Issues in Political Economy, Vol 30(2), 2021, 29-45

Convergence and economic growth: does income inequality matter?

Kenneth Giraldo, Furman University
Yuhan You, Furman University

Convergence and steady states are important theoretical and empirical concepts in the sphere of economic growth. Convergence theory suggests that over time as countries develop, their wealth grows. In the process, poorer economies will experience some catch-up effect with the wealthier economies measured by their level per-capita GDP. Steady state economic growth implies that in the long-run, given savings and population growth rates, a country will no longer be able to increase and achieve sustained economic growth rate without the introduction of other factors such as technological innovation. Economic growth first developed as a robust theoretical concept by Solow (1956), has since expanded to empirically test why some countries are poor while others are rich. These tests include factors such as technology, human capital, trade, and income inequality among other variables, as possible explanations in the growth literature.

The research idea was inspired from previous studies on the topic. For instance, Senhadji (2000) found some evidence for convergence, whereas Ben-David and David H. (1998) found no evidence of convergence among poorer countries. The inconsistent conclusions have thus left room for our further exploration on the topic of convergence. We recognize that countries vary in their momentum, pace, and sustainability of growth over time. We also recognize that growth rates of countries can be attributed to many different factors including physical or human capital accumulation, technological progress, the presence of inequality, trade factors and institutional factors. Hence, our research aims to empirically examine the concept of convergence, with a focus on the effect of one major explanatory variable, income inequality, in explaining a country's growth.

Studies related to income inequality and growth also provided insight for our analysis. In a recent review paper, Johnson and Papageougiou (2020) made a compelling case that starting with the formal origins of Solow (1956), the area of convergence "remains today a perennial research topic although only perhaps, under the new lamppost of global inequality." The simplest argument for the causal relationship flowing from inequality to growth comes from the political economy literature, which suggests that inequality leads to a redistribution of wealth and this in turn undermines growth. Likewise, in convergence testing, the effect of inequality on economic growth is also somewhat ambiguous. Persson & Tabellini (1994) found a large negative relation between inequality and growth, implying inequality is harmful for growth. In contrast, Li & Zou (1998) found a positive and significant relationship between income inequality and growth in their panel dataset containing 217 observations covering 46 countries. These two opposing conclusions among others such as, Banerjee and Duflo (2003), Barro (2000), Halter, Oechslin, & Zweimüller (2014), Rehme (2007), and Panizza (2002), triggered our interest in examining the relationship between income inequality and economic growth. In distinguishing our research from the previous ones, we take advantage of a carefully constructed panel data set for 182 countries with a focus on the period 2000-2017. The Madison-Penn World Table (PWT) dataset was recently updated in 2015. Similarly, data on income inequality (Gini coefficients) from the World Income Inequality Database (WIID) has also been updated and now covers observations up to the year 2018. This study is therefore a more contemporary 21st century data approach of this literature.

The rest of our research paper is developed as follows. Section II conducts a detailed review of previous studies carried out on convergence as well as studies that prominently include inequality as an independent variable. Here we will also look at other control variables in the growth convergence literature such as institutions and trade. Section III provides a discussion of the data, variables included, and our model specifications and methodology for running convergence and growth inequality tests. In section IV we show our main results and tables. Section V provides some concluding remarks.

II. Growth and Income Inequality Literature Review

Since our empirical work begins with some basic convergence tests, before we review some of the research on the effect of income inequality on growth, we cite some general research on convergence and divergence patterns across countries. In his survey article, Pritchett (1997) uses a country fixed effects model and concludes there was divergence 'big time' as he calls it, for less developed and non-industrialized countries for the period of 1870-1994. For developed countries, he identifies a strong pattern and considerable convergence in absolute income levels. He suggests that in theorizing about economic growth, researchers need to ask whether the country is a stable advanced economic leader, or an emerging but booming industrializing economy, or a semi-industrialized economy, distinct from others trying to escape a poverty trap into sustained growth.

Mankiw, Romer, & Weil (1992) build on and take seriously the basic implications of the Solow model. They show that growth of income is a function of the determinants of the ultimate steady state and the initial level of income. By regressing the change in the log of income per capita between 1960 and 1985 on the log of income per capita in 1960, with and without controlling for investment, growth of the working-age population, and school enrollment, the authors found that there was a strong tendency for convergence, or, for poor countries to grow faster than rich countries. Furthermore, they were among the first to augment the Solow model by including human capital as well as physical capital. They conclude that the augmented Solow model with savings, education, and population growth explain most of the international variation in per-capita income.

Given our main focus, which is the link between inequality and growth, there are some interesting studies that have examined this particular question. Barro (2000) concludes that inequality retarded growth in poor countries but promoted growth in wealthy countries. His empirical approach to capture variations in growth rates includes, per-capita GDP, Rule of Law Index, Democracy Index, Inflation Rate, Years of Schooling, and Growth Rate of the Terms of Trade as explanatory variables for three time periods between 1965 and 1995 for a broad set of countries. Separately he runs growth rates on a set of quintile–based income inequality measures. Barro (2000) concludes that the Kuznets curve hypothesis does emerge as a clear empirical regularity but technological considerations and other measures such as the stock of human capital, and investment to GDP ratio explain a large proportion of the variation in growth across countries over time.

Among the papers that look at the link between inequality and growth, Banerjee and Duflo (2003) provide a theoretical basis for this relationship but in addition look at why the effect of inequality and growth has given different and conflicting answers across various studies. They point to the fact that the choice of models used in previous studies is likely the main reason that some studies find a positive relation between growth and income inequality, while others show that it is negative. Their study conducts various empirical tests and demonstrate that imposing a linear specification on the relationship between inequality and growth is inaccurate and problematic at many levels including the simultaneity bias between the two variables. Banerjee and Duflo (2003), state, “OLS estimations of this equation are likely to be biased by a correlation between inequality and the error term. If this is indeed the real structure of the data, it is possible to solve some of these identification problems by exploiting the panel structure of the data.” They also suggest that this problem could be addressed by taking first differences in the growth equation and inequality estimation, or by introducing appropriate inequality lagged variables and quadratic terms. In their empirical work they find that inequality lagged 1-period leads to a negative relationship with growth rates. Following other studies including Barro (2000), they also find that there is a non-linear relationship between inequality and economic growth best captured by a quadratic term. The main reason why this paper is influential in our own research and in our model specification is because of the explicit warning they issue against the automatic specification of a linear models in empirical testing of growth rates and inequality.

Besides studies that have analyzed the issue at a global level, Birdsall, Ross & Sabot (1995) pay close attention to the region of East Asia for the period of 1965-90, including countries that were once recognized as "high performing Asian economies” like Hong Kong, Indonesia, Japan, the Republic of Korea, Malaysia, Singapore, Taiwan (China), and Thailand. In exploring the connection between rapid economic growth and low inequality for this set of countries over the 25-year period, they find that education (or primary and secondary school enrollment rates), as well as other elements such as a country’s development strategy, play a critical role in contributing to the overall positive economic environment. The regional focus of this paper is a reminder for us to not just run cross-country regressions for our data on 182 countries. Rather, to also carry out sub-group regressions based on geographical region and income levels, in an attempt to make comparisons that are more meaningful between regions.

Another interesting paper that contributed to our model specification is “Income Inequality: Does Inflation Matter?” by Aleš (2001). This study runs cross-country regressions with the effect of inflation and hyperinflation on inequality. They find that low levels of inflation or price stabilization decreased income inequality, and that this relationship between inequality and inflation is most pronounced in countries with hyperinflation and reductions in inflation appeared to have diminishing returns to improvements in inequality. Inclusion of some measure of government spending as a control variable is motivated by the paper “Mortality, Fertility, and Persistent Income Inequality” by Sarkar (2008), which describes a feedback loop due to the cost of healthcare services that perpetuate income inequality. Since parents care about the health and education of their children, yet health care workers are paid skilled level wages. The implication here is that healthcare across the world is subsequently costlier for the poor than the rich. Since healthcare is important in raising survivability, poor parents have higher child mortality relative to wealthy parents. Sarkar (2008) notes that this feedback loop is more impactful in areas with lower levels of PCGDP. Our model includes a health care expenditure variable as a percent of GDP.

Helpman et al (2010) offers some insight on how trade openness in a country could affect inequality. The paper looks at the relationship between trade openness and inequality and unemployment, and finds that the opening of trade can either increase or decrease unemployment depending on the current level of trade openness. Their model considers firms to be heterogeneous where not all firms participate in international trade resulting in some businesses enjoying increased revenue from trade participation. Their theoretical model shows trade openness increases inequality initially, but then decreases with more trade openness as more firms begin to participate. Due to the relationship trade openness has with inequality, as well as it's clear implications with economic growth, we have also added a trade openness variable to our model.

Finally, while not directly related to the growth literature, Acemoglu et al. (2001) looked at the impacts of colonization on the economies of various countries. The authors conclude that mortality rates significantly affected the type of institution established in European colonies, which led to long term palpable negative economic effects. They also conclude that there are benefits to adopting institutions that are more suited for economic growth.

Given our above literate review and following Barro (2000) and Banerjee and Duflo (2003) in particular, our data collection and model specification focused on the democracy index, inflation, trade openness, healthcare expenditures, in addition of course to growth rates, per-capita GDP and income inequality. The next section provides more details on our data sources and variables, and the motivation for our model specifications.

III. Data, Model Specification, & Methodology

Based on our literature review in Section I above, our study is divided into three main parts:

(1) A set of standard growth convergence tests.

(2) Conditional convergence tests with a focus and inclusion of inequality as the main variable of interest (conditional convergence test includes a measure of the per-capita GDP level, along with other independent variables).

(3) A non-linear specification with an introduction of a quadratic term on inequality closely following previous studies by Banerjee & Duflo (2003) and Barro (2000).

Accordingly, the first model specified below is typically referred to as the standard convergence test in the growth literature. A negative sign on $RGDP_{1k2000}$ indicates that the lower the initial level of income in a country, the higher its growth rate will be, matching up with what the concept of convergence suggests—over time, poorer countries tend to catch up with the richer ones in minimizing the gap in between.

$$(1) \qquad \widehat{GROWTH\_RATE} = \beta_0 + \beta_1 * RGDP_{1k2000} + \beta_2 * POP + \mu$$

The first set of regressions are for a cross section of 182 countries where the dependent variable is the average growth rate. The main independent variable of interest is the level Real GDP in 2000 and given the variation in population across this set of countries, we have included population as an additional control variable.

The convergence test is testing the null hypothesis that the coefficient on $rgdp_{1k2000}$, $\beta = 0$, versus the alternative that $\beta \neq 0$. Statistical significance implies that in our sample, if $\beta < 0$, those countries that were initially poorer in 2000, grow faster, than those countries that are or were initially richer in 2000, signaling convergence. Following the cross-country regression for Model (1), we use the same regression for the entire panel resulting in a larger sample size of 3,250 observations.

The second set of regressions are for a simple linear growth model to test the fit and significance of the inequality growth relationship. The full model is specified below.

$$(2) \qquad \widehat{GROWTH\_RATE} = \beta_0 + \beta_1 * GINI\_LAG + \beta_2 * RGDP1k\_LAG(or\ RGDP_{1k2000}) + \beta_3 * DEMOCRACY\_INDEX\_LAG + \beta_4 * TRADE\_INDEX\_LAG + \beta_5 * HEALTHSPEND\_LAG + \beta_6 * INFLATION\_LAG + \mu$$

The standard practice in estimating the relationship between growth and inequality is to assume a linear relationship. Instead of $RGDP_{1k2000}$ in testing for conditional convergence, Model (1), here we will lag real GDP one period, *RGDP1K_LAG*. Furthermore, we divide our real GDP variable by 1000 for better scaling and interpretation of the coefficients and since GDP is specified in the data in $ millions. We also lag the inequality (Gini coefficient). Given the Standard Kuznets curve hypothesis that growth influences inequality, specifying a lagged structure minimizes the problem of simultaneity bias. Once again, this is the standard preliminary procedure for example in Banerjee & Duflo (2003) and Barro (2000). Following the cross-country analysis, we run the model as a panel, with time and country fixed effects to account for unobservable factors that vary over time across all the countries or to account for unobservable factors that remain the same over time but vary across individual countries. The fixed effects regression is an estimation procedure in a panel data that allows us to control for time-invariant unobserved individual characteristics that can be correlated with the observed independent variables. In addition, if there are omitted variables in our model, and these variables are correlated with our included variables, then a fixed effects models may provide a means for controlling for omitted variable and simultaneity bias.

Our final model is similar in structure to Model (2) with one notable exception, the inclusion of inequality as a quadratic term.

$$(3) \quad \widehat{GROWTH\_RATE} = \beta_0 + \beta_1 * GINI\_LAG + \beta_2 * (GINI\_LAG)^2 + \beta_3 * RGDP1K\_LAG(or\ RGDP_{1k2000}) + \beta_4 * DEMOCRACY\_INDEX\_LAG + \beta_5 * TRADE\_INDEX\_LAG + \beta_6 * HEALTHSPEND\_LAG + \beta_7 * INFLATION\_LAG + \mu$$

Banerjee & Duflo (2003), state that OLS estimations of cross-country data are likely to be biased by correlation between inequality and the error term. In their paper, they closely examine and review past studies on the relationship between inequality and growth and conduct various theoretical tests to determine if a linear monotonic relationship between these two variables is accurate and therefore well specified in previous research. Furthermore, they acknowledge as we do in our appendix that inequality is not easy to measure. Starting with Deininger and Squire (1996) study, data on inequality have been continuously improving, including the data set that we use in this research, but the potential for error remains. The question remains whether the data on inequality and growth can be appropriately compared over time for a given set of countries. Finally, Banerjee & Duflo (2003), ask what are the right time lags for estimating such models? For example, Banerjee & Duflo (2003), state the net effect on growth coming from the initial reduction in inequality is obviously much smaller than the impact effect, and we can clearly have shocks to inequality that are costly in the short run but beneficial over a longer horizon. For these reasons, we think that the relationship between inequality and economic growth is best specified with quadratic inequality term and an appropriate set of control variables. Barro (2000) states that the choice of the control variables can be critical in such studies because the inequality measure, the Gini coefficient could proxy for omitted variables.

Accordingly, Model (3) specified above include our dependent variable as the growth rate, and the main explanatory variable is income inequality, represented by Gini coefficients. Other control variables include real level GDP, a country's population, government health spending, democracy index, trade index, and inflation. Data collected for the creation of variables *RGDP* (and *RGDP_LAG*), *GROWTH_RATE*, and income group specifications came from the Penn World Table (PWT) dataset by Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer (2015). This dataset allowed for a more contemporary look into our variable of interest while still maintaining a high amount of observations when compared to the Madison Dataset, which also has a large number of observations but scarcely any data with consistency post 2000. Using the PWT dataset, a year by year growth rate was calculated for all countries by using a growth formula that accounted for varying gaps in time.

$$G = \left(\frac{Y_t}{Y_{t-n}}\right)^{\frac{1}{T_t - T_{t-n}}} - 1$$

Where G is the growth rate, Y_t is RGDP at time T and Y_{t-n} is output from n years before, T_t being the year Y_t is recorded in and T_{t-n} being the year Y_{t-n} is recorded in. With this representation of G, varying gaps in time can be overlooked when calculating growth rates for large datasets such as the PWT dataset. Moreover, using the population data that came with the PWT allowed for the calculation of per capita GDP (*PCGDP*) and allocating for high income and low income group countries by percentiles.

Data on the variable income inequality (*GINI_LAG*) was collected from The UNU-WIDER World Income Inequality Database (WIID). Initially compiled in 1997-99 for a UNU-WIDER project, this dataset has been updated and now covers observations up to the year 2018 in 200 countries. This dataset provides information including country name and code, year the survey was taken, Gini coefficients as reported based on microdata, quintile measures of income distribution, unit of analysis (weighted by person or household), population and area coverage, region and income group classification based on World Bank and United Nations, source of observations, and other less relevant information. This is a very comprehensive source of data on income inequality overall, however, the measurement issues referred to by Banerjee & Duflo (2003) are likely applicable in our study as well. We make reference to this in Appendix A.

As per policy recommendation from research conducted by Sarkar (2008), health spending or education spending, both combat the feedback loop on inequality. For our study, healthcare spending was chosen as the data largely because of more observations than education spending. The data was collected from Ortiz-Ospina, Esteban, and Max Roser (2017) and was constructed using two datasets in order to create government spending on health as a percent of GDP, multiplying (government health spending as a percent of total health expenditure) * (total health expenditure as a percent of GDP).

Following Acemoglu, Daron, et al. (2001), we chose to use the *DEMOCRACY_INDEX* to measure difference in political institutions across our 182 countries. The data was collected from *Gapminder*, which compiles the annual ratings released by the Economist Intelligence Unit. Data on the variable *INFLATION* was collected from The World Bank database. Here inflation is measured by the consumer price index, reflecting the annual percentage change in the cost to the average consumer of acquiring a basket of goods and services.

Finally, for the trade openness variable, rather than simply look at trade as a percentage of GDP and given our sample of 182 countries, our goal was to generate an index that is as all-encompassing as possible. Therefore, we decided on including three sub-variables for the index construction: trade as a percentage of GDP, average tariff rate applied across all products, a country's status of membership in The World Trade Organization (WTO). Among the three, data on trade as a percentage of GDP was gathered from the World Bank – World Development Indicators, with the variable being measured as the sum of exports and imports of goods and services as a share of gross domestic product. Data on weighted mean applied tariff on all products was also collected from the same source. For the two variables, tradeshare and average tariff rate, the time span was selected to match the period between as early as 1980 to the most recent 2017 or 2016. Due to the relative complex nature of index construction, a detailed description of how we constructed the trade index can be found in Appendix B.

IV. Results

The regression results for the basic convergence test are presented in Table 1. Here, *RGDP1k_2000* represents the initial level of real GDP (in millions of dollars) in the year of 2000 for all the countries, scaled or divided by 1000 to overcome the scaling issue and to interpret the coefficient more easily. At both the cross-country and the panel level, the relationship between initial level of income in 2000 and growth rate over the 17-year period between 2000-2017 is negative. Specifically, for example in column (2), for every 1 billion dollar increase in GDP, growth rate is expected to decrease by 0.000487% (or for every 1 trillion dollars increase in GDP, growth rate will decrease by 0.487%), significant at the 5% significance level, other things constant. This suggests that convergence among the 182 countries likely occurred during the 17-year period. The magnitudes of the coefficients are in fact not so trivial.

The rest of the table displays the convergence tests for sub-groups of countries classified by geographical regions by the World Bank. Except for South Asia, North America, the Middle East and North Africa, all the other four regions show some within-region convergence. Among those, the effect of base level GDP on growth rate is significant at 1% level for East Asia and Pacific, Europe and Central Asia, and Sub-Saharan Africa; while it is significant at the 5% level for Latin America and the Caribbean. In general, the results suggest that in these specific regions, the countries exhibit shrinking of the gap between poorer and richer nations over time.

Model (2) is our conditional convergence test, where we added our main independent variable, *GINI_LAG* (Gini lagged one year), followed by other control variables health spending, population, democracy index, trade index, inflation, and real GDP, with all being lagged one year in order to match with the lagged Gini**.** Table 2. summarizes our cross-country, panel level, and fixed effects regressions. In general, while the coefficient on inequality (*GINI_LAG*), is positive, it does not seem to have a significant effect on growth rate. Among the control variables, government spending on health, population, the extent to which a country is democratic, inflation, and level GDP do appear to contribute to growth rate. Column (3) includes both time-invariant and country-specific fixed effects. As stated earlier, the fixed effect model allows us to account for unobservable factors that vary over time among countries, as well as unobservable factors that remain the same over time but do vary across different countries. The results of the fixed effect model in column (3) appear to improve on the results from the simple panel regressions. The variable *GINI_LAG* is now significant at the 5% level, and interestingly, the sign has turned from positive to negative. This means that for each 1 percentage increase in Gini, the growth rate is expected to decrease by 0.102%, other things constant. Furthermore, with fixed effects, adjusted-R^2 surged from 0.152 to 0.536, suggesting a much larger proportion of variation in the dependent variable being explained by the explanatory variables. Similar to the empirical procedure in the convergence test process, we regrouped the 182 countries based on their income level to minimize the potential wide variation in income and with an eye towards considering differing levels of development. Thus, the two groups we ran regressions on are low-income and high-income countries, and the last two columns of Table 2. show the results. However, no statistical significance in inequality is detected.

Table 1. Convergence at Global and Regional Level

	(1) (Cross Country)	(2) (Panel)	(3) (East Asia and Pacific)	(4) (South Asia)	(5) (Europe and Central Asia)	(6) (Middle East and North Africa)	(7) (Sub_Saharan Africa)	(8) (North America)	(9) (Latin America and the Carribean)
VARIABLES	Growth_Rate	Growth_Rate	Growth_Rate	Growth_Rate	Growth_Rate	Growth_Rate	Growth_Rate	Growth_Rate	Growth_Rate
rgdp1k_2000	-0.000487**	-0.000487**	-0.00140***	-0.00196	-0.00311***	0.00144	-0.00974***	-0.000556	-0.00586**
	(0.000201)	(0.000201)	(0.000318)	(0.00222)	(0.000940)	(0.00188)	(0.00254)	(0.000833)	(0.00278)
pop	0.00459***	0.00485***	0.00745***	0.00546	0.0637**	-0.0217	0.0406***	0.0266	0.0659**
	(0.000997)	(0.000972)	(0.000916)	(0.00515)	(0.0273)	(0.0220)	(0.0107)	(0.0358)	(0.0315)
Constant	3.883***	3.896***	5.305***	5.523***	3.541***	4.013***	4.440***	1.176	2.566***
	(0.168)	(0.165)	(0.611)	(0.545)	(0.337)	(0.627)	(0.325)	(0.690)	(0.264)
Observations	182	3,250	359	126	802	357	808	54	690
R-squared	0.081	0.080	0.303	0.112	0.249	0.028	0.175	0.317	0.078
Adj. R-squared	0.0708	0.0790	0.300	0.0975	0.247	0.0227	0.173	0.290	0.0756

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.

Following Model (1), we included the level real GDP *RGDP1k_2000* instead in place of the lagged variable *RGDP1k_lag*. This result is shown in Table 3. Once again, the level of real GDP *RGDP1k_2000* is only significant in the panel data (column (2)) at the 10% level. Similarly, later in Table 5. (column (2)), *RGDP1k_2000* is again significant in the panel data at the 5% level, confirming conditional convergence.

Table 2. Full OLS Model with Income Group Regressions (rgdp1k_lag)

VARIABLES	**Cross Country** Growth_Rate	**Panel** Growth_Rate	**Fixed Effects** Growth_Rate	**F.E. Low Income** Growth_Rate	**F.E. High Income** Growth_Rate
gini_lag	0.0285	0.00728	-0.102**	-0.0364	-0.0866
	(0.0519)	(0.0192)	(0.0483)	(0.0613)	(0.0718)
healthspend_lag	-0.869***	-0.466***	-0.0312	-1.383***	0.399
	(0.265)	(0.119)	(0.418)	(0.392)	(0.267)
pop_lag	-0.00149	0.00273***	0.0376*	0.00175	0.252***
	(0.00130)	(0.000895)	(0.0205)	(0.00791)	(0.0623)
democracy_index_lag	-1.883	-1.409**	2.446*	2.936**	1.652
	(1.592)	(0.606)	(1.463)	(1.449)	(2.441)
trade_index_lag	-0.00336	0.000424	0.0151	0.0266	0.0274
	(0.00621)	(0.00291)	(0.0160)	(0.0233)	(0.0177)
inflation_lag	-0.0746*	-0.0514**	-0.0148	-0.0176	-0.0509
	(0.0375)	(0.0214)	(0.0164)	(0.0187)	(0.0425)
rgdp1k_lag	0.000103	-0.000150**	-0.000254	6.80e-05	-0.00226***
	(0.000118)	(7.29e-05)	(0.000241)	(9.56e-05)	(0.000565)
Constant	7.936**	6.738***	4.011	5.722	-3.405
	(3.119)	(1.268)	(3.638)	(4.193)	(5.056)
Observations	80	820	820	344	476
R-squared	0.364	0.159	0.612	0.603	0.666
Adj. R-squared	0.302	0.152	0.536	0.468	0.604
SER	3.228	3.325	2.459	2.622	2.129

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 3. Full OLS Model with Income Group Regressions (rgdp1k_2000)

VARIABLES	**Cross Country** Growth_Rate	**Panel** Growth_Rate	**Fixed Effects** Growth_Rate	**F.E. Low Income** Growth_Rate	**F.E. High Income** Growth_Rate
gini_lag	0.0291	0.00691	-0.105**	-0.0348	-0.0902
	(0.0515)	(0.0191)	(0.0479)	(0.0599)	(0.0687)
healthspend_lag	-0.863***	-0.470***	-0.0432	-1.374***	0.399
	(0.258)	(0.118)	(0.417)	(0.387)	(0.272)
pop_lag	-0.00105	0.00242***	0.0283**	0.00412	0.0963
	(0.000843)	(0.000801)	(0.0114)	(0.00759)	(0.0636)
democracy_index_lag	-1.930	-1.384**	2.432*	2.947**	2.495
	(1.582)	(0.606)	(1.465)	(1.444)	(2.547)
trade_index_lag	-0.00327	0.000289	0.0154	0.0266	0.0287
	(0.00623)	(0.00290)	(0.0160)	(0.0233)	(0.0178)
inflation_lag	-0.0751**	-0.0518**	-0.0149	-0.0175	-0.0452
	(0.0377)	(0.0214)	(0.0164)	(0.0186)	(0.0422)
rgdp1k_2000	0.000118	-0.000164*			
	(0.000133)	(8.77e-05)			
Constant	7.915**	6.769***	4.496	5.457	-0.925
	(3.105)	(1.259)	(3.440)	(3.950)	(4.666)
Observations	80	820	820	344	476
R-squared	0.364	0.160	0.611	0.603	0.660
Adj. R-squared	0.302	0.152	0.536	0.470	0.598
SER	3.228	3.325	2.459	2.617	2.145

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Following our Model (3) specification stated earlier, we projected a non-linear relationship between inequality and growth, and included a quadratic term Gini-squared, *GINI_2.* Table 4. illustrates that at the cross-country level, *GINI_LAG* was not significant. On the other hand, the linear panel fixed effect data regression run shows that a 1 percentage increase in Gini increases growth rate by 0.213%, other things constant, at the 5% level. The quadratic term is also statistically significant in this regression. In columns (4) and (5), we show the regressions for low-income countries and high-income countries. Inequality is statistically significant for the lower income countries but not for the higher income countries, suggesting that changes in inequality have no effect on wealthy countries, but has some effect on poorer countries.

Overall, we conclude that including the quadratic provides a better specification of the relationship between inequality and economic growth. Our result seems to be consistent with a non-linear specification with an introduction of a quadratic term prompted largely by the Banerjee & Duflo (2003) study. Table 5. shows similar regression runs as in Table 4., but with *RGDP1k_2000*, the initial level of GDP, to test for conditional convergence.

Table 4. Non-Linear Model (rgdp_lag)

VARIABLES	**Cross Country** Growth_Rate	**Panel** Growth_Rate	**Fixed Effects** Growth_Rate	**F.E. Low Income** Growth_Rate	**F.E. High Income** Growth_Rate
gini_lag	0.348	0.213**	0.427	0.840**	0.00491
	(0.214)	(0.106)	(0.258)	(0.356)	(0.232)
gini_2	-0.00376*	-0.00238**	-0.00628**	-0.00987**	-0.00121
	(0.00218)	(0.00114)	(0.00297)	(0.00381)	(0.00284)
pop_lag	-0.00122	0.00289***	0.0346*	-0.00275	0.245***
	(0.00131)	(0.000892)	(0.0188)	(0.00627)	(0.0632)
healthspend_lag	-0.820***	-0.433***	0.0289	-1.154***	0.409
	(0.271)	(0.121)	(0.414)	(0.420)	(0.265)
democracy_index_lag	-1.745	-1.286**	2.612*	3.159**	1.627
	(1.601)	(0.620)	(1.382)	(1.286)	(2.455)
trade_index_lag	-0.00253	0.000773	0.0148	0.0281	0.0273
	(0.00657)	(0.00292)	(0.0157)	(0.0229)	(0.0178)
inflation_lag	-0.0636	-0.0496**	-0.00865	-0.00925	-0.0501
	(0.0388)	(0.0213)	(0.0169)	(0.0182)	(0.0423)
rgdp_lag	6.62e-08	-1.79e-07**	-1.84e-07	9.94e-08	-2.21e-06***
	(1.24e-07)	(7.66e-08)	(2.32e-07)	(8.88e-08)	(6.04e-07)
Constant	1.101	2.317	-6.494	-13.09	-4.845
	(5.905)	(2.682)	(6.293)	(9.297)	(6.354)
Observations	80	820	820	344	476
R-squared	0.374	0.164	0.617	0.619	0.666
Adj. R-squared	0.303	0.156	0.542	0.487	0.603

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 5. Non-Linear Model (rgdp1k_2000)

VARIABLES	**Cross Country** Growth_Rate	**Panel** Growth_Rate	**Fixed Effects** Growth_Rate	**F.E. Low Income** Growth_Rate	**F.E. High Income** Growth_Rate
gini_lag	0.349	0.212**	0.433*	0.840**	0.0869
	(0.213)	(0.105)	(0.257)	(0.355)	(0.252)
gini_2	-0.00376*	-0.00237**	-0.00637**	-0.00984**	-0.00234
	(0.00218)	(0.00114)	(0.00296)	(0.00380)	(0.00321)
pop_lag	-0.000941	0.00251***	0.0279**	0.000712	0.0903
	(0.000850)	(0.000793)	(0.0116)	(0.00694)	(0.0579)
healthspend_lag	-0.816***	-0.438***	0.0212	-1.142***	0.418
	(0.263)	(0.119)	(0.412)	(0.415)	(0.271)
democracy_index_lag	-1.775	-1.257**	2.604*	3.174**	2.406
	(1.594)	(0.621)	(1.382)	(1.285)	(2.573)
trade_index_lag	-0.00248	0.000617	0.0150	0.0281	0.0284
	(0.00658)	(0.00290)	(0.0158)	(0.0228)	(0.0179)
inflation_lag	-0.0639	-0.0501**	-0.00862	-0.00917	-0.0439
	(0.0391)	(0.0214)	(0.0170)	(0.0182)	(0.0423)
rgdp1k_2000	7.44e-05	-0.000195**			
	(0.000143)	(9.17e-05)			
Constant	1.076	2.372	-6.302	-13.43	-3.826
	(5.880)	(2.662)	(6.278)	(9.175)	(6.279)
Observations	80	820	820	344	476
R-squared	0.374	0.164	0.617	0.619	0.661
Adj. R-squared	0.303	0.156	0.543	0.489	0.597

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Additionally, the presence of a quadratic term in Table 4. allows us to empirically capture where growth is maximized for low income countries. We derive the process as follow. First, we can rewrite the growth equation with inequality as a function of $gini\_lag$ and $(gini\_lag)^2$ as:

$$\widehat{GROWTH\_RATE} = \beta_1 * GINI\_LAG + \beta_2 * (GINI\_LAG)^2$$

Then, taking the first derivate and setting it equal to zero allows us to find the highest value of Gini that maximizes economic growth:

$$\beta_1 + 2\beta_2 * gini = 0$$
$$2\beta_2 * gini = -\beta_1$$
$$gini = -\frac{\beta_1}{2\beta_2}$$

Regression	**Coeff. on Gini '$\boldsymbol{\beta_1}$'**	**Coeff. on (gini)^2 '$\boldsymbol{\beta_2}$'**	**Gini* = $-\frac{\beta_1}{2\beta_2}$**	**Growth max**
World	0.427	-0.00628	33.99%	7.26%
High Income	0.00491	-0.00121	2.03%	.0049%
Low Income	0.840	-0.00987	42.55%	17.87%

We should note that in regressions for the world and high-income group, the Gini coefficient is not significant. It is, however, significant for low-income countries. We could therefore interpret the above result to mean that for low-income countries, a target Gini of 42.55% will maximize growth rate to 17.87% as they try to converge towards the high-income countries. Once they reach the PCGDP threshold to be classified as a high-income country, the effect of Gini on growth rate will no longer be significant.

Our overall remark from this research is as follows. From Model (1) and the results in Table 1., we found that there is evidence for convergence for the list of 182 countries over the period 2000-2017. Furthermore, our more detailed approach conducting regressions on countries based on their geographical regions revealed even stronger evidence for the existence of convergence during the same time frame. In Model (2) and (3), where we tested for conditional convergence and added inequality as the main explanatory variable, we found that in the case of a simple linear model, growth is primarily negatively associated with income inequality across countries. When adopting a non-linear quadratic model, we observe the relationship between inequality and growth to be also primarily negative, and the two kinds of relationship hold true at both the cross-country and the panel level. Additionally, there is some evidence suggesting that countries exhibit convergent growth pattern in the setting of a conditional convergence test.

V. Conclusion

In their extremely comprehensive recent paper in the Journal of Economic Literature, "What remains of Cross-Country Convergence ?", Johnson and Papageougiou (2020) state that convergence is hard to pin down, "because the concept can be operationalized in many ways and second, because econometric approaches and data measurement issues remain a challenge in empirical tests of convergence." This would also be the case when including inequality as an explanatory variable into the growth literature. Considering our results, we have a better appreciation for the empirical testing process but also the ambiguity within this area. One important finding from our research is that the link between inequality and growth does not appear to be a factor for the wealthy or high-income countries in our model. Johnson and Papageougiou (2020) cite Branko Milanovic's work on global inequality where Milanovic finds that while inequality is rising for the high-income countries, "global inequality of incomes though huge, has been falling particularly since 2000" (Milanovic, 2016).

In specifying a 21st century approach starting with the year 2000 in our current research, we believe we are capturing this observation quite adequately. Furthermore, we follow up on the wealth effect discussion in Banerjee and Duflo (2003) in our empirical approach where inequality matters more in poorer countries due to initial higher interest rates and difficulties in accumulating capital. Bannerjee and Duflo (2003) derive a theoretical model in which they show that over a long enough period, two economies that start at the same average wealth level will exhibit the same average growth rate, since they both would have gone from the initial mean wealth to an average wealth in the long-run, say $w*$.

By using a more contemporary dataset for the period 2000-2017, while acknowledging the limitations involved in empirical testing of growth convergence theory in general, and the relationship between inequality and growth in particular, our research offers some useful contributory inferences. More elaborate empirical techniques, and utilizing cleaner and better inequality and growth data as they become available will almost certainly continue to be on the agenda for future researchers.

VI. References

Acemoglu, Daron, Simon Johnson, and James A. Robinson. "The Colonial Origins of Comparative Development: An Empirical Investigation." The American Economic Review 91, no. 5 (2001): 1369-401

Aleš Bulíř. "Income Inequality: Does Inflation Matter?" IMF Staff Papers 48, no. 1 (2001): 139-59

Banerjee, Abhijit V., and Esther Duflo. "Inequality and Growth: What Can the Data Say?" Journal of Economic Growth 8, no. 3 (2003): 267-99

Barro, Robert J. "Inequality and Growth in a Panel of Countries." Journal of Economic Growth 5, no. 1 (2000): 5-32

Ben-David, Dan, and David H. Papell. "Slowdowns and Meltdowns: Postwar Growth Evidence from 74 Countries." The Review of Economics and Statistics 80, no. 4 (1998): 561-71

Birdsall, Nancy, David Ross, and Richard Sabot. "Inequality and Growth Reconsidered: Lessons from East Asia." The World Bank Economic Review 9, no. 3 (1995): 477-508

"Democracy Index." Democracy Index | Gapminder. https://www.gapminder.org/data/documentation/democracy-index/.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. "The Next Generation of the Penn World Table." The American Economic Review 105, no. 10 (2015): 3150-182

Helpman, Elhanan, Oleg Itskhoki, and Stephen Redding. "INEQUALITY AND UNEMPLOYMENT IN A GLOBAL ECONOMY." Econometrica 78, no. 4 (2010): 1239-283

Stephen P. Jenkins. "World income inequality databases: an assessment of WIID and SWIID." The Journal of Economic Inequality *13* (4), (2015): 629-671

Johnson, Paul, and Chris Papageougiou. "What remains of cross-country convergence?" Journal of Economic Literature *58*(1), (2020): 129-175

Mankiw, N. Gregory, David Romer, and David N. Weil. "A Contribution to the Empirics of Economic Growth." The Quarterly Journal of Economics 107, no. 2 (1992): 407-37

Branko Milanovic. "Global inequality: A new approach for the age of globalization". Cambridge, Massachusetts; London, England: Harvard University Press, (2016)

Ortiz-Ospina, Esteban, and Max Roser. "Financing Healthcare." Financing Healthcare - Our World in Data, (2017). https://ourworldindata.org/financing-healthcare.

Persson, Torsten, and Guido Tabellini. "Is Inequality Harmful for Growth?" The American Economic Review 84, no. 3 (1994): 600-21

Pritchett, Lant. "Divergence, Big Time." The Journal of Economic Perspectives 11, no. 3 (1997): 3-17

Sarkar, Jayanta. "Mortality, Fertility, and Persistent Income Inequality." Southern Economic Journal 75, no. 2 (2008): 332-50

Senhadji, Abdelhak. "Sources of Economic Growth: An Extensive Growth Accounting Exercise." IMF Staff Papers 47, no. 1 (2000): 129-57

Solow, Robert M. "A Contribution to the Theory of Economic Growth." The Quarterly Journal of Economics 70, no. 1 (1956): 65-94

Appendix A: Clarification of Income Inequality Database

In a study, "World Income Inequality Databases: An Assessment of WIID and SWIID," Stephen Jenkins (2015) discusses several potential issues in the WIID database. First, the definitions of the source of income inequality are different. Both the income and consumption inequality measures of inequality are utilized in this database. The income measure is often more easily available for high-income countries, while the consumption expenditure measure is more commonly used for low-income countries. Moreover, the difference between net income and gross income used when measuring income distribution raises another potential problem. Third, the reference unit—whether inequality is measured against household income or individual income within a household is another potential limitation in the data. In working with this inequality dataset, we followed the suggestion by Jenkins, and limited observations to where the *Quality Score =1 (a column of quality rating is part of the dataset; a score of 1 refers to the highest quality).* For some less developed countries where none of these criteria were met, we used the Gini reported based on consumption expenditure, and also obtained data when missing in the WIID database from sources such as FRED (St Louis Federal Reserve's Economic Database) and the World Bank's POVCAL package. Though the process of cleaning the inequality data reduced our overall number of observations, what remained, totaling to approximately 2000, still provided a good-sized number of observations on the variable *gini*.

Appendix B: Construction of Trade Openness Index

In constructing the trade openness index, we first approached how to assign weights to each of the three variables. First, we regressed growth rate on trade share, tariff rate, and membership status individually with both time and country-specific fixed effects, and the resulting coefficients on trade share, tariff rate, and membership were 0.004848, -0.0012887, and 0, respectively. By adding the three coefficients (absolute value of the effect size) together, we get a total number of 0.0061367. We then divided each of the three coefficients by this total value to get a percentage share of each. Specifically, trade share takes up around 79% (0.004848 / 0.0061367) of the weight, and tariff rate accounts for roughly 21% (0.0012887 / 0.0061367), with membership having no weight. Our *trade_index* variable was then generated by multiplying each variable by its corresponding weighted percentage, i.e. tradeshare*0.79, tariff_rate*0.21, and adding them together.