employment, education or training), and inactivity (out of labor force) shares to the population data. To be specific, we first recover the population in the labor force using the inactivity (out of labor force) share. Then, we calculate the number of employed/unemployed by using the unemployment rate on the population in the labor force. Finally, we indirectly extract student shares by subtracting the share of employed and the share of NEETs (Not in Employment, Education, or Training) from one (1) and apply this share to the population. The rest of the population (after applying final employed, unemployed, and student shares) are called "inactive", which due to our separate student category is somewhat more narrowly defined than in most ILO statistics. Whenever it was necessary to make use of a particular definition of labor market indicators, we followed [ILOSTAT's guidelines](#).

Finally, once these groups were calculated, we further grouped the employed population into their respective sectors, again using ILO's shares. Additional data sets on formality of employment, working poverty, and urban/rural unemployment rates were collected and prepared, but not integrated into the primary, comprehensive dataset due to a lack of combined microdata (i.e. the available data structure did not allow for comprehensive merging).

(ii) Harmonizing subnational surveys with national level information

The objective of integrating national surveys with subnational information is to harmonize the available information across several focus countries, in order to provide comparability and

¹ILO considers people to be “inactive” if they are out of the labor force, which includes students. In our case, we explicitly calculate who is a student or in training, which means that our “inactive” group is slightly narrower than ILO’s.

²WDL uses IIASA’s population forecasts rather than the UN’s World Population Prospects due to IIASA’s inclusion of education as a demographic dimension relevant for fertility as well as their provision of a multitude of scenarios (Shared Socioeconomic Pathways (SSPs)).

consistency. To achieve this, we stayed as true to the microdata as possible, but aggregated local educational and sector definitions to conform with international groupings, and scaled population numbers proportionally and iteratively³ to match our results from (i).

Our source data for Rwanda were labor force surveys from the National Institute of Statistics of Rwanda ([NISR](#)) for the years 2017 - 2022. For Kenya, we accessed labor market data from the Kenya Population and Housing Census (2019 and 2021) from the Kenya National Bureau of Statistics ([KNBS](#)). Subnational data will become available for Ghana, Nigeria, Uganda, Senegal, and Ethiopia, where we will follow the same approach.

(iii) Forecasting disaggregated numbers up to 2030

Similar to forecasting approaches used by, for example, [ILO \(2010\)](#) or [Huruta \(2023\)](#), we forecast unemployment, inactivity, and NEET shares. We make use of a widely established relationship between these indicators and GDP (Gross Domestic Product) per capita, and employ an ARIMAX (Autoregressive Integrated Moving Average with an eXogenous Regressor) model. The model predicts the change in the rates of the variables of interest based on historical values and the influence of GDP per capita growth rates on the time series. This was the best performing model out of a variety of tested approaches based on out-of-sample mean-squared error (MSE).

We leveraged two pre-existing datasets which provided us with population and GDP forecasts up to 2030.

- a. Youth population projections disaggregated by single age, gender and educational attainment up to 2030, forecasted by IIASA. (Please consult their website for further information on assumptions and methodology).
- b. National GDP per capita growth rates, forecasted up to 2028 By the International Monetary Fund (IMF’s) World Economic Outlook (WEO) team, and extended to 2030 by WDL. To extend the IMF’s forecasts, we follow [Crespo Cuaresma \(2017\)](#), which uses a simple production function with human capital as the primary driver. We furthermore use a smoothing model for the years 2028/2029 for a sensible transition between the short-term and long-term projections.

In **subnational datasets** we furthermore computed compound annual growth (CAGR) rates of employment by sector to estimate these values until 2030. Similar to what was outlined in (ii), we rescale subnational forecasts iteratively to match national-level forecasts.

Overall, around 6 - 8 years are forecasted depending on the latest year of data available from the original source data.

Important assumptions and treatment of missing data

³ Scaling had to be done iteratively because subnational data included merged information of sector and educational attainment shares, whereas in the national data this information was separate. Therefore, scaling considered shares by educational attainment and by sector in turns.

In some cases, labor force shares were only available for 15-24 and 25+ years, requiring our data science experts to make a decision on what rates might be most reasonable for the subgroup of 25-35 year-old youth. For the initial version of the tool we opted to assume the available rates for people aged 25+ applies to youth 25-35 for most of the data. For the NEET share, however, we decided to use age-based patterns found in survey data for Kenya/Rwanda. Specifically, we apply the average ratio (15-24 : 25-35) to all the other countries. We understand that these assumptions may affect the results significantly and are currently in the process of acquiring better source data for the age-range of interest.

Missing values in *national-level data* due to irregular surveying were imputed using either linear interpolation, regional averages, or forward and backward filling; depending on the existing data structure and using our best judgment for what is the most appropriate method in each specific case.

In *subnational microdata* missing data points were estimated using linear interpolation (for time-related missing values) or Multivariate Imputation by Chained Equations (MICE) (for cross-sectionally missing values).

Further limitations may arise from the source data used in this clock. WDL uses data from different sources such as ILO's modeled estimates, IIASA's population and IMF's Gross Domestic Product projections and as such the limitations of the data highlighted by these organizations may apply to the data on this Clock. More from [ILO](#) and [IIASA](#).

Glossary of Terms

Please note that this is a summarized version of the full definition from ILO by WDL, for the complete and full definition please refer to the official documentation from ILO as linked in the column named "Source".

Terminology	Definition	Source
Employed	Part-time, seasonal, or full-time employment within an organization which commences after an individual secures income from this work. Includes 3 subsets: new wage, sustained wage, and improved wage employment. Note also that this category includes self-employment (youth engaged in entrepreneurship).	ILO
Unemployed	Persons of working age who are without work, seeking work (carried out activities to seek employment during a recent past period), and currently available for work.	ILO
Inactive	Persons outside the labor force (they are jobless but not available and/or not looking for a job), including people engaged in household or family duties full-time.	ILO
Youth in Education or Training (Student)	Persons in education programs or school-based technical and vocational training programs. Excluded from this group are persons in employer-based training, youth in unemployment and also in education or Training Note: This is a derived indicator by WDL based on the NEET shares from ILO.	ILO
Rural/Urban	There is no internationally agreed definition for rural and urban areas. Therefore, the differentiation between these area types is made according to national definitions. Since data are obtained from household surveys, the determination for the area type is based on the location of the household (i.e., the dwelling) rather than that of the job (if any of the household members are employed).	ILO
Formal/Informal Jobs	Informal employment is defined as any activity of persons to produce goods or provide services for pay or profit that is -in law or in practice- not covered by arrangements such as commercial laws, procedures to report economic activities, income taxation, labor legislation and social security laws and regulations providing protection against economic and personal risks associated with carrying out the activities. Formal employment is defined as any activity of persons to produce goods or provide services for pay or profit that is -in law or in practice- covered by arrangements such as	ILO (1) ILO (2)

	<p>commercial laws, procedures to report economic activities, income taxation, labor legislation and social security laws and regulations providing protection against economic and personal risks associated with carrying out the activities.</p> <p>Informal sector: All workers in unincorporated enterprises that produce at least partly for the market and are not registered. It excludes households that produce exclusively for own final use, subsistence agriculture, construction of own dwellings, etc.</p> <p>Formal sector: all workers in incorporated enterprises.</p>	
Working Poverty	<p>Working poverty reveals the proportion of the employed population living in poverty despite being employed, implying that their employment-related incomes are not sufficient to lift them and their families out of poverty and ensure decent living conditions. For the purpose of international comparability, extreme poverty is the category of employed youth living below the absolute international poverty line of US\$2.15 per capita per day at 2017 PPP, moderate poverty is the category of employed youth living above this international poverty line, but below the estimated line that is more typical of lower-middle-income countries at US\$3.65 per capita per day at 2017 PPP. The "not poor" category of employed youth living above both these lines.</p> <p>Working poverty rate is the share of employed persons living in households with consumption or income that is below the international poverty line of US\$2.15 per-capita (2017 PPP).</p>	<p>ILO</p> <p>UNST ATS</p>
Labour Force (LF)	Composed of the employed and unemployed population.	ILO
Not in Employment, Education, or Training (NEET)	Persons not in employment, education or Training.	ILO