\*\*\*\*\*\*\*\*\*\*

Country level file

Emigration and immigration rates 2000, 2010 and 2020

Code edited for WDR by Narcisse Cha’ngom

\*\*\*\*\*\*\*\*\*\*

Data Files

The final datafile named “emigration-immigration-rates.xlsx” reports at country level the total number of emigrants and immigrants by gender and education level. This dataset is restricted to foreign born population aged 15+ and thus disregards children.

***Important note***: In 2020, there is 5.7 million more immigrants than emigrants. This is solely due to the stateless who reside in well specified destination countries but do not have a specified origin country. So, they appear in the destination, but not at the origin. The way we circumvent this issue in the bilateral migration matrix has been to create a residual origin country called “Other” to which we attributed them. Given the primary purpose of country level datafile, i.e., calculate migration rates, we decided not to create an origin country called “Other” as a population size cannot be attributed to it and consequently emigration rates could not be calculated.

Data are combined in the following way:

1. The first step consists of collecting human capital data for 2000, 2010 and 2020. The benchmark data source for human capital (i.e. the gender specific proportion of tertiary educated in the population aged 15+) is Barro and Lee dataset. But as for 2020, Barro and Lee do not exist, we retrieve human capital data from three different sources: 2020 ACS for the US, 2020 EULFS for European countries and Wittgenstein Centre for the rest of the countries. It is worth noting that for as for 2000 and 2010, Barro and Lee do not provide data for all countries in the World. So, for countries missing in Barro and Lee, we use data from Wittgenstein Centre.
2. In the second step, we gather data on socio-economic and demographic characteristics from The World Bank World Development Indicators. This includes total population, adult population (15+), children (0-14), income per capita
3. In the third step, we combine the first two steps allow to compute the gender specific, education specific labor force.
4. The fourth step consists of aggregating the bilateral migration data along the dimensions described above. The aggregation is done at origin-year level to obtain the number of emigrants by gender and education. It is then done at destination-year level to obtain the number of immigrants.
5. In the fifth step, using steps 3 and 4, we could compute the emigration and immigration rates as follow.

5.1- emigration rates

Let be the emigration rate in the sending country for education level , gender at the year .

Similarly, gender specific immigration rate is obtained as:

Where is the immigration rate in destination country for gender at year .

The code should be run in the sequence presented above.