MSR

Working Papers 001-2022

Feb 2022

Economic Development and Income Inequality: Findings from Regional Panel Data Analyses

John Michael Riveros-Gavilanes, Ryan Jacildo, Brian Kiberu, Imen Nouira, and Sherif M. Hassan

Publisher and Distributor

M&S Research Hub institute

Carlo Mierendorff Str. 43, 34132 Kassel, Germany

Telephone +49 (0)56149941680,

Email: info@ms-researchhub.com,

Web: www.ms-researchhub.com

The findings, interpretations and conclusions expressed in this publication are entirely those of the author(s).

Economic Development and Income Inequality: Findings from Regional Panel Data Analyses

John Michael Riveros-Gavilanes[1], Ryan Jacildo[2], Brian Kiberu[3], Imen Nouira[4], and Sherif M. Hassan[5]

23/01/2022

Abstract

The inequality-economic development nexus has been the focus of a number of research undertakings in the past. Nonetheless, the variations in regional dynamics are still arguably understudied. Employing semi-parametric regressions, this research finds that the classical hypothesis of Kuznets is not well-established in all regions of the world. This reflects the heterogeneities present across economies and regional groups. The regions where the inverted U-shape relationship between income inequality and economic development (Kuznets hypothesis) can be observed are East Asia and the Pacific and Latin America and the Caribbean. In Sub-Saharan Africa, South Asia, Europe and Central Asia, the data suggest an N-shape relationship. In the Middle East and North Africa, the data interestingly present a negative correlation. Meanwhile, in North America inequality appears to increase as the real income per capita increases.

Introduction

The inequality-economic development nexus remains a key policy area in many economies, especially with the speed of structural transformation that is transpiring and the increasing importance of digitalization. The subject has been the focus of a number of research undertakings in the past. Nonetheless, the variations in regional dynamics are still arguably understudied. It is on this ground that this research is undertaken.

In pursuing the research objectives, the study appeals to the proposition of Kuznets (1955, 1963) that posits an inverted U-shape relationship between inequality and economic development. Incidentally, there is no established consensus on this matter in the empirical literature. The analysis in this study utilizes a semi-parametric approach to take into account

[1] M&S Research Hub
[2] M&S Research Hub
[3] University of Nairobi
[4] Ministry of finance/Tunisia, Statistician,
[5] M&S Research Hub, British University in Egypt, CNMS-Marburg University, ERF

the non-linearities and it draws from the panel dataset of 151 economies that are categorized into seven geographic regions.

The subsequent elements of the study are organized as follows: Section 2 lays out the literature on the subject matter; Section 3 presents the analytical framework; Section 4 discusses the data, data sources, and the data characteristics; Section 5 details the empirical exercise methodology; Section 6 outlines the results and findings; And Section 7 concludes.

Review of Related Literature

The relationship between inequality and economic growth has been a subject of numerous studies over the years. The inverted U-shaped relationship between these variables as proposed by Kuznets (1955, 1963) is a common starting point of many analyses on the subject matter. The resulting Kuznets curve suggests that economic development tends to be firstly associated with an increase in economic inequality. However, after a certain threshold is breached, economic development tends to lower inequality.[6]

Some theories on the association of economic growth and income distribution appeal to the level of maturity of economic and technological development in a country (Ahluwalia 1976, Robinson 1976, Gupta and Singh 1984, and Barro 2000). In relation to Kuznets' findings, it is purported that inequality tends to increase during the early stage of economic technological development, and to decrease as the economy moves to the mature stage.

The direction of impact similarly goes the other way. For instance, it is argued that inequality can affect growth by boosting aggregate saving (Kuznets 1955; Kaldor 1955), (ii) boosting research and development (Foellmi and Zweimüller 2006); and (iii) inducing high-return projects (Rosenzweig and Binswanger 1993) Conversely, inequality can force governments to implement expensive redistribution mechanisms (Perotti 1993) and induce political instability (Agnello, Castro, Jalles, and Sousa 2017) and inefficient government bureaucracy (Acemoglu, Ticchi, and Vindigni 2011). Furthermore, given the link between inequality and poverty, and considering that the poor are credit constrained, it is suggested that high income inequality reduces investment and hampers growth in the long run (Galor and Zeira 1993, Piketty 1997, Panizza 2002). There is also evidence that inequality increases poor people's fertility, thus, limiting their ability to accumulate and build human capital (de la Croix and Doepke 2003).

Empirically, the dynamics of the relationship can change depending the definition of inequality metric used, the quality of the data, and the processes employed. For instance, in some studies, market inequality is the focus of some research, whereas in others, it is structural inequality or net inequality.

[6] In a related albeit different take on the issue, some studies have examined the growth in income-by-income group, i.e. the dynamics of income distribution. In this respect, one emerging observation is the elephant curve, which purports that income growth tends to rise with income group initially (up to a particular income group), then it declines before rising again sharply in highest income group (see: Alvaredo et al. 2017).

Banerjee and Duflo (2003) use non-parametric methods to investigate the relationship between income inequality and growth using cross country data. The study finds economic growth rate is an inverted U-shaped function of net changes in inequality. It also underscores that this non-linear relationship "is sufficient to explain why previous estimates of the relationship between the level of inequality and growth are so different from one another."

With an almost similar perspective, Forbes (2000) examines the relationship between income inequality and economic growth using a panel estimation approach. The approach enables the control of time-invariant country-specific effects, thereby removing a source of omitted variable bias. The study discovers that an increase in a country's income inequality has a significant positive relationship with subsequent economic growth in the short and medium-term.

Meanwhile, Halter, Oechslin, and Zweimüller (2014) use a system generalized method of moments to estimate the short term and long-term effect of income inequality. The dataset is from 1965 to 2005 with an interval of 5 years and covers 106 countries. The study finds a positive relationship between higher income inequality and economic growth in the short term and a negative relationship between the two in the long run.

In a study with a slightly broader scope, Kyosuke and Takashi (2011) examine the relationship between poverty, income inequality, and economic growth using a generalized method of moments. The study uses province-level panel data derived from microdata on household expenditure in Thailand (1988–2004) and the Philippines (1985–2003). The study concludes that inequality slowed per-capita consumption growth and that differences in inequality account for a sizable portion of the gap between Philippines and Thailand in terms of growth and poverty reduction since the late 1980s.

Using a construct of net income inequality, Berg et al. (2018) examine the relationship between inequality, redistribution, and growth using panel growth regressions. The authors employ a dataset that distinguishes between market (pre-tax and transfer) and net (post-tax and transfer) inequality to calculate redistributive transfers for advanced and developing countries. They find that lower net inequality is strongly correlated with faster and more durable growth, even when the level of redistribution is controlled for, and that higher inequality appears to retard growth.

Climent (2010), on the other hand takes into account human capital. Using a dynamic panel model to examine the effect of human capital and income inequalities on economic growth in various world regions, the study posits that inequality affects growth depending on the country's stage of development. Both human capital and income inequalities have a negative effect on economic growth in low and middle-income economies. In contrast, they have a positive impact on economic growth in higher-income countries.

There are also studies that look into endowment and wealth inequality. For example, Easterly (2007) uses cross-country data to investigate the relationship between inequality and development–the study instruments inequality with agricultural endowments. Agricultural endowments are measured using an abundance of land suitable for growing wheat relative to that suitable for growing sugarcane. The study finds higher inequality is a barrier to prosperity, high schooling, and good quality institutions.

Separately, Alesina and Rodrik (1994) examine the relationship between economic growth and redistribution of wealth using a political economy model of growth based on the median voter theorem. In a model that requires the government to choose a median voter's tax rate, the study estimates wealth distribution using land distribution and income distribution measures. After controlling for initial income and human capital levels, the authors find a statistically significant negative relationship between land distribution inequality and economic growth, as well as a negative relationship between initial income inequality and growth of production.

Analytical Framework

The starting point of the analysis is the Kuznets (1955) hypothesis that posits that the extent of inequality is a function of economic development as discussed in the previous section. This perspective considers how economies transition in terms of sectoral activities. The Kuznets proposition is underpinned by the idea that economies start their development path relying on primary-extractive production. At this stage, high levels of inequality can be experienced until a certain point; however, as industrialized production gains more traction, inequality declines (Baymul and Sen 2020). This view is encapsulated in the inverted U-shape curve that describes the relationship between real income per capita and the inequality.

In the context of the neoclassical growth model with exogenous saving rates, Barro and Sala-i-Martin (2004) provide a theoretical base to empirically analyze the inequality dynamics. In the model, the production function of the general form in per capita terms is given by Equation 1.

$$y_i(t) = f\big(k_i(t)\big) \tag{1}$$

In Equation 1, y is the output per capita for a certain country i at time t is defined as a function of the levels of capital per worker k. Notably, $f(.)$ should satisfy all neoclassical assumptions. Equation 2 defines the growth of the economy in per capita terms.

$$\gamma_y = \frac{\Delta y_i(t)}{y_i(t)} \tag{2}$$

In this framework, the starting levels of capital $k_i(0)$ for the country i determine the path and pace of convergence towards the steady state of the economies. Growth rates γ_y of the richer countries (i.e. countries with larger initial capital stock) will be lower than the growth of poorer countries, i.e. $\gamma_y^{rich} < \gamma_y^{poor}$. This framework is in concordance with Kaldorian approach wherein more industrialized economies (characterized by having larger capital stock) tend to grow at slower rates (Kaldor 1966) in comparison to developing and non-industrialized economies. Against this backdrop and as introduced by Barro and Sala-i-

Martin (2000), inequality can be represented by the variance of the sample economies (Equation 3).

$$D_t \equiv \frac{1}{N} \sum_{i=1}^{N} \left[\log y_i(t) - \bar{y}\,(t)\right]^2 \tag{3}$$

In Equation 3, $\bar{y}$ is the cross-country average of the natural logarithm of per capita income at time t, y_i is the per capita income of country i, and D_t is the inequality at time t. The above expression can be interpreted as the dispersion of per capita income at time t. Equation 3 entails that when $\log y_i(t) > \bar{y}\,(t)$, the dispersion is positive or inequality exists; and when $\log y_i(t) = \bar{y}\,(t)$, inequality is nil. It is worth noting that this approach does not imply that inequality will necessarily converge considering the difference in initial conditions (Barro and Sala-i-Martin, 2004; Barro and Sala-i-Martin, 1992). The view is that the extent of inequality is a direct consequence of the characteristics of the production function $f\big(k_i(t)\big)$ and the starting level of capital accumulation, which includes all types of capital and not only physical capital in endogenous growth models.

In summary, inequality could be interpreted as deviations between economies; and the dispersion is dependent on the stages of economic development. The analysis of Baymul and Sen (2020) is relevant in emphasizing the role of structural transformation in examining the inequality dynamics. Meanwhile, in the framework of Barro and Sala-i-Martin (2004), the association between inequality and economic development is influenced by the initial conditions, particularly by the level of capital stock.

Data, sources, and summary statistics

The data used in this study are sourced from the Penn World Table 10.0 (Feenstra et al. 2015), International Monetary Fund World Economic Outlook October 2021 database (2021) and the World Bank World Development Indicators database (2021). The resulting dataset covers 151 economies from 1990 to 2019, clustered into seven regional groups (Table 1) based on the definitions of the World Bank. More than half of the economies in the dataset belong to Europe and Central Asia and Sub-Saharan Africa. The remaining countries are subsumed under East Asia and the Pacific, Latin America and the Caribbean, Middle East and North Africa, North America, and South Asia.

Globally, inequality has declined in the last thirty years and the dispersion across countries has narrowed. As can be gleaned from Table 1, the average Gini coefficient index of the 151 economies dipped from about 40.8 in the baseline period to 38.4 in recent years. At the same time, the standard deviation decreased by about 2 index points. The drop in the inequality, notably, coincides with rise in real per capita income, average years of schooling, total investment-to-GDP ratio, and employment rate (or the decline in unemployment rate).

Table 1. Number of Economies by Region in the Dataset

Region	Number of economies with data

East Asia and the Pacific	11
Europe and Central Asia	47
Latin America and the Caribbean	25
Middle East and North Africa	15
North America	2
South Asia	7
Sub-Saharan Africa	44

Note: Only economies with Gini coefficient data are included in the calculation. The regional groupings are based on the definitions of the World Bank.
Source: Authors.

Table 2. Summary Statistics, All economies, 1990-2019

Global sample	Number of observations	Mean	Standard Deviation	Minimum	Maximum
Earliest available data					
Gini coefficient index	151	40.8	10.01	20.7	65.8
Real GDP per capita	149	11,972	13,029	475.4	62,879
Human capital, average years of schooling	131	2.228	0.702	1.045	3.476
Total investment, % of GDP	138	23.04	10.27	-7.957	61.5
Unemployment rate, %	83	9.991	6.22	1.38	32.17
Latest available data					
Gini coefficient	151	38.35	8.188	24.6	63
Real GDP per capita	151	17,942	18,412	619.3	91,850
Human capital, average years of schooling	131	2.608	0.723	1.193	3.821
Total investment, % of GDP	141	24.38	7.991	7.261	54.33
Unemployment rate, %	94	8.32	5.256	1	26.9

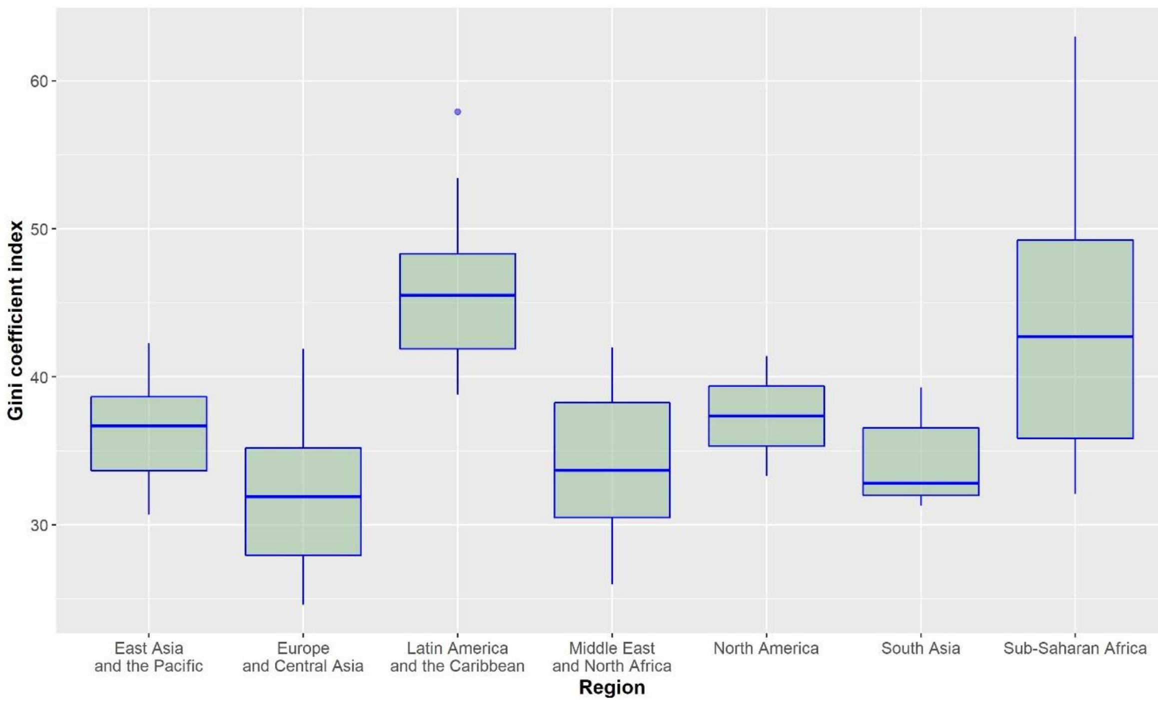
GDP = gross domestic product.
Notes: Earliest data refer to the first data points available from 1990 to 2019 by country, while the latest data refer to the most recent data points available. The data for GDP per capita, human capital, total investment, and unemployment rate included in the calculation correspond to the period of the available Gini coefficient data for every country. For instance, if the latest Gini coefficient data of a country is for 2015, the latest GDP per capita included in the calculation for the same country is also for 2015. There are periods when Gini coefficient data are available, but not the other variables.
Source: Authors, based on Feenstra et al. (2015), World Economic Outlook October 2021 database (2021), and the World Bank World Development Indicators database (2021).

Zooming in on the Gini coefficient indices and the real per capita income indicates some divergence across the regions. The most recent data reveal that the extent of income inequality in Latin America and the Caribbean and Sub-Saharan Africa is the most concerning of all the regions (Figure 1a). East Asia and the Pacific, Middle East and North Africa, North America, and South Asia are middling. Meanwhile, income distribution is most favorable in Europe and Central Asia.

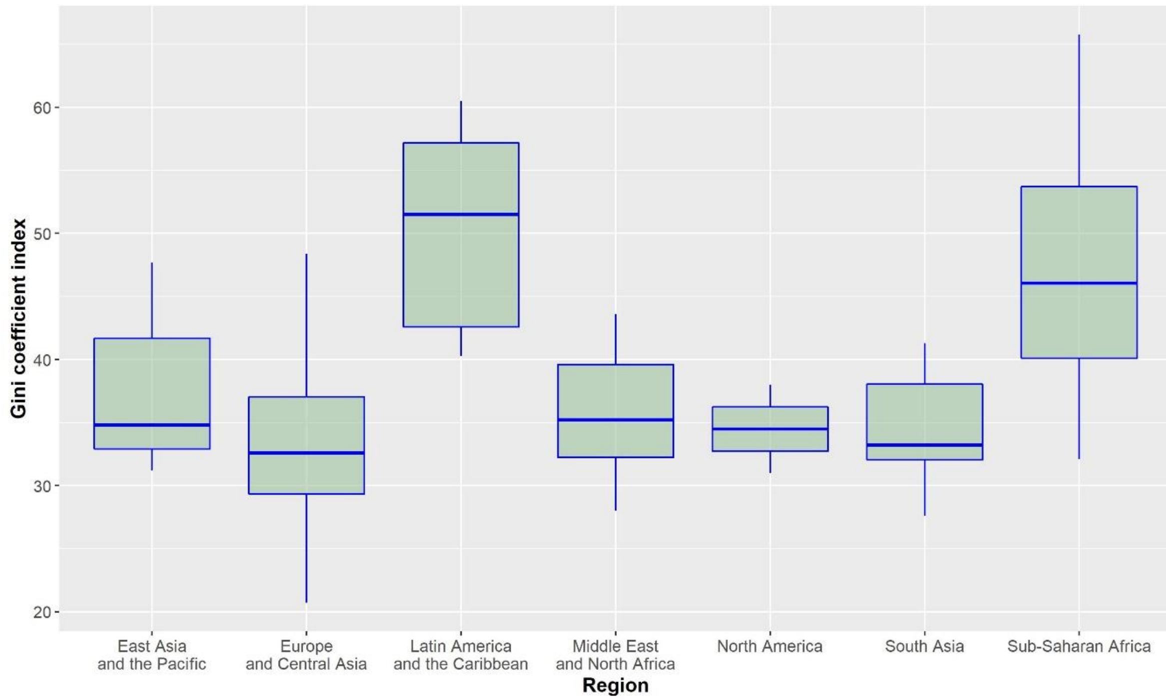
Relative to the average baseline values (Figure 1b), i.e. earliest data points from 1990-2019, the improvement in income distribution based on the median values is apparent in Latin America and the Caribbean and Sub-Saharan Africa; and to some extent, in Europe and Central Asia. Change is hardly noticeable in East Asia and the Pacific, Middle East and North Africa, and South Asia; whereas in North America, inequality has somewhat widened. Roughly the same story is conveyed by the averages or the mean values (see Table A1 in the Appendix). With the exception of North America, the distribution of the Gini coefficient indices has also become less dispersed in recent years within the regional groups.

Figure 1. Gini coefficient index by Region

a. Latest data available, 1990-2019



b. Earliest data available, 1990-2019



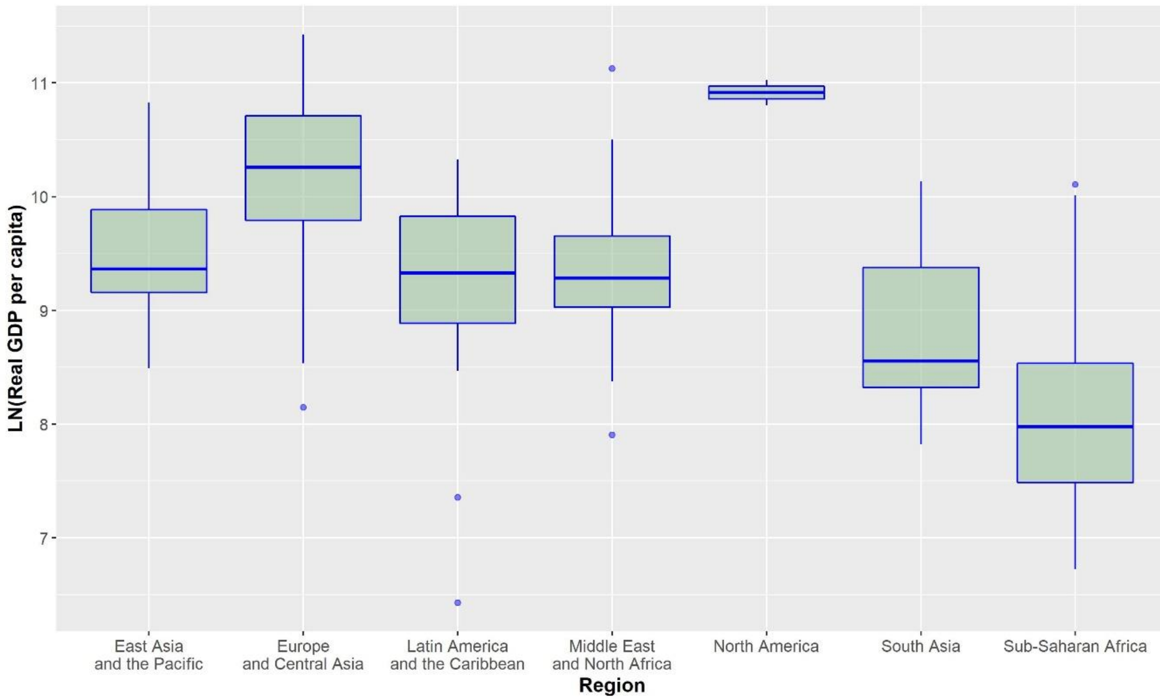
Note: Earliest data refer to the first data points available from 1990 to 2019 by country, while the latest data refer to the most recent data points available.
Source: Authors, based on Feenstra et al. (2015), World Economic Outlook October 2021 database (2021), and the World Bank World Development Indicators database (2021).

Incidentally, compared to the baseline, the average real per capita income has risen across the regional groups. The increase is also marked in all regions except in Sub-Saharan Africa (Figures 2a and 2b). Recent data (Figure 2a) show that average real income per capita is highest in North America and Europe and Central Asia while South Asia and Sub-Saharan Africa trail all the other regions.

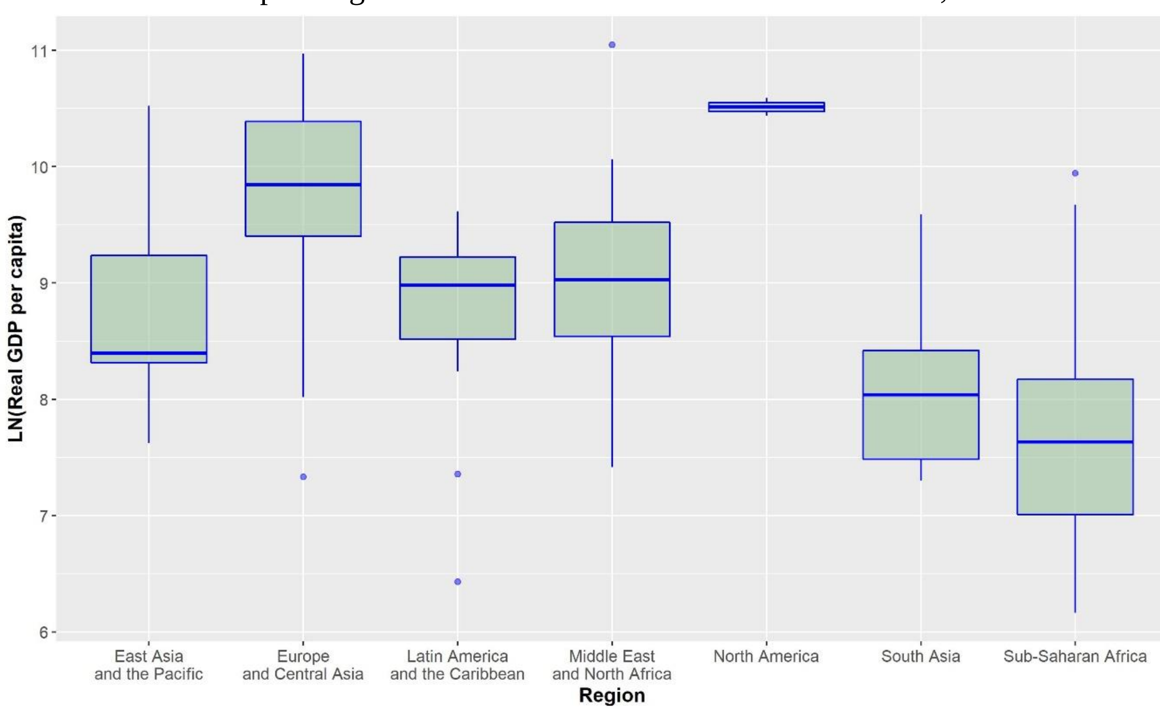
Moreover, compared to the baseline, the increase in human capital, investment-to-GDP ratio, and employment rate can be generally observed in all regions (see: Table A1 in the Appendix). Exceptions are the investment-to-GDP ratio in Middle East and North Africa, which has marginally slid, and the average employment rate in Sub-Saharan African economies, which has also minimally declined (i.e. average unemployment rate increased slightly).

Figure 2. Real GDP per capita by Region

a. Data corresponding to the latest available Gini coefficient data, 1990-2019



b. Data corresponding to the earliest available Gini coefficient data, 1990-2019



LN = natural logarithm.
Notes: The GDP per capita data included in the calculation correspond to the period of the available Gini coefficient data for every country. For instance, if the latest Gini coefficient data of a country is for 2015, the latest GDP per capita included in the calculation for the same country is also for 2015.
Source: Authors, based on Feenstra et al. (2015), World Economic Outlook October 2021 database (2021), and the World Bank World Development Indicators database (2021).

Methodology for the Empirical Exercise

In examining the relationship between economic development and income inequality in a more in-depth manner, the study applies a semi-parametric approach that takes into account the possible nonlinearities (Kosorok 2009). The general model is specified by Equation 4.

$$G_{i,t} = \boldsymbol{x}_{i,t}^{'}\beta + m\left(y_{i,t}\right) + \epsilon_{i,t} \tag{4}$$

In the model, G is proxied by the Gini coefficient as a measure of income inequality; $\boldsymbol{x}$ is a vector of control variables; and $y_{i,t}$ is a measure of per capita income in real terms, proxied by GDP per capita, for each country $i = 1,2, \dots n$ and time $t = 1,2, \dots, T$ years. Of notable importance in the specification is the function $m\left(y_{i,t}\right)$. It is a flexible and smooth function that is partially specified by the observable variable $y_{i,t}$. The function captures both linear and non-linear relationships, which renders it more advantageous than the polynomial type regression models used in a number of related literature.

Following Hastie, Tibshirani, and Friedman (2001), the functional form $m\left(y_{i,t}\right)$ under the context of regression splines implies that $m\left(y_{i,t}\right)$ is a piecewise-polynomial of order M and contains at most M–2 continuous derivatives. The positive disjoint regions under the knots $\xi_j = 1,2, \dots, K$ in a spline type model set-up are defined by Equations 5 and 6.

$$m_j(y) = y^{j-1}, j = 1,2, \dots, M \tag{5}$$

$$m_{M+\tau}(y) = \left(y - \xi_\tau\right)_+^{M-1}, \tau = 1, \dots, K \tag{6}$$

The order of the polynomial is M with M–1 degrees and the number of knots ξ is defined up to K knots selected by an error criterion. The possible $m_j(y)$ polynomials are obtained by dividing the domain of y in "contiguous intervals" where each interval has a representative M polynomial form (see: Hastie, Tibshirani and Friedmann (2017) section 5.2 "Piecewise Polynomials and Splines" for further explanation). Considering the potential bias that could arise from the general linear model specified by Equation 4, a selection procedure is done to identify the number of degrees of the polynomial in this B-Spline as basis for the regression. In order to reduce the likelihood of bias, the Akaike error criterion is used as suggested by Kleiber and Zeileis (2011). In this study, the data are segmented by country group to appropriately capture the heterogeneities in each geographic cluster.

The methodology allows for nonlinearities at the cost of the population point-estimates at least for the $m_j(y)$ in the process of capturing the dynamics of the relationship of income inequality and economic activity. By using the fitted values after estimating Equation 4 restricted to the economic production variable $y_{i,t}$, the behavior of inequality against real economic production can be pinned down. The regression outputs are also presented with the selected covariates for each region using the general linear model in Equation 4.

Some specification issues

It is worth noting that Equation 4 specifies the variables in levels and not in growth terms. This approach follows the previous literature of semi-parametric and non-parametric studies of Zhou and Li (2011) and Li and Zhou (2014) related to income inequality and real per capita income, where the variables in levels are analyzed. In contrast, the approaches of Barro (2000); and Cunha Neves and Tavares Silva (2013) included dynamic terms in the regressions.

Aside from the original work of Kuznets (1956) and the previously mentioned literature that used non-parametric methods, more specification issues are discussed by Banarjee and Duflo (2003). One of which is the challenge of extracting long-term relationships. For instance, estimating Equation 4 is likely to show long-run relationships. However, if the panel dataset used has large number of countries and relatively short time period, cointegration tests would be weaker in terms of the asymptotic performance. On the other hand, estimating Equation 4 with partial adjustment specifications or with dynamic terms in the specification may neglect the existence of the long-run relationships even when they may actually exist.

The aforementioned difficulty in identifying long run relationships can be overcome by putting together consistent dataset over a longer period. And in order to compare results with the short-run specification, Equation 7 is also estimated.

$$\Delta \ln G_{i,t} = \boldsymbol{x}_{i,t}^{'}\beta + m\left(\Delta \ln y_{i,t}\right) + \epsilon_{i,t} \tag{7}$$

In Equation 7, the growth of income inequality and real per capita income are measured by the log-differences of these variables given by the term, $\Delta \ln G_{i,t}$ and $\Delta \ln y_{i,t}$, respectively[7]. Estimating Equations 4 and 7 allows contrasting of the results and identification of the sources of variations. At the same time, the analysis of the long run dynamics based on Equation 4 as in the original study of Kuznets (1956) benefits from the short-term results based on Equation 7 to provide context.

The segmentation by region is done in order to reduce the potential aggregation bias commonly encountered in this kind of analysis as discussed by Piketty and Saez (2003) and Piketty (2014), who examined the asymmetric forms of the Kuznets curve at different aggregation levels. The world aggregated effect is not necessarily undesirable. For instance, Barro and Sala-i-Martin (2004) show that the reduction in poverty has been accompanied by the rise in inequality in a number of countries. However, the analysis conveys heterogeneous patterns across countries and country groups.

[7] The error correction forms under this panel structure $N > T$ represents a problem in terms of the possible detection of cointegration, therefore we remain cautious about this procedure since it can result in inaccurate depiction of the actual relationship between inequality and growth.

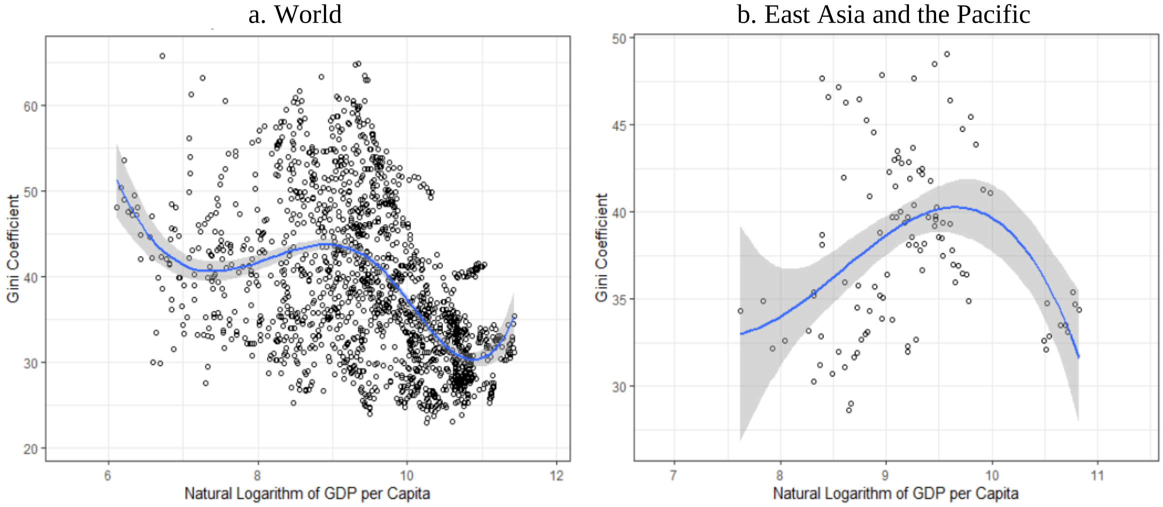
Results

The pairwise relationship of the Gini coefficient across countries and regions using the regression splines are presented in Figure 3. The panel of charts indicate heterogeneous patterns across regions. Overall, evidence of the original inverted U-shape relationship described by Kuznets is not clearly established in all regional groups and the world as a whole. It is only visible in East Asia and the Pacific (Figure 3b) and Latin America and the Caribbean (Figure 3c).

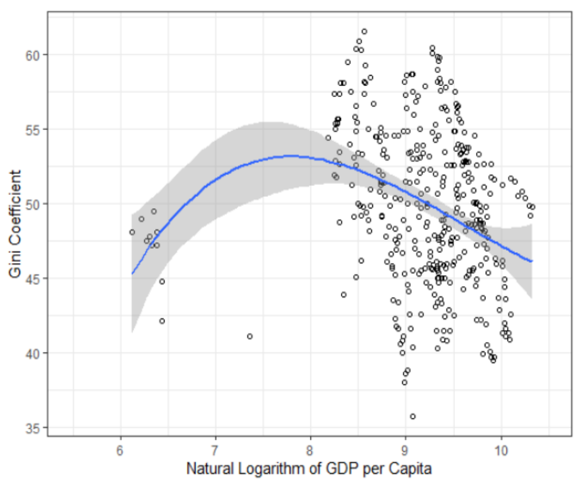
The results based on the global sample (Figure 3a) indicate a heterogeneous pattern. In North America (Figure 3d), inequality tends to increase with the rise in the levels of economic production per capita. In Sub-Saharan Africa (Figure 3e) and South Asia (Figure 3h), an N-shaped relationship can be observed. The inverted U-shape relationship of inequality and per capita economic production is only visible in one segment of the regional datasets. The charts for Middle East and North Africa (Figure 3g) indicate a quasi-linear inverse relationship between inequality and economic production. Meanwhile, the data for Europe and Central Asia (Figure 3f) seem to be the most heterogeneous—a result that could be due to the number of countries in the dataset that belong to this region. Serious fluctuations can be gleaned from the charts with a lasting pattern of simultaneous increases in inequality and per capita economic production. The varying relationships of economic production per capita and the Gini coefficient by region implies that different local and structural factors play critical roles across regional groups.

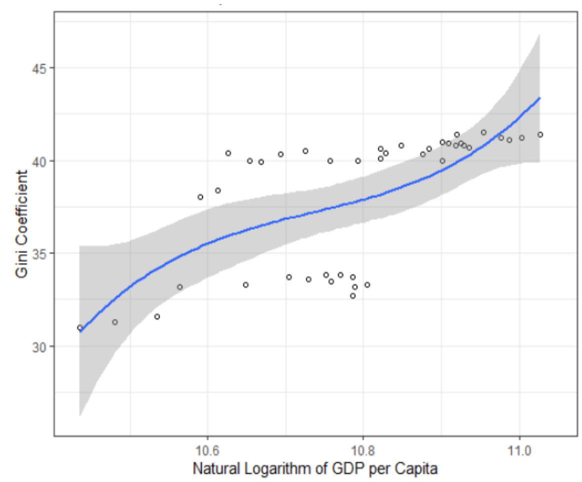
Figure 3. Regression Spline Outputs by Region, Gini and GDP per capita

a. World b. East Asia and the Pacific

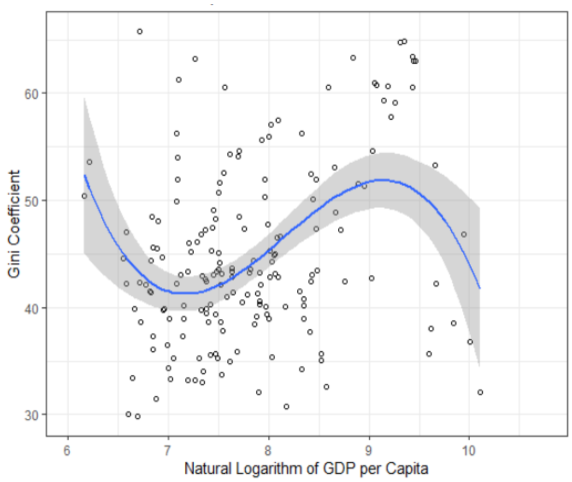


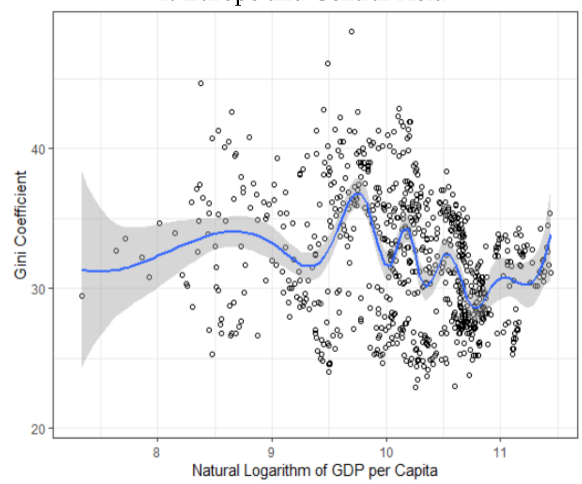
c. Latin America and the Caribbean d. North America



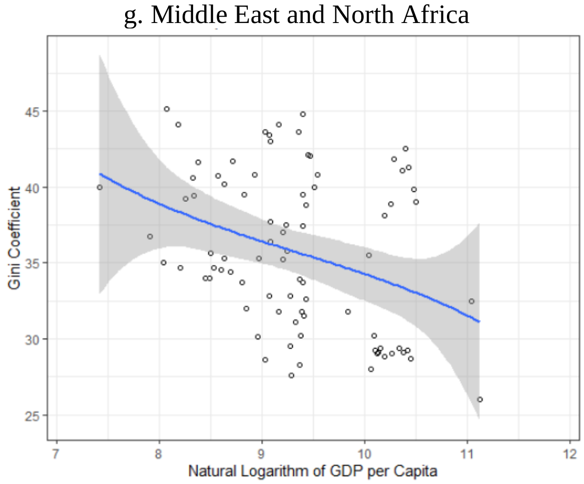


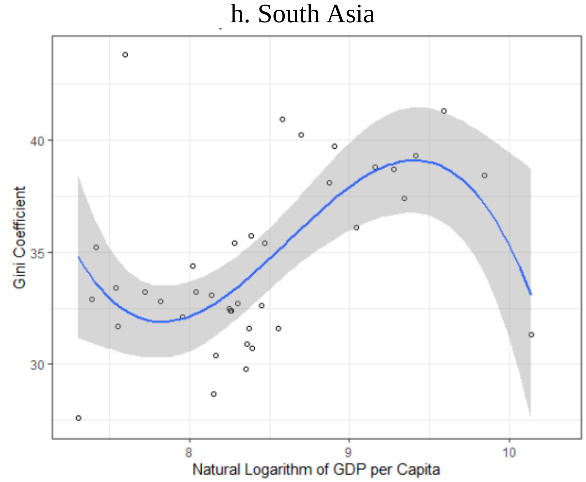
e. Sub-Saharan Africa f. Europe and Central Asia





g. Middle East and North Africa h. South Asia





Source: Authors.

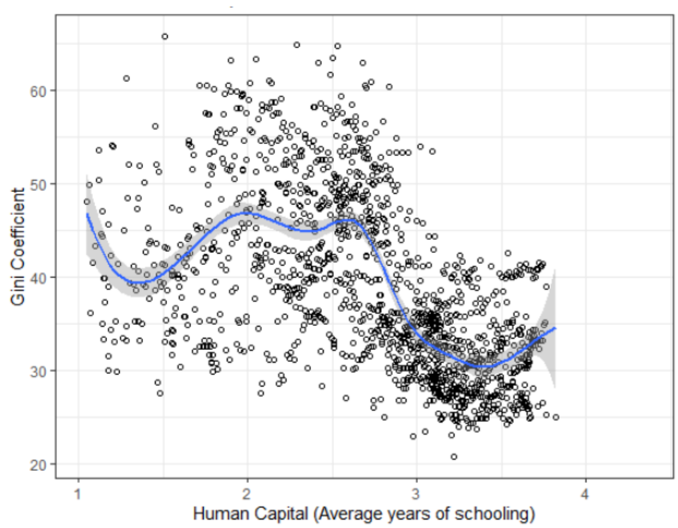
The regressions abide by the Akaike error criterion to reduce the bias. The selection of degrees of freedom are likewise set differently across regions. The regression results in levels, which are in Appendix A, indicate that piecewise polynomials are jointly statistically significant at a 5% level of significance. However, in the spline regression of inequality and per capita economic output growth, all the models according to the F-tests are not significant. The short-run dynamics has notably changed drastically to higher orders of the piecewise polynomials. Since growth rates are used, this implies higher frequency dynamics and effectively higher uncertainty about the evolution of per capita income growth and growth in inequality, further research is recommended.

Overall, the results indicate that the proposed inverted U-shape relationship between income inequality measured by the Gini Coefficient and the levels of economic production given by the real GDP per capita is not well established across the regions based on the empirical results. This finding is in line with Piketty (2003), who emphasize the role of the aggregate trend behaviors and possible misleading empirical facts about the evolution of inequality. Given the extensive vector of control variables used in the estimation, this finding also raises questions on the drivers of the variations and whether these driving factors are changing in terms of importance over time.

Moreover, the global pattern captured by the traditional square or cubic parametric regressions as discussed by Barro (2000), Barro (2008), and Barnerjee and Duflo (2003) is heavily dependent on the functional form of the specification. The spline type estimation reveals heterogeneities across the regions and the heterogeneity is particularly evident in areas with the highest levels of real income per capita. Incidentally, this phenomenon tends to be neglected by researchers when they perform their parametric analysis through regressions using nonlinear terms. Notably, the direction of influence is potentially from income inequality towards per capita income growth and not vice versa at certain stage of the country's economic development.

As anticipated, the empirical results of the semi-parametric regressions indicate a somewhat negative correlation between income inequality and human capital accumulation that is proxied by the average years of schooling (Figure 4). This happens to be a common control used in the empirical literature on this subject matter (e.g., Delbianco, Dabús, and Caraballo 2014, Levi and Renelt 1992, Barro 2000). However, as can be observed from Figure 4, the negative relationship does not appear to be robust across levels of human capital. It would indeed be interesting to look into the country contexts in examining this relationship further.

Figure 4. Scatterplot Gini Coefficient and Human Capital Accumulation



Source: Authors.

This effect of human capital on income inequality has been the subject of a few studies. One of which is Aiyar and Ebeke (2019), who posit that "unequal access to education, unequal access to labor markets and unequal access to finance, separately or in various combinations, could amplify the negative impact that a worsening of the income distribution has on growth." This finding is hardly surprising considering the theoretical constructions of the growth models and innovation such as the Solow residual in the Solow Growth model. What it is noticeable is the strong correlation that exists between inequality and real per capita income growth in the high-income economies, which can be explored further in the future. It could be the case that educational levels may have an important and unexplored correlation with the income inequality.

Conclusion

Based on the empirical analysis using semi-parametric regression, it is shown that the classical hypothesis of Kuznets is not well-established in all regions. This reflects the heterogeneities present across economies and regional groups.

The regions where the Kuznets hypothesis can be observed are East Asia and the Pacific and Latin America and the Caribbean. In Sub-Saharan Africa, South Asia, Europe and Central Asia, the data suggest an N-shape relationship between inequality and per capita income. In the Middle East and North Africa the data interestingly present a negative correlation between income inequality and economic development. Meanwhile, in North America inequality appears to increase as the real income per capita increases. In the latter three cases, it would be informative to carry out further research at the country-level to examine the peculiarities and the drivers.

The mechanism in which the inequality correlates with income levels remains the focus of many studies. The same can be said of the income redistribution dynamics within economies. This empirical exercise emphasizes that the heterogeneity across economies and regions must be taken in account in analyzing this particular issue. There are also ample merits in investigating further the relationship between human capital accumulation and inequality in different country contexts.

Moving forward, it would be interesting to conduct a similar exercise using different measures of inequality rather than the Gini coefficient, although, it could be empirically laborious. Doing so will provide comparable results and establish if the findings presented in this research are robust.

References

Acemoglu, D., Ticchi, D., and Vindigni, A. (2011), Emergence and Persistence of Inefficient States, *Journal of European Economic Association*, 9(2), pp. 177–208, https://doi.org/10.1111/j.1542-4774.2010.01008.x.

Agnello, L., Castro, V., Jalles, J. T., and Sousa, R. M. (2017), Income Inequality, Fiscal Stimuli and Political (In)stability, *International Tax and Public Finance*, 24, pp. 484-511, https://doi.org/10.1007/s10797-016-9428-x.

Aiyar, S. and Ebeke, C. (2019), Inequality of Opportunity, Inequality of Income and Economic Growth, *IMF Working Paper*, WP/19/34, Washington DC: International Monetary Fund. https://www.imf.org/-/media/Files/Publications/WP/2019/WPIEA2019034.ashx.

Alesina, A., and Rodrik, D. (1994), Distributive Politics and Economic Growth, *The Quarterly Journal of Economics*, 109(2), pp. 465-490, https://doi.org/10.2307/2118470.

Ahluwalia, M.S. (1976), Inequality, Poverty and Development, *Journal of Development Economics*, 3(4), pp. 307–342, https://doi.org/10.1016/0304-3878(76)90027-4.

Alvaredo, F., Chancel, L., Piketty, T., Saez, E., and Zucman, G. (2017), The Elephant Curve of Global Inequality and Growth, *WID.world Working Paper Series*, No. 2017/20, World Inequality Database, https://wid.world/document/elephant-curve-global-inequality-growth-wid-world-working-paper-2017-20/

Banerjee, A.V. and Duflo, E. (2003), Inequality and Growth: What Can the Data Say?, *Journal of Economic Growth*, 8(3), pp. 267-299.

Barro, R. J. (2000). Inequality and Growth in a Panel of Countries. Journal of Economic Growth, 5(1), 5–32. http://www.jstor.org/stable/40216021.

Barro, R.J. (2008), Inequality and Growth Revisited, *ADB Working Paper Series on Regional Economic Integration*, No. 11, Manila: Asian Development Bank, https://www.adb.org/sites/default/files/publication/28468/wp11-inequality-growth-revisited.pdf.

Barro, R.J., and Xavier Sala-i-Martin (1992), Convergence, *Journal of Political Economy*, 2(100), April, pp. 223–251. https://doi.org/10.1086/261816.

Barro, R. and Sala-i-Martin, X. (2004), *Economic Growth*, 2nd ed, Boston: Massachusetts Institute of Technology.

Baymul, C. and Sen, K. (2020), Was Kuznets Right? New Evidence on the Relationship between Structural Transformation and Inequality, *The Journal of Development Studies*, 56(9), pp. 1643-1662, https://doi.org/10.1080/00220388.2019.1702161.

Berg, A., Ostry, D.J., Tsangarides, C.G., and Yakhshilikov, Y. (2018), Redistribution, Inequality and Growth: New Evidence. *Journal of Economic Growth*, 23, pp. 259–305, https://doi.org/10.1007/s10887-017-9150-2.

Climent, A.C. (2010), Inequality and Growth in Advanced Economies: An Empirical Investigation. *Journal of Economic Inequality*, 8, pp. 293–321. https://doi.org/10.1007/s10888-010-9133-4.

Cunha Neves, P. and Tavares Silva, S. M. (2014), Inequality and Growth: Uncovering the Main Conclusions from the Empirics, *The Journal of Development Studies*, 50(1), pp. 1-21, https://doi.org/10.1080/00220388.2013.841885.

Castelló, A. and Doménech, R. (2002), Human Capital Inequality and Economic Growth: Some New Evidence, *The Economic Journal*, 112(478), pp. C187–C200, https://doi.org/10.1111/1468-0297.00024.

de la Croix, D., and Doepke, M. (2003). Inequality and Growth: Why Differential Fertility Matters, *The American Economic Review*, 93(4), pp. 1091-1113, https://doi.org/10.1257/000282803769206214.

Delbianco, F, Dabús, C, and Caraballo, M. Á. (2014), Income Inequality and Economic Growth: New Evidence from Latin America, *Cuadernos de Economía*, 33(63), pp. 381-398. https://doi.org/10.15446/cuad.econ.v33n63.45338.

Easterly, W. (2007), Inequality Does Cause Underdevelopment: Insights from a New Instrument. *Journal of Development Economics*, 84(2), pp. 755-776, https://doi.org/10.1016/j.jdeveco.2006.11.002.

Feenstra, R.C., Inklaar, R., and Timmer, M.P. (2015), The Next Generation of the Penn World Table, *American Economic Review*, 105(10), pp. 3150-3182, https://doi.org/10.1257/aer.20130954.

Foellmi, R., and Zweimüller, J. (2006). Income Distribution and Demand-Induced Innovations, *Review of Economic Studies*, 73(4), pp. 941–960, http://www.jstor.org/stable/4123254.

Forbes, K.J. (2000). A Reassessment of the Relationship between Inequality and Growth. *The American Economic Review*, 90(4), pp. 869-887, https://doi.org/10.1257/aer.90.4.869.

Galor, O. and Zeira, J. (1993), Income Distribution and Macroeconomics, *The Review of Economic Studies*. 60(1), pp. 35–52, https://doi.org/10.2307/2297811.

Gupta, G.S., and Singh, R.D. (1984), Income Inequality Across Nations Over Time – How much and why, *Southern Economic Journal*, 51(1), pp. 250–277, https://doi.org/10.2307/1058336.

Halter , D., Oechslin, M., and Zweimüller, J. (2014), Inequality and Growth: The Neglected Time Dimension. *Journal of Economic Growth*, 19(1), pp. 81-104, https://www.jstor.org/stable/44113419.

Hastie T, Tibshirani R, Friedman J (2001), The Elements of Statistical Learning. New York: Springer-Verlag.

International Monetary Fund (2021), IMF Data, https://www.imf.org/en/Data.

Kaldor, N. (1966), *Causes of the Slow Rate of Economic Growth of the United Kingdom. An Inaugural Lecture*. London: Cambridge University Press and Louvain: Recherches Economiques De Louvain/Louvain Economic Review.

Kleiber, C. and Zeileis, A. (2011). *Applied Econometrics with R*, Vienna: Springer.

Kosorok M. R. (2009), What's So Special About Semiparametric Methods?, Sankhya. Series B. [Methodological.], 71-A(2), pp. 331–353, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2903063/

Kuznets, S. (1955), Economic Growth and Income Inequality, *American Economic Review*, 58(1), pp. 1-28, https://assets.aeaweb.org/asset-server/files/9438.pdf

Kyosuke , K. and Kurosaki T. (2011), Dynamics of Growth, Poverty and Inequality: A Panel Analysis of Regional Data from Thailand and the Philippines. *Asian Economic Journal*, 25(1), pp. 3-33, https://doi.org/10.1111/j.1467-8381.2011.02046.x.

Levine, R. and Renelt, D. (1992), A Sensitivity Analysis of Cross-country Growth Regressions, *American Economic Review*, 82(4), 942-963, https://www.jstor.org/stable/2117352

Panizza, P. (2002), Income, Inequality and Economic Growth: Evidence from American Data, *Journal of Economic Growth*, 7(1), pp. 25–41, https://doi.org/10.1023/A:1013414509803.

Perotti, R. (1993). Political Equilibrium, Income Distribution, and Growth. *Oxford Journals*, 60(4), pp. 755-776, https://doi.org/10.2307/2298098.

Perotti, R. (1996), Growth, Income Distribution and Democracy, *Journal of Economic Growth*, 1(2), pp. 149-187, https://www.jstor.org/stable/40215914.

Piketty, T. (1997), The Dynamics of the Wealth Distribution and the Interest rate with Credit Rationing, *The Review of Economic Studies*, 64(2), pp. 173–189, https://doi.org/10.2307/2971708.

Piketty, T. (2003), Income Inequality in France, 1901–1998, *Journal of Political Economy, 2003*, 111(5), http://www.piketty.pse.ens.fr/fichiers/public/Piketty2003b.pdf.

Piketty, T. (2014), *Capital in the Twenty-First Century*, Cambridge, MA and London, England: Harvard University Press, https://doi.org/10.4159/9780674369542.

Piketty, T. and Saez, E. (2003), Income Inequality in the United States, 1913–1998, The *Quarterly Journal of Economics*, 118(1), pp 1–41, https://doi.org/10.1162/00335530360535135.

Rosenzweig, M.R. and Binswanger, H.P. (1993). Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments, *The Economic Journal*, 103(416), pp. 56-78, https://doi.org/10.2307/2234337.

Robinson, S. (1976), A Note on the U-hypothesis Relating Inequality and Economic Development, *American Economic Review*, 66 (3), pp. 437–440.

Rosenzweig, M.R., and Binswanger, H.P. (1993). Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments, *The Economic Journal*, 3(416), pp. 56-7, https://doi.org/10.2307/2234337.

World Bank (2021a). World Development Indicators, https://data.worldbank.org/indicator.

World Bank (2021b). World Bank Countries and Economies, https://data.worldbank.org/country.

Xie, Y. and Zhou, X. (2014), Income Inequality in Today's China, *Proceedings of the National Academy of Sciences of the United States of America*, 111(19), pp. 6928-6933. https://doi.org/10.1073/pnas.1403158111.

Zhou, X. and Li, K. (2011), Inequality and Development: Evidence from Semiparametric Estimation with Panel Data, *Economics Letters*, 113(3), pp. 203-207. https://doi.org/10.1016/j.econlet.2011.07.013.

Appendix

Table A1. Summary Statistics by Region

a. East Asia and the Pacific

	Number of observations	Mean	Standard Deviation	Minimum	Maximum
Earliest Data					
Gini coefficient	11	37.75	6.299	31.2	47.7
Real GDP per capita	11	11,113	12,864	2,047	37,184
Human capital, average years of schooling	11	2.354	0.63	1.552	3.474
Total investment, % of GDP	10	29.46	9.373	15.68	42.5
Unemployment rate, %	7	5.548	3.287	2.5	11.18
Latest Data					
Gini coefficient	11	36.47	3.675	30.7	42.3
Real GDP per capita	11	17,846	14,133	4,885	50,478
Human capital, average years of schooling	11	2.712	0.553	1.817	3.529
Total investment, % of GDP	10	30.53	6.679	23.73	42.65
Unemployment rate, %	9	4.768	1.995	1	7.8

b. Europe and Central Asia

Earliest Data	Number of observations	Mean	Standard Deviation	Minimum	Maximum
Gini coefficient	47	33.27	6.178	20.7	48.4
Real GDP per capita	46	22,053	14,193	1,533	58,359
Human capital, average years of schooling	40	2.946	0.331	1.844	3.476
Total investment, % of GDP	44	22.46	7.495	-7.957	39.05
Unemployment rate, %	40	9.749	6.059	1.38	32.17
Latest Data					
Gini coefficient	47	31.8	4.53	24.6	41.9
Real GDP per capita	47	33,190	20,400	3,453	91,850
Human capital, average years of schooling	40	3.32	0.304	2.495	3.821
Total investment, % of GDP	45	23.08	4.961	13.34	41.53
Unemployment rate, %	43	7.708	4.823	2.243	20.73

c. Latin America and the Caribbean

Earliest Data	Number of observations	Mean	Standard Deviation	Minimum	Maximum
Gini coefficient	25	50.2	6.943	40.3	60.5
Real GDP per capita	24	7,821	3,639	620.7	14,995
Human capital, average years of schooling	23	2.154	0.346	1.638	2.697
Total investment, % of GDP	22	19.58	4.22	11.9	27.45
Unemployment rate, %	17	8.748	3.972	2.83	15.97
Latest Data					
Gini coefficient	25	46.09	4.803	38.8	57.9
Real GDP per capita	25	12,199	7,297	619.3	30,610
Human capital, average years of schooling	23	2.62	0.376	1.638	3.108
Total investment, % of GDP	23	21.43	5.823	14.64	39.29
Unemployment rate, %	20	8.162	3.09	3.334	12.8

d. Middle East and North Africa

Earliest Data	Number of observations	Mean	Standard Deviation	Minimum	Maximum
Gini coefficient	15	35.62	4.82	28	43.6
Real GDP per capita	15	13,151	15,344	1,666	62,879

	Number of observations	Mean	Standard Deviation	Minimum	Maximum
Human capital, average years of schooling	12	1.986	0.653	1.14	3.178
Total investment, % of GDP	13	25.35	6.566	9.103	34.26
Unemployment rate, %	8	16.11	7.928	6.817	28.1
Latest Data					
Gini coefficient	15	34.11	5.088	26	42
Real GDP per capita	15	17,198	17,105	2,709	68,036
Human capital, average years of schooling	12	2.493	0.591	1.576	3.766
Total investment, % of GDP	13	24.81	8.636	7.832	38.06
Unemployment rate, %	10	11.77	6.382	3.658	26.9

e. North America

Earliest Data	Number of observations	Mean	Standard Deviation	Minimum	Maximum
Gini coefficient	2	34.5	4.95	31	38
Real GDP per capita	2	36,915	4,033	34,063	39,767
Human capital, average years of schooling	2	3.382	0.0997	3.311	3.452
Total investment, % of GDP	2	19.73	0.544	19.34	20.11
Unemployment rate, %	2	8.584	2.452	6.85	10.32
Latest Data					
Gini coefficient	2	37.35	5.728	33.3	41.4
Real GDP per capita	2	55,386	8,719	49,220	61,551
Human capital, average years of schooling	2	3.725	0.0267	3.706	3.744
Total investment, % of GDP	2	22.35	1.701	21.15	23.55
Unemployment rate, %	2	5.137	1.761	3.892	6.383

f. South Asia

Earliest Data	Number of observations	Mean	Standard Deviation	Minimum	Maximum
Gini coefficient	7	34.61	4.986	27.6	41.3
Real GDP per capita	7	4,563	4,635	1,485	14,596
Human capital, average years of schooling	6	1.626	0.439	1.349	2.505
Total investment, % of GDP	7	27.25	15.08	17.48	60.67
Unemployment rate, %	3	6.942	7.783	1.8	15.9
Latest Data					
Gini coefficient	7	34.36	3.128	31.3	39.3
Real GDP per capita	7	9,258	7,983	2,493	25,191

Human capital, average years of schooling	6	2.076	0.453	1.559	2.867
Total investment, % of GDP	7	32.52	12.8	17.34	54.33
Unemployment rate, %	3	4.363	1.206	3.138	5.55

g. Sub-Saharan Africa

	Number of observations	Mean	Standard Deviation	Minimum	Maximum
Earliest Data					
Gini coefficient	44	47.31	9.198	32.1	65.8
Real GDP per capita	44	3,556	4,301	475.4	20,836
Human capital, average years of schooling	37	1.575	0.384	1.045	2.416
Total investment, % of GDP	40	22.66	14.39	5.375	61.5
Unemployment rate, %	6	14.14	7.136	3.325	22.16
Latest Data					
Gini coefficient	44	43.55	8.13	32.1	63
Real GDP per capita	44	4,875	5,438	833.1	24,512
Human capital, average years of schooling	37	1.864	0.475	1.193	2.834
Total investment, % of GDP	41	24.54	9.615	7.261	52.85
Unemployment rate, %	7	14.78	8.138	3	25.1

GDP = gross domestic product.

Notes: Earliest data refer to the first data points available from 1990 to 2019 by country, while the latest data refer to the most recent data points available. The data for GDP per capita, human capital, total investment, and unemployment rate included in the calculation correspond to the period of the available Gini coefficient data for every country. For instance, if the latest Gini coefficient data of a country is for 2015, the latest GDP per capita included in the calculation for the same country is also for 2015. There are periods when Gini coefficient data are available, but not the other variables.

Source: Authors, based on Feenstra et al. (2015), World Economic Outlook October 2021 database (2021), and the World Bank World Development Indicators database (2021).

Table (), Splines Regression Output in Levels

	Dependent variable:							
	Gini Coefficient							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Piecewise Polynomials (Results Omitted)	**	**	**	**	**	**	**	**
hc	-8.078***	3.313	-2.978***	-37.119***	-6.690	-5.307***	1.314	10.130**
	(0.534)	(3.192)	(0.977)	(10.849)	(14.463)	(0.431)	(2.638)	(3.841)
invest	-0.215***	-0.048	0.262***	-0.910***	2.076	-0.153***	0.387***	-0.046
	(0.032)	(0.057)	(0.048)	(0.275)	(1.082)	(0.030)	(0.083)	(0.149)
unemp_rate	-0.120***	-0.395	0.546***	-0.491	2.566***	0.048	0.502***	-0.638

	(0.045)	(0.271)	(0.077)	(0.383)	(0.514)	(0.032)	(0.187)	(0.399)
Constant	108.261***	20.393	26.973***	176.009***	-3,834.189	48.977***	-10.424	33.073
	(7.031)	(27.090)	(6.273)	(28.695)	(1,987.893)	(3.924)	(30.171)	(28.265)
Observations	1,266	90	367	42	11	675	62	19
R^2	0.458	0.177	0.259	0.863	0.987	0.412	0.480	0.840
Adjusted R^2	0.455	0.118	0.247	0.840	0.968	0.397	0.423	0.760
F Statistic	151.760*** (df = 7; 1258)	2.983** (df = 6; 83)	20.984*** (df = 6; 360)	36.864*** (df = 6; 35)	51.021*** (df = 6; 4)	28.766*** (df = 16; 658)	8.446*** (df = 6; 55)	10.519*** (df = 6; 12)

Note: Statistical Significance *p<0.1 **p<0.5 ***p<0.01. Piecewise polynomials coefficients are omitted. Controls Included Human capital accumulation based on the average years of schooling, the percentage of investment of the economy compared to the GDP, and the unemployment rates. Control's selection is a mixture on the mention of the Perotti (1996) and Barro (2000) specifications. Source: Own Elaboration using R.

Column Models represent:

- (1) = World
- (2) = East Asia and pacific
- (3) = Latin America and the Caribbean
- (4) = North America
- (5) = Sub-Saharan Africa
- (6) = Europe and Central Asia
- (7) = Middle East and North Africa
- (8) = South Asia

Table () Splines Regression Output in Log Differences.

	Dependent variable:					
	Log Difference (Gini Coefficient)					
	(1)	(2)	(3)	(4)	(5)	(6)
Piecewise Polynomials (Results Omitted)	…	…	…	…	…	…
gr_hc	0.002	0.007	-0.005	-0.148*	-0.002	-0.041
	(0.003)	(0.035)	(0.007)	(0.076)	(0.006)	(0.046)
gr_inv	0.007	-0.099*	-0.001	0.176	0.020	0.011
	(0.013)	(0.051)	(0.020)	(0.344)	(0.019)	(0.052)
Gr_unemp rate	0.0002	0.005*	0.001*	-0.0002	0.00002	-0.002

	(0.0003)	(0.002)	(0.001)	(0.004)	(0.0004)	(0.006)
Constant	0.063	0.228	6.397	3.618	0.415	-0.136
	(11.603)	(0.868)	(8.439)	(2.942)	(2.113)	(4.143)
Observations	959	43	291	33	572	18
R^2	0.022	0.441	0.069	0.456	0.030	0.324
Adjusted R^2	0.003	0.060	0.008	-0.244	0.002	-0.044
F Statistic	1.159 (df = 18; 940)	1.158 (df = 17; 25)	1.127 (df = 18; 272)	0.651 (df = 18; 14)	1.063 (df = 16; 555)	0.880 (df = 6; 11)

Note: Statistical Significance *p<0.1 **p<0.5 ***p<0.01. Piecewise polynomials coefficients are omitted, and also are not significant. Controls in growth rates included Human capital accumulation based on the average years of schooling, the growth in percentage of investment of the economy, and the growth of unemployment rates. Source: Own Elaboration using R.

Column Models represent:

(1) = World
(2) = East Asia and pacific
(3) = Latin America and the Caribbean
(4) = North America
(5) = Sub-Saharan Africa
(6) = Europe and Central Asia
(7) = Middle East and North Africa
(8) = South Asia