The inequality measure used in this research is the WIID Companion dataset from the United Nations (2023). This dataset provides information on annual per capita income distributions at both the country and global levels. The variables of interest are the standardized Gini Index and the difference in income share of the top and bottom 20 percent earners. The dataset partially covers 201 countries from 1890 to 2022. These two variables are chosen because the Gini Index offers a comprehensive measure of overall income inequality, while the income shares of the top and bottom 20 percent highlight disparities in wealth distribution, allowing us to assess inequality from both a comprehensive and a distribution-sensitive perspective, and to check whether the results are specific to one particular measure.

Our explanatory variable, which is the reliance of a country’s economy on remittances, is represented by the personal remittances received as a percentage of GDP (World Bank, 2024). These values are based on World Bank staff estimates, using IMF balance of payments data as well as World Bank and OECD GDP estimates. IMF data on worker remittances is based on official banking reports and may underestimate actual flows, as informal transfers are not accounted for. Consequently, the estimated coefficients likely represent a lower bound of remittances' true impact on income inequality (Koechlin and Leon, 2006). The dataset partially covers 266 countries from 1890 to 2023.

Following Murodova (2018), the skill composition of migrant workers is considered an important factor when examining the impact of international remittances on inequality. To account for this, I use school life expectancy as a proxy for education levels across countries. School life expectancy refers to the total number of years of schooling a child entering the education system can expect to receive, assuming current age-specific enrollment ratios remain unchanged. The data are drawn from two datasets covering expectancy for ISCED levels 1–3 and ISCED levels 5–8 (United Nations, 2016). ISCED levels 1–3 correspond to primary through high school level education, while ISCED levels 5–8 are equivalent to short-cycle tertiary education through to doctoral studies. The former reflects general accessibility to basic education, whereas the latter indicates the availability and development of higher education. Both datasets provide partial coverage for the period 1975–2016, with the ISCED 1–3 dataset covering 216 countries and the ISCED 5–8 dataset covering 193 countries.

Additionally, we aim to include data that reflect a country's overall economic conditions. For this purpose, we use the World Economic Outlook database (IMF, 2024), which provides partial coverage of 44 different economic indicators across 196 countries from 1980 to 2023. To ensure cross-country and temporal comparability, we retain only those variables expressed in indices, percentages, or purchasing power units. These variables are included in the regression as control variables to account for differences in economic status across countries. While they serve a similar role to country fixed effects by capturing country-specific heterogeneity, their advantage lies in avoiding potential collinearity with time fixed effects. In addition, to minimize multicollinearity among the control variables themselves, we performed a correlation check and excluded variables with pairwise correlations exceeding 0.6. The final set of control variables includes Real GDP, GDP per capita, Total investment, Gross national savings, Inflation, Unemployment rate, General government total expenditure, General government net lending/borrowing, and Current account balance.

It is important to note that all the datasets we use are incomplete. Therefore, our final research dataset is constructed as the intersection of all the aforementioned datasets and transformed into a panel data format. After processing and cleaning, the final dataset consists of 58 observations, covering 72 different countries over the period from 1980 to 2015. The definitions of all datasets are provided in the Variable Descriptions Table.