University of Macedonia

Department of Economics

Interdepartmental postgraduate program

Master in Economics

A Reinvestigation of the Existence of Multiple Cross-Country Growth Regimes

Dalakouras Alexandros

Supervisor: Boikos Spyridon

A thesis submitted in partial fulfillment of requirements for the degree of Master in Economics

August 2022

Acknowledgments

I would like to express my undivided gratitude to my professor, Mr. Boikos Spyridon, for his support on multiple levels along with his precious advices throughout my studies. I would like to thank him and of course the rest of the educational staff who have been providing to me useful information allowing me to submit this master thesis.

Furthermore, I would like to thank my family, my friends and my bosses for boosting me throughout this tough process.

Abstract

The purpose of this study is to replicate and extend a part of the study ʺ Multiple regimes and cross-country growth behaviorʹʹ carried out by Steven N. Durlauf and Paul A. Johnson (1995). Despite the dominant theory of a single-regime model existence in cross-country growth theory, they have found strong evidence of multiple-regimes. Firstly, we are exploring the literature that are directly or indirectly related to this topic. On the second part, we present in detail the S.N. Durlauf and P.A. Johnson (1995) theory framework. On the third part, we follow the reasoning of the authors and empirically try to reassess their method, focusing on estimating the unconstrained regressions accompanied with the relative Wald tests. These estimations cover the 1960-1985 (96 countries), 1960-2015 (96 countries) and 1960-2015 (108 countries) time frames. Our findings are not always consistent to the main study findings, due to mixed evidence regarding the regimes in the growth process. Finally, we draw the conclusion based on the most important results of our empirical approach. Overall, we point out that: i) for the 1960-1985 time frame, the results are mixed, regarding the multiple regime existence, as we reject the single regime specification for some cases (under literacy rate in two-way split and when the variables interact) but we do not for some others (under the output per capita in the two-way split and in all cases in the three-way split) ii) for the expanded set of countries and span of time, we still produce mixed results (rejection of the single regime only when under output per capita we break the sample at three parts and when we let the variables interact with each other), showing a contradiction between the control variables and iii) for the 1960-2015 using the original sample, we find exclusively evidence of a single regime existence.

Keywords:

Convergence, multiple-regimes, growth, subsamples

Contents

1. Introduction

The main goal of this study is to empirically examine the existence of multiple regimes in cross-country growth behavior. This objective can be achieved following the pattern used by Steven N. Durlauf and Paul A. Johnson in their research ''*Multiple regimes and cross-country growth behavior*'' (1995) according to which the same data were for the same time period. Additionally, the study examines the existence of multiple regimes in an expanded sample both in selected countries and in a given time period.

Specifically, S.N. Durlauf and P.A. Johnson -based on the Summers-Hestons dataset- focus on finding evidence regarding the existence of multiple regimes in the cross-country growth behavior. In fact, they examine a group of 96 countries[1] during the 1960-1985 time period. Taking into consideration the analysis described by Mankiw-Romer-Weil (1992) (hereinafter mentioned as M-R-W), they split the data into subgroups based on the initial values of real GDP per capita (Y/L_{1960}) and literacy rate (LR_{1960}) and they run specifications tests, examining whether the model parameters are equal across these subgroups. Results provide strong evidence, like for example the rejection of the hypothesis that all the countries in the analysis follow a common specification model. That, contradicts the currently dominant theory[2] according to which the cross-country growth behavior follows a common linear model. Despite the failure to establish these results as formal ones, they provide enough arguments on why the economic science needs to search deeper for the factors shaping up the cross-country growth pattern.

Furthermore, Durlauf and Johnson (1995) pursue locating economies that follow a common linear model and they group them together. This process can only be held by examining the data endogenously. Thus, the regression tree analysis technique, which has been described for the first time in Breiman *et al.* (1984), helps the authors categorize

[1] They use the sample of countries straight from M-R-W (1992) analysis excluding Botswana and Mauritius due to lack of data regarding literacy.

[2] Some papers with the most important that of Azariades and Drazen (1990), provide theoretical mechanism which questioning the common linear specification for all countries.

countries that follow the same regime. A strong argument to use this technique is that allows researchers to examine the existence of an infinite number of potential regimes[3]. As a result, this technique provides the authors with overall four non-equal groups of countries. The sample is split mostly based on the initial value real GDP per capita, which means that $(\frac{Y}{L})_{i,1960}$ is more useful than initially literacy rate when identifying regimes. Finally, the results indicate that different production functions exist in economies with different initial conditions.

Following the reasoning of the authors, we use the data we have obtained from various databases and for different time periods. The databases in question include the latest version of Penn World Table, the World Bank database as well as the Barro-Lee table. The sample of the countries used in the study has been retrieved from the M-R-W (1992) research and includes the countries referred to as the non-Oil producing ones (the other two categories are the Intermediate and the OECD countries). For our calculations we use the Eviews econometric program. Firstly, we split the sample properly into subgroups and then we run Ordinary Least Squares regressions for the sample as a whole and for the subgroups separately. The results are not always consistent quantitatively with the ones produced by Steven N. Durlauf and Paul A. Johnson (1995), but the numerical indications seem to be much more accurate. We then employ Wald tests to check the equality of parameters across the indicated subgroups. At the last step, we use an expanded set of variables which we have recovered from Barro (1991). Running again the OLS and Wald tests in the augmented equations, we draw the conclusion that the results don't differ from what we expected. At this point, it is wise to mention that we focus on estimating the unconstrained equations and we exclude the constrained versions which impose non-linear restrictions to the parameters of the model.

The rest of the study is structured as follows: Section 2 explores the literature review and the presentation of the main study framework. In Section 3 the data and the variables that

[3] No currently existing theory provides a numerical limit about the potential regimes.

we used in our empirical approach are described along with a concise presentation of the used methodology. In Section 4 the empirical results are presented whilst in Section 5 the most important outcomes are summarized. Appendices A and B provide some extra information about the data and the specification tests, respectively.

2. Theoretical background

2.1 Literature Review

In the past but nowadays more intensively-due to globalization, many people choose mostly for employment purposes to relocate to richer and more developed countries. In these countries, jobs are constantly being created and the working environment is mature. That means, obviously, that the standard of living is different between rich and poor countries. But, why are some countries richer than others? Are the underdeveloped and poorer countries able to catch up with the richer ones? If yes, how fast can they converge to richer ones? Are there specific factors that promote economic growth in all economies or do different factors have different effects on different economies? These are some of the most debatable questions among economists to which they respond by developing a big variety of theories in their endless and tough pursuit to help underdeveloped countries put an end to their poverty.

Real Growth Domestic Product (RGDP) is used to give a first description of the economy of a country. The majority of the countries seek to attract investments in order to raise their RGDP. At this point, it is useful to mention, that RGDP *alone* isn't a very clear factor of the economic situation of a country. Governments and authorities should also focus in the income distribution as well, taking into consideration instruments- among others- such as the Gini or Theil Index. Despite the great number of theoretical models and instruments that help policy-making, it is very challenging for economists and governments to build the perfect path for each country.

Most of the researchers whilst studying cross-country growth behavior focused on finding evidence of convergence among economies. Convergence is defined as the condition in which underdeveloped countries are growing up much faster than the developed ones, as far as their GDP is concerned, and they finally reach a steady state. Baumol at his article *''Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show''* (1986) argues that there are signs of convergence among the industrialized countries. Examining postwar data, Baumol (1986) finds that the convergence phenomenon can be spotted in the intermediate level countries as well as the centrally-planned economies. The results regarding the poorest countries don't indicate great convergence. Following Baumol, many scientists are engaged in studying cross-country growth behavior. Currently, in the growth theory there are two aspects dividing the economic theory. On the one hand, there are those who believe that all of the countries under analysis follow a single growth equation. The famous Solow-Swan (1956) closed economy model works under a common linear specification. At this point it is wise to mention that this model is an extension of the Harrod-Domar (1946) model, in the sense that labor is allowed to substitute physical capital. Solow (1956) assumes that consumption-savings and investment decisions[4] are exogenous as are labor and technology accumulation as well. Furthermore, input markets and goods markets[5] are assumed to be in perfect competition. It is worth mentioning that the Solow-Swan model is the neoclassical (Cobb-Douglas) production function hereof, and it is based on constant-returns-to-scale and on diminishing returns regarding its input elements. This model includes, also, constant depreciation rate of the physical capital. Nonetheless, the most important contribution of this model is that it suggests conditional convergence. Solow argues that the rate of return is fairly lower in countries with higher capital accumulation, and thus capital will flow to poorer countries where rates of return in physical (and human) capital tend to be high. Also, underdeveloped countries have the chance to benefit by spreading knowledge, especially in globalized world-conditions. They

---

[4] In every single time period, savings equal investments and that excludes the flow of savings into non-investing decisions, like for example profiting from currencies imbalances.

[5] In the Solow model there is production of only one good examined.

can adopt technologies invented by technological giant countries without having to spend resources of their own, in the sense of investing in research and development from scratch. Yet, the Solow model isn't perfectly structured as even though it performs well regarding the effects that savings and population growth have, it doesn't perform well regarding the magnitudes. For instance, more than half of the cross-country variation in income can be explained by savings and population growth. Convergence (also known as catch-up effect) is examined in the study of M-R-W (1992), where authors experiment with an augmented version of the Solow-Swan model. Their findings seem to be consistent with what Solow suggested even though they were trying to produce a model performing better with regards to the magnitudes of the effects. Therefore, they add the human capital factor in their augmented model. Interestingly, they find that human capital is correlated with savings and population growth. Their model explains variation in income of about 80 percent. It gives therefore a pretty clear explanation of why there are rich and poor countries. In their effort to identify convergence, authors include in their research some countries that most likely haven't exhibited any convergence in previous scientific works. Surprisingly, empirical process reveals the existence of convergence, despite the initial thoughts of the authors that there would be no convergence. Their paper, to sum up, is questioning the beliefs of some economists engulfing exclusively endogenous growth models.

On the other side, however, there are some new alternative ideas that examine cross-country growth behavior under the belief that after a certain point of human or physical capital accumulation, the aggregate production function stops being concave. In other words, these new growth models claim that the cross-country growth behavior is not explained by a single growth equation, but rather from various ones. In a study pioneered by Azariades and Drazen (1990), the authors examine cross-country growth behavior under an augmented neoclassical growth model in which they include a feature that they call ''*technological externalities with a <<threshold>> property*''[6] [Azariades and Drazen (1990),

[6] [Azariades and Drazen (1990), p.503]

p.503]. With this function they manage to document steep changes in the production that occur when some of the explanatory variables surpass a certain break point. They include in their model externalities that being produced from different types of capital *'' in order to capture in a primitive way the notion of <<infrastructure>>''*[7] [Azariades and Drazen (1990),p.504]. Additionally, they are taking into account the effects of the externalities that the human capital aggregation has. The assumption of perfect credit markets excludes every potential intervention from the capital market sector to growth. In a similar spirit, Durlauf and Johnson (1995) examine data for a large group of countries pursuing on finding evidence that multiple regimes exist. Initially, it was necessary to split data into subgroups based on two different control variables[8]. Afterwards, the authors run specification tests considering a single-regime model as the null hypothesis. Their empirical approach is based on examining both a constrained and an unconstrained version of the Solow model. Their results reinforce the rejection of the use of a common linear model when studying cross-country growth behavior. Intuitively, their results suggest that the subsamples under analysis exhibit different production functions. Many criticized these results, arguing that omitted variables lead to inaccurate estimates and therefore to the presence of multiple regimes. A way to verify these results is to extend the number of the included explanatory variables and to test whether the results differ. Barro (1991) proposes a number of variables to be included in the model. These variables are not pure economic, but rather variables that include geographical, social and educational factors. The authors re-estimate the constrained and unconstrained augmented equations and they are running again specification tests. Thus, the results don't differ from the ones produced without the broader set of variables. Since they strengthened their claims regarding the existence of multiple regimes, they pursue to group together countries that follow the same regime. To achieve that goal, they use the regression tree analysis, a Breiman's (1984) idea.

[7] [Azariades and Drazen (1990), p.503]

[8] The split of variables is quite necessary, as DeLong (1988) stated, to avoid the selection bias problem. If a selection bias problem occurs, it means that there is a failure considering the proper characteristics of the subgroups under analysis within the total sample. Therefore, the outcomes of the analysis would be inaccurate and the relationships between the variables would probably be explained in a wrong way.

Interestingly enough, this process shows a higher preference to output per capita, rather than literacy, as a helpful variable in identifying various regimes in the sample, as mentioned before.

With the passage of time, even more researchers focus on studying cross-country growth. The research of Corrado L., Stengos T., Weeks M. and Yazgan M. (2018) proposes a *'' multiple pairwise comparisons method based on a recursive bootstrap*[9] *to test for convergence with no prior information on the composition of the convergence clubs''*[10]. The use of bootstrap provides more powerful evidence of convergence existence. Similarly, Beylunioğlu F., Stengos T. and Yazgan M. (2018) propose an extraordinary method of locating convergence groups, using maximum clique and maximal clique algorithms. Their method is quite better than the majority of the others, providing evidence of convergence existence. The most recent research of Omerovic S., Friedl H. and Grun B. (2022) with title *''Modelling multiple regimes in economic growth by mixtures of Generalised Nonlinear Models''* exhibit -under the analysis of Generalised Nonlinear Models- different laws of motion between subgroups of economies. Club convergence is examined in the research of Galor (1996) as well. Studies such as Corrado L., Martin R., Weeks M. (2005), Dufrenot G., Mignon V., Naccache T. (2012), Durlauf S.N., Johnson P.A., Temples J.R. (2005), Fritsche U., Kuzin V. (2011), Hausmann R., Pritchett L., Rodrik D. (2005) and Johnson P., Papageorgiou C. (2017), to name a few, provide precious knowledge on this matter.

2.2 The Growth Puzzle

Before proceeding to the analysis of the S.N.Durlauf and P.A.Johnson model and the empirical approach of this thesis, it is beneficial to discuss how countries can influence their economic growth. As it is become clear from literature[11] and empirical experience, there are lot of things that may hold back or slow down the growth process. Every country has its own specific features- or in some cases a group of countries share some common features-

[9] The bootstrap is a way to properly ''manipulate'' the big number of cross-sectional units within a short time frame.
[10] Corrado L., Stengos T., Weeks M., Yagzan M. (2018),
[11] Information extracted from *''Richard S.J. Tol (2009), Harvanek Tomas, Horvath Roman, Zeynalov Ayaz (2016), Siebert Horst, Stolpe Michael (2001), Corden W.M. (1984).*

which determine the growth course in the short- and in the long-run. However, not all countries experience growth and the countries that are currently growing might experience shrinkage periods. Looking back to history, one may understand fairly easily that economic empires rise and fall depending on the conditions that co-exist at the specific time period. Wars, pandemics, lack of social and political stability, existence or not of natural resources, corruption, investments in physical and human capital along with technology, play a key role on a country's growth process. Research has shown that technology has the greatest involvement among these factors, but it is unwise to focus just on the latter. Since a country cannot keep growing based exclusively to capital and labor accumulation due to diminishing returns, there is a huge need to find alternative paths so that the growth can keep occurring. For instance, better technology in crucial sectors of economies such as manufacturing, natural resources extraction and logistics (or other services in general) should make the production faster and easier. In addition to all of that, the costs of production and transportation would fall significantly, making the goods and services cheaper and more accessible to the population. This very specific economy would experience an increase in its exports -due to the lower price which in its turn is providing the country with a comparative advantage- and eventually it would bring an increase in their GDP. For that it is, obviously, necessary that it is an open economy without imposing quotas, high customs or any other factors that can intervene negatively in the international trade. The example of Germany simply makes it clearer. Despite the almost complete destruction of its infrastructure[12] during the two World Wars, the country managed to recover enormously -due to the massive technological boost- and thus be the leader among its counterparts in European Union in the modern world. Regarding the natural resources, a country needs to be very careful how to handle this situation. A subsoil rich in natural resources may affect the economic growth in a positive way. Yet, it is not always possible to properly extract and store all this wealth, and this is why many countries leave all these

[12] Germany not only have had its infrastructure destroyed, but also its resources almost completely consumed. The huge needs during war preparations as well as during the fighting period, forced the German Empire to seek resources mostly in Caucasus, Middle-East and North Africa. Two battles are considered to have drained out the German resources: the ''Normandy Landings'' and the ''Battle of the Bulge''.

resources untouched. Furthermore, should a country make of game-changer discovery of oil or natural gas fields, this does not guarantee that its economy will thrive in the years to come. The danger of the so-called ''Dutch disease''[13] is lurking and it might hold back the economy as a whole in favor of only one sector. One should not forget that political stability associated with extremely low rates of corruption, could contribute in a crucial way in economic growth. Companies and funds are more likely to invest in a country where institutions function with transparency, laws about copyrights are being respected and local authorities give motives to potential investors. Therefore, it is more than useful to promote ideas for a healthy social-political sector. Last but not least, we should take into consideration the climate change and how this affects in various ways the production of a country. The extreme production of carbon dioxide emissions around the globe has led eventually in overheating the planet and destroy its ecosystems. It is due to raised temperatures that extreme weather phenomena appear such as floods, wildfires and many more. All these destroy -most of the times- the infrastructure of factories, companies, the crops that produce raw materials for other products. One may easily understand that the effort of governments and authorities to boost growth is a very tricky work to do, due to the multicomplex riddle of combining many factors together.

2.3 S.N. Durlauf and P.A. Johnson framework analysis

The framework of Durlauf and Johnson is based mostly on the M-R-W (1992) analysis. The total output of a country is being extracted from a Cobb-Douglas production function:

$$Y_{i,t} = \varphi K_{i,t}^{\alpha} H_{i,t}^{\gamma} (A_t L_{i,t})^{1-\alpha-\gamma}$$

where:

- Y_t is the aggregate output
- A_t is the level of technology at time t

[13] The economic magazine ''The Economist'' is the ''mastermind'' behind this term, while trying to understand and explain the fall of the manufacturing sector in the Netherlands after the discovery of the huge natural gas field in the Chroningen subsoil.

- $L_{i,t}$ is the labour input
- $K_{i,t}$ is the physical capital input
- $H_{i,t}$ is the human capital input

Technology grows at a constant rate *g* and labor grows at constant rate *n*. Physical and human capital accumulate at constant rates s_i^k and s_i^h respectively. They both depreciate at the same rate δ which is estimated to 0.03. Therefore, physical and human capital accumulation can be explained using the following equations:

- $\frac{dK_{i,t}}{dt} = s_i^k Y_{i,t} - \delta K_{i,t}$
- $\frac{dH_{i,t}}{dt} = s_i^h Y_{i,t} - \delta H_{i,t}$

From *T* to *T+t* time period the variable $Y/L_{i,t}$ is as follows:

$$\ln(\frac{Y}{L})_{i,T+\tau} - \ln\left(\frac{Y}{L}\right)_{i,T} = g\tau + (1 - e^{-\lambda\tau})x[\Theta + \frac{\alpha}{1-\alpha-\gamma}\ln(s)_i^k + \frac{\gamma}{1-\alpha-\gamma}ln(s)_i^h - \frac{\alpha+\gamma}{1-\alpha-\gamma}\ln(n+g+\delta) - \ln\left(\frac{Y}{L}\right)_{i,T}]$$

Where:

- $\Theta = \frac{1}{(1-\alpha-\gamma)\ln(\varphi)} - \ln(A)_0 - gT$
- $\lambda_i = (1 - \alpha - \gamma)(n_i + g + \delta)$, which is the convergence rate

This specific equation explains cross-country growth behavior while countries share technology. This equation is referred to as the ''constrained Solow equation'' due to the non-linear restrictions it places on its parameters. Assuming that λ remains constant, we can create the ''unconstrained Solow equation'':

$$ln(\frac{Y}{L})_{i,T+\tau} - ln\left(\frac{Y}{L}\right)_{i,T} = \zeta + \beta ln\left(\frac{Y}{L}\right)_{i,T} + \Pi X_i + \varepsilon_i \qquad i = 1,2,\dots,N$$

where: $X_i = [\ln(s)_i^k, ln(s)_i^h, \ln(n + g + \delta)]$

If $-\frac{(1-e^{\lambda,\tau})(\alpha+\gamma)}{1-\alpha-\gamma} < 0 \; or \; \beta < 0$, in the restrained equation or in the unconstrained respectively, it means that convergence is occurring. If those quantities are negative and far from zero, convergence takes place with high rates. Therefore, the gap between countries is closing with faster pace. However, the authors follow the reasoning of Azariades and Drazen (1990) in which they claim the existence of human and/or physical capital thresholds. Practically, this means that after a certain break point the indicated production function transforms, so, there are different steady states for different initial conditions. The production function under this point of view, gets transformed compared to the previous one:

$$Y_{i,t} = \varphi(K)_{i,t}^{a_j}(H)_{i,t}^{\gamma_j}(A_t L_{i,t})^{1-a_j-\gamma_j}$$

What is special about this production function is that quantities *α and γ* are set to change if physical and human capital, respectively, surpass the threshold. Specifically:

1. Physical capital threshold is set as $\bar{K}(t)$
2. Human capital threshold is set as $\bar{H}(t)$

If $K_{i,t} < \bar{K}(t)$➔ $a_j = a_1$, otherwise if $K_{i,t} > \bar{K}(t)$ -which means that the physical capital ''crossed'' the threshold point- then $a_j = a_2$

If $H_{i,t} < \bar{H}(t)$➔ $\gamma_j = \gamma_1$, otherwise if $H_{i,t} > \bar{H}(t)$ -which means that the human capital surpassed the threshold point- then $\gamma_j = \gamma_2$

The constrained equation of the Solow model, takes a new form such as:

$$ln\left(\frac{Y}{L}\right)_{i,T+\tau} - ln\left(\frac{Y}{L}\right)_{i,T} = g\tau + (1 - e^{-\lambda_{i,\tau}})x\left[\Theta_i + \frac{a_i}{1 - a_i - \gamma_i}\ln(s_i^k) + \frac{\gamma_i}{1 - a_i - \gamma_i}\ln(s_i^h) - \frac{a_j - \gamma_j}{1 - \alpha_j - \gamma_j}\ln(n_i + g + \delta) - ln\left(\frac{Y}{L}\right)_{i,T}\right]$$

Therefore, it all depends on the evolution of the physical and human capital variables.

Proceeding with the framework empirical analysis, data have been extracted from the Summers-Heston (1988) dataset. The Penn World Table, as it is commonly referred to, is created to provide information about national accounts over a quite large number of countries spanning from 1950 to 2019[14]. To be exact, it covers 183 countries and includes data mostly about productivity, population, capital and employment, allowing this way researchers to make comparisons over the standard of living among countries. Furthermore, the authors are using the Barro-Lee (2021) table which includes precious information about educational attainment over the population. The data in this table are being calculated every five years in comparison with the PWT data that were calculated on yearly basis. Furthermore, World Data Bank provides data for the working-age population enrolled into secondary school. Having all the required data already available, in the next step the variables are constructed. Namely:

- $(\frac{Y}{L})_{i,t}$, which is the real GDP per capita (population aged 15-64), country *i* at time *t*
- $(\frac{I}{Y})_{i}$, which is the fraction of real GDP engaged in investments for country *i*, average for years 1960-1985
- n_i, which is the growth rate of the working-age population for country ***i,*** average for years 1960-1985
- $SCHOOL_i$, which is the fraction of the working-age population enrolled in a secondary school for country ***i***, average for years 1960-1985
- $LR_{i,1960}$, which is the adult literacy rate for country *i*, in 1960

At this point, it must be clarified that for some countries the researchers are using the literacy rate of the year 1975 because there are no available data for 1960. Countries exhibiting literacy rate over 90% it is unlikely to have biased effects during the calculations. Furthermore, Botswana and Mauritius are excluded because there are no data at all. It is being assumed that *g*=0.02 and δ=0.03**.**

[14] Not all countries have available data especially for the 1950-1959 decade.

The goal of the following process is to identify whether the data exhibit multiple regimes. To do so, it is required to split of the total sample into subsamples. The split is based on two control variables. The first control variable is the per capita output in 1960 [$(\frac{Y}{L})_{i,1960}$]. The second one is the adult literacy rate in 1960 [$(LR_{i,1960})$]. The use of literacy as a control variable ''*makes sense if one thinks of the potential regimes in the data as stemming from differences in the level of social and economic development rather than the current level of economic activity''*[15] [S.N Durlauf and P.A Johnson (1995),p.369]. In other words, these two variables are able to detect ''*unobserved physical and human capital stocks''*[16] [S.N Durlauf and P.A Johnson (1995),p.369].

After splitting is completed, specification tests will reveal whether there are multiple regimes in the data. In other words, it is needed to test whether the parameters of the model exhibit equality across subgroups. Therefore, the hypothesis test is as such:

- $H_0: Single\ regime\ model\ (\ equality\ of\ parameters\ )$
- $H_1: Multiple\ regime\ model\ (\ inequality\ of\ parameters\ )$

The authors test for the equality of parameters carrying out multiple Wald tests, covering all data splits. Specifically, splits using initial output and literacy as control variables are:

1. Two groups (low group and high group) of 48 members for each one of the control variables (Two-way split).
2. Three groups (low, intermediate and high group) of 32 members for each one of the control variables (Three-way split).
3. Four groups (high-output/high-literacy, high-output/low-literacy, low-output/high-literacy, low-output/low-literacy) when variables interact with each other (Four-way split). The high-output/low-literacy and low-output/high-literacy are excluded from the analysis due to the small number (6) of observations. The remaining groups include 42 members each.

[15] [S.N Durlauf and P.A Johnson (1995),p.369]
[16] [S.N Durlauf and P.A Johnson (1995),p.369]

With the use of Wald tests, the significance level of the null hypothesis that all parameters are equal across the subgroups, is being examined. Before running the Wald tests, heteroscedasticity tests exhibit presence of heteroscedasticity. In the following tests, homoscedasticity is being assumed for the outcome to be more accurate and the results are being presented in the table 1:

Splits and subgroups	*Unconstrained regressions*	*Constrained regressions*
Two-way split for: $\left(\frac{Y}{L}\right)_{i,1960}$ $LR_{i,1960}$	0.009 0.011	0.218 0.112
Three-way split for: $(\frac{Y}{L})_{i,1960}$ $LR_{i,1960}$	0.029 0.404	0.011 0.000
Four-way split both for: $LR_{i,1960}$ *and* $(\frac{Y}{L})_{i,1960}$	0.000	0.000

Table 1: Marginal significance levels of Wald tests testing whether model parameters remain the same across subgroups .

Source: Table I page 369 from Durlauf and Johnson (1995).

Results are quite interesting and important. Three out of four splits regarding the initial output variable led to the rejection of the null hypothesis that the parameters are equal across the indicated subgroups. The hypothesis is rejected when significance level is 3%. Regarding the literacy rate control variable, the null hypothesis is rejected in two out of four tests, when significance level is 1%. When output and literacy interact, there is absolutely zero evidence of presence of a single regime model. Alongside with these results, it is wise to make a reference for the break points of the control variables. Regarding the two-way split, the output breaks at 1950$ while the literacy rate breaks at 54%. Regarding the three-way split, the output includes the low group ($(\frac{Y}{L})_{i,1960}$<1150$), the intermediate group (1150$ ≤$(\frac{Y}{L})_{i,1960}$≤2750$) and the high group ($(\frac{Y}{L})_{i,1960} > 2750\$$). As for the literacy rate, it includes the low group ($LR_{i,1960}$<26%), the intermediate group (26%≤$LR_{i,1960}$≤72%) and the high group ($LR_{i,1960}$>72%).

From all the groups that are created, the most ''attractive'' ones are those when the two control variables are interacting. Therefore, Durlauf and Johnson (1995) ran OLS for these two groups as well as for the sample that MRW analysis used. The table 2 presents the results of OLS in these three cases:

Unconstrained regressions	*M-R-W*	$(\frac{Y}{L})_{i,1960}$*<1950$ and* $LR_{i,1960}$*<54%*	$(\frac{Y}{L})_{i,1960}$ *≥1950$ and* $LR_{i,1960}$ *≥54%*
Observations	98	42	42
Constant	3.04**	1.40	0.450
	(0.831)	(1.85)	(0.723)
$ln(\frac{Y}{L})_{i,1960}$	-0.289**	-0.444**	-0.434**
	(0.062)	(0.157)	(0.085)
$ln(\frac{I}{Y})_i$	0.521**	0.310**	0.689**
	(0.087)	(0.114)	(0.170)
$ln(n+g+\delta)_i$	-0.505	-0.379	-0.545
	(0.288)	(0.468)	(0.283)
$ln(SCHOOL)_i$	0.233**	0.209**	0.114
	(0.060)	(0.094)	(0.164)
$\bar{R}^2$	0.46	0.27	0.48
σ_t	0.33	0.34	0.30

$Table\ 2: OLS\ results\ for\ unconstrained\ regressions\ in\ M-R-W\ and\ sample\ splits$

Source: Table II, page 370 from Durlauf and Johnson (1995).

It is quite noticeable that amounts differ amongst unconstrained regressions, however the numerical indications exhibit a ''common'' way of affecting the dependent variable. For instance, the coefficients of the initial output are almost the same in the two samples produced by the split, but they are much higher in absolute numbers (-0.444 for the lower group and -0.434 for the higher group) than the coefficient in the M-R-W regression (-

0.289). However, these three numbers are affecting negatively the dependent variable and they are all statistically significant when α=5%. Something special about these numbers is that they reveal faster any convergence within the subgroups rather than in the whole sample due to a bigger gap in the dependent variable $\mathbf{ln(\frac{Y}{L})_{i,1985}}$-$\mathbf{ln(\frac{Y}{L})_{i,1960}}$. Examining the coefficients regarding the fraction of the investment devoted to real GDP, one may notice that they differ in an important way. The coefficient in the high group is more than a double (0.689) compared to the low group (0.310). The coefficient of the M-R-W analysis is between the numerical values of the low and high group (0.521). Surprisingly, none of the coefficients of the $\ln(n + g + \delta)_\iota$ is statistically significant, thus the variable shouldn't count as useful for the regression. Also, the $\ln(SCHOOL)_i$ coefficients of the M-R-W and low group analysis are statistically significant at α=5% and almost equal but the variable seems not to be useful for the high group analysis, where is not statistically significant. Knowing that the $\bar{R}^2$ is a very helpful instrument to measure the effectiveness of the explanatory variables in the regression, it seems that for the M-R-W group as well as for the high group, the adjusted-R squared is almost at 50%. In other words, almost the fifty percent of the variability of the dependent variable is explained collectively by the explanatory variables. In the low group, the $\bar{R}^2$ is a bit lower at 27% which might be not quite desirable. Simply put, the higher the $\bar{R}^2$ is the better the model is.

Furthermore, the results for the constrained regressions exhibit different ''paths'' for the groups.

Constrained regressions	*M-R-W*	$(\frac{Y}{L})_{i,1960}$<1950$ *and* $LR_{i,1960}$<54%	$(\frac{Y}{L})_{i,1960}$ ≥1950$ *and* $LR_{i,1960}$ ≥54%
Observations	98	42	42
Θ	-2.56**	2.29	-0.395
	(1.14)	(1.17)	(1.24)
α	0.431**	0.275**	0.509**
	(0.061)	(0.097)	(0.098)
γ	0.241**	0.217**	0.108
	(0.046)	(0.061)	(0.094)
$\bar{R}^2$	0.42	0.28	0.50
σ_t	0.34	0.34	0.29

$Table\ 3: OLS\ results\ for\ constrained\ regressions\ in\ M-R-W\ and\ sample\ splits$

Source: Table II, page 370 from Durlauf and Johnson (1995).

For instance, the physical capital share in GDP, the α variable, is extremely larger (0.509) for the high group than the low one (0.275) and a bit higher from the one estimated in the M-R-W analysis (0.431). However, the low group coefficients for Θ (human capital investment) and for γ (human capital share in GDP) are larger compared to those in high group. Thus, it is safe to say that there are different laws of motions between the subgroups. In other words, these results strengthen the belief that there are multiple regimes across the data do exist.

In response to the criticism claiming that the presence of multiple regimes in the data arises from important variables that are omitted in the analysis, authors add a bigger number of variables in the regression, that can affect directly or indirectly the growth process. These specific variables arise from the research of Barro (1991) and cover quite completely the factors than are involved in growth. These variables are:

1. $AFRICA_i$, a dummy variable that gets 1 if the country is in sub-Saharan Africa[17]
2. $ASSASS_i$, number of assassinations per million per year
3. $(\frac{G^c}{Y})_i$, ratio of government consumption -which includes spending on defense and education- to GDP
4. $LATAMER_i$, a dummy variable that gets 1 if the country is in Latin America[18]
5. $LIT60_i$ the adult literacy rate in 1960. This variable differs from the $LR_{i,1960}$
6. $MIXED_i$, that gets 1 if the country has a combination of free/centrally planned economy system
7. $PPI60DEV_i$, deviation from sample mean of the 1960 purchasing power parity
8. $PRIM60_i$, the 1950 primary school enrollment rate
9. REV_i , the number of revolutions and coups per year
10. $SEC60_i$, the 1960 secondary school enrollment rate
11. SOC_i , gets 1 if the country has pure centrally planned economy
12. $STTEAPRI_i$, the 1960 primary school student-teacher ratio
13. $STTEASEC_i$, the 1960 secondary school student-teacher ratio

These variables may play a key role in the growth process but they might as well not reflect their real impact. This happens because in many countries – especially the African countries - no data have been collected for decades in all these sectors. Despite this problem, the variables still remain the most representative indicators for the analysis. Adding them in the constrained and unconstrained equations and testing again the significance level, the results are consistent with the ones produced without the extra variables.

To sum up, specification tests exhibit the existence of multiple regimes. The challenge now is to spot those economies with common laws of motion and to group

[17] As sub-Saharan countries are stated those countries that are fully or partly situated in the south part of Sahara. The map of the sub-Saharan countries is listed in the Appendix. The term separates these countries from the north Africa, also known as the Arab World.

[18] Latin America includes all countries below Mexico and the Caribbean states, with the exception of the Guyana, Suriname and French Guyana.

them together. Splitting the data exogenously isn't helpful. Therefore, there is a need to carry out this process endogenously. S.N. Durlauf and P.A. Johnson use a fairly complicated algorithm which has first been established by Breiman (1984), that is the regression tree analysis. Since there is no prior theory suggesting how many regimes there are, this algorithm allows to search for an infinite number of regimes. The procedure reveals four groups of countries that have common laws of motion. The indicated groups are being presented in the Appendix. It is noteworthy that these four groups have a common characteristic: geographical homogeneity. Tables 4 and 5 present the OLS estimations for each one of these groups, for unconstrained and constrained regressions respectively:

Unconstrained regressions	*Group 1 (Terminal node 1)*	*Group 2 (Terminal node 2)*	*Group 3 (Terminal node 3)*	*Group 4 (Terminal node 4)*
Observations	14	34	27	21
$Constant$	3.46	-0.915	0.277	-7.26**
	(2.27)	(1.79)	(1.42)	(1.59)
$ln(\frac{Y}{L})_{i,1960}$	-0.791**	-0.086	-0.316**	0.069
	(0.269)	(0.131)	(0.123)	(0.139)
$ln(\frac{I}{Y})_i$	0.314**	0.129	1.110**	0.475**
	(0.109)	(0.159)	(0.165)	(0.119)
$ln(n+g+\delta)_i$	-0.429	-0.390	0.059	-1.75**
	(0.678)	(0.489)	(0.451)	(0.270)
$ln(SCHOOL)_i$	-0.028	0.469**	-0.114	0.341**
	(0.073)	(0.095)	(0.167)	(0.141)
$\overline{R^2}$	0.57	0.52	0.57	0.82
σ_t	0.16	0.28	0.28	0.12

$Table\ 4: OLS\ estimations\ for\ the\ endogenously-created\ terminal\ groups-Unconstrained\ regressions$

Source: Table V, page 375 from Durlauf and Johnson (1995).

One may notice heterogeneity among the variables. Coefficients vary significantly along with the numerical indications. Notwithstanding this variation, there is an improvement over the fit compared to the single regime model. All adjusted R-squared are higher than the 46% that the M-R-W analysis produces. That means that the independent variables explain more of the variation of the dependent variable. In addition, the coefficients of the $\ln(\frac{Y}{L})_{i,1960}$ are greatly dissimilar, exhibiting convergence for the first and third group, while for the fourth group there is no convergence. Similar results are produced when examining constrained regressions.

Constrained regressions	*Group 1 (Terminal node 1)*	*Group 2 (Terminal node 2)*	*Group 3 (Terminal node 3)*	*Group 4 (Terminal node 4)*
Observations	14	34	27	21
Θ	4.107**	0.539	-3.95	-11
	(0.552)	(1.809)	(2.67)	(7.64)
α	0.306**	0.186	0.758**	0.333**
	(0.083)	(0.123)	(0.095)	(0.100)
γ	-0.034	0.416**	-0.073	0.455**
	(0.083)	(0.080)	(0.114)	(0.103)
$\overline{R^2}$	0.64	0.40	0.55	0.71
σ_t	0.19	0.32	0.30	0.18

$Table\ 5: OLS\ estimations\ for\ the\ endogenously-created\ terminal\ groups-Constrained\ regressions$

Source: Table V, page 375 from Durlauf and Johnson (1995).

The great gap regarding the physical capital share, according to the authors, is explained by the different technologies that the economies have access to. Gap is also spotted when examining the human capital share. The source of this difference might be the $\ln(SCHOOL)_i$ variable which only measures the secondary school enrollment. Primary and college education might also add something important when identifying regimes, therefore these factors should be taken into account.

Summarizing, the effort of S.N. Durlauf and P.A. Johnson led to three basic results. First, they reject the null hypothesis of the single regime model with and without the additional variables proposed by Barro (1991). Secondly, they used the regression tree analysis group for economies that exhibit common laws of motion. That indicates that economies have different technological levels. Finally, their results cannot verify the presence or absence of any convergence.

3. Data, Variables and Methodology

To fulfill the aim of this study, which is to reassess the unconstrained equations analyzed in S.N Durlauf and P.A. Johnson framework, we derive the required data from various databases. The main databases include the latest version of Penn World Table[19], the World Data Bank and the Barro-Lee table[20]. A few other databases such as FRED database and SIPRI Military Expenditure Database[21] played a supportive role when the main databases didn't store data regarding some categories. To be more exact, we need data for the following variables:

- *Y,* real GDP per country, annual averaged
- *L,* population aged 15-64 per country (labor force), annual averaged
- *I,* variable of investment (including government investment), annual averaged
- n_i, growth rate of working age population, annual averaged
- $SCHOOL_i$, fraction of the working-age population enrolled in secondary school, annual averaged
- $LR_{i,1960}$, control variable that is being used for breaking the sample

Additionally, we need data for the extra set of variables:

[19] The Penn World Table's data are being updated and maintained by two universities: Davis University of California and the University of Chroningen. The latest update of the PWT took place on 18th of June, 2021.
[20] Latest update of the Barro-Lee table took place at September 2021.
[21] The SIPRI (Stockholm International Peace Research Institute) Database is an independent organization who collects data regarding military from various sources, such as governments and authorities, International Monetary Fund statistics, NATO annual reports for expenditures on military defense and various information from papers and journalists.

- G^c, government expenditure on defence and education, annual averaged
- $PRIM60_i$, primary school enrollment rate at 1960
- $STTEAPRI_i$, primary school student-teacher ratio at 1960
- $STTEASEC_i$, secondary school student-teacher ratio at 1960

The AFRICA and LATAMER are dummy variables and can be created manually within the econometric program.

Having multiplied the Y (stated as *rgdpna* in PWT) and I (constructed using *rnna* from PWT) with 1.000.000 so the variables are homogenous, we proceed on constructing the variables of the model -using the Eviews Econometric Program- as follows:

- $\ln(\frac{Y}{L})_{i,1960}$: For every country within the sample, we divide the 1960 values of output and population and then we logarithm the new quantity.
- $\ln(\frac{Y}{L})_{i,j}$: For every country within the sample, we divide the *j* values of output and population and then we logarithm the new quantity, whereas *j=1985 or j=2015*
- $\ln(\frac{I}{Y})_i$: To create this variable, we first need to construct separately the *I* following these steps:
    - As mentioned before, to construct the term *I* we make use of the *rnna* which we set it equal with K.
    - $K_t = K_{t-1} * (1 - \delta_{t-1}) + I_{t-1} \rightarrow I_{t-1} = K_t - K_{t-1} * (1 - \delta_{t-1})$
    - Afterwards, for every country we extract the average for output and investment and we divide the two variables. Finally, we logarithm the new quantity.
- $\ln(n_i + g + \delta)$: We calculate the n_i for every country and we extract the average quantity. It is clear from literature that *g=0.02 and δ=0.03.* In our case the PWT provides numerical values for the delta variable for each year. Thus, we summing up these values and we logarithm the new quantity.

- $\ln(SCHOOL)_i$: Extracting the numerical values from the Barro-Lee table regarding the working age population enrolled in secondary school, we calculate the average for every country and then we calculate the new quantity.

We first focus on the 1960-1985 time period, as the paper we based on does. For this time period, we use the as sample the countries referred as the Non-Oil in M-R-W (1992). In addition to that, we expand the time spectrum to 1960-2015 using again the same sample as before. Finally, for the last part of the empirical approach and for 1960–2015-time spectrum we expand the sample of countries, adding up these countries that data are available for the variables described above.

1. The 1960-1985 sample includes 96 countries (Botswana and Mauritius are excluded). However, these two countries are needed when running the M-R-W sample.
2. The 1960-2015 analysis includes two versions of sample. First analyzing the 96 countries and then adding 12 more countries for the sample to become 108.

As a first step, we need to split the sample the same way S.N. Durlauf and P.A. Johnson do in their paper which is the base of this study. For this process, we insert into separate Excel files the countries along with their initial values of output and literacy rate. Afterwards, we rank the values in ascending order and we break the sample in two equal groups, where each subgroup consists of 48 countries, for both of the control variables. The same process is being followed for the three-way split, creating three subgroups of 32 countries, for each one of the control variables. Finally, we let output and literacy to interact with each other. The outcome includes four subgroups. From these four groups we focus on high-output/high-literacy and on low-output/low-literacy. This pattern is being followed for every time spectrum we examining in this study. In every case (two-way split, three-way split, four-way split), we combine the indicated subgroups into a single Excel file so we can make the comparisons through the Eviews program. The rest of the process includes the OLS estimation of each equation, conduct of heteroscedasticity tests and correction of it where it exists. Afterwards, we carry out Wald specification tests to examine whether the model parameters are equal between the subgroups. Finally, we add the extra variables and we repeat the process. The results are being presented with details in the next section.

4. Empirical results

4.1 Results for 1960-1985 and 96 countries

In this section we try to estimate from scratch the unconstrained equations for each one of the samples for years 1960 to 1985. Table 6 presents the OLS estimates for the control variables in the two-way split.

Dependent variable: $\ln(Y/L)_{I,1985} - \ln(Y/L)_{I,1960}$

Method: OLS	Low group (LR_{1960})	High Group (LR_{1960})	Low group (Y/L_{1960})	High Group (Y/L_{1960})
Constant	3.45**	1.47	3.36**	1.58
	(1.20)	(2.54)	(1.57)	(1.05)
$\ln(Y/L)_{I,1960}$	-0.17	-0.85**	-0.28**	-0.07
	(0.10)	(0.03)	(0.12)	(0.11)
$\ln(I/Y)_i$	0.06	-0.24	-0.002	0.32**
	(0.14)	(0.27)	(0.13)	(0.11)
$\ln(n+g+\delta)_i$	0.37	-2.71**	0.03	-0.13
	(0.21)	(0.67)	(0.24)	(0.33)
$\ln(School)_i$	0.22	0.41**	0.28**	0.07
	(0.11)	(0.16)	(0.10)	(0.09)
Adjusted R^2	0.10	0.96	0.14	0.21
Observations	48	48	48	48

$Table\ 6: Two-way\ splits\ based\ on\ LR\ and\ \frac{Y}{L}\ 1960\ initial\ values$

The sample breaks at 46.15% when we take into account the literacy rate variable and at 5777.33$ when we split the sample according to initial output. From table 6, one may understand that the constant is relatively high and statistically significant in α=5% only for the low group of each control variable. The coefficient of the initial output is negative in all cases but only significant in the high group when literacy is the control variable and in the low group when output is the control variable. The negative coefficient is a sign that convergence may occur in these subgroups. Furthermore, the estimates for the $\ln(\frac{I}{Y})_i$, are almost zero for both of the low groups and negative- but not significant in any of significance levels- for the high group. Only for the group with high initial output the coefficient is statistically significant. This groups consists of mostly rich and developments economies, which are investing a lot in physical capital. As for the $\ln(n+g+\delta)_i$ coefficients, only the estimate of the high group of literacy is significant and negative. Last but not least, all of the coefficients in the variable that reflects the enrollment in the secondary school are all positive. One may expect this variable to play a positive role in the growth, however it is only significant in two out of four groups. Also, the adjusted-R squared gives information about how good the model is. Only for the high group in literacy sample breaks the explanatory variables seem to explain almost completely the variation of the dependent variable. For the other cases, the adjusted-R squared is quite low. Moving on to table 7, we examine the results when we break the sample using the literacy rate variable into three equal subgroups.

Dependent variable: $\ln(Y/L)_{I,1985} - \ln(Y/L)_{I,1960}$

Method: OLS	*Group A (LR split) (Low)*	*Group B (LR split) (Intermediate)*	*Group C (LR split) (High)*
Constant	4.04**	-0.66	2.96**
	(1.42)	(2.45)	(1.36)
$ln(Y/L)_{I,1960}$	-0.24**	0.03	-0.27**
	(0.12)	(0.12)	(0.08)
$ln(I/Y)_i$	0.04	0.12	0.37**
	(0.10)	(0.18)	(0.17)
$ln(n+g+\delta)_i$	0.25	-0.36	-0.41
	(0.18)	(0.83)	(0.43)
$ln(School)_i$	0.35**	-0.05	0.23
	(0.10)	(0.22)	(0.14)
Adjusted R^2	0.30	0	0.27
Observations	32	32	32

Table 7: Three – way split based on Literacy Rate

From these three groups, the intermediate group seems to be completely meaningless since none of the variables is significant. As for the low and high group, the constant is relatively high and significant to the analysis. The coefficients of the initial output variable are negative, significant and almost equal for the low and high group. The $\ln(\frac{I}{Y})_i$ seems to be meaningful only for the high group, whilst the population growth isn't playing a key role in any of the subgroups. Lastly, the variable of school is significant only for the low group. Obviously, the adjusted-R squared is zero for the intermediate group and around 0.30 for the other two groups. The groups exhibit break points at 21.85% (low-intermediate) and at 70.25% (intermediate-high). We then, run the same process splitting the sample based on the initial output. Table 8 exhibits the results on this matter:

Dependent variable: $\ln(Y/L)_{I,1985} - \ln(Y/L)_{I,1960}$

Method: OLS	*Group A (y split) (Low)*	*Group B (y split) (Intermediate)*	*Group C (y split) (High)*
Constant	1.52	1.46	0.94
	(3.54)	(3.89)	(1.33)
$\ln(Y/L)_{I,1960}$	-0.19	-0.03	-0.12
	(0.20)	(0.40)	(0.13)
$\ln(I/Y)_i$	0.07	0.09	0.22**
	(0.17)	(0.15)	(0.09)
$\ln(n+g+\delta)_i$	-0.45	0.02	-0.46
	(0.86)	(0.37)	(0.32)
$\ln(School)_i$	0.25**	0.27**	0.03
	(0.11)	(0.13)	(0.11)
Adjusted R^2	0.14	0.006	0.16
Observations	32	32	32

$Table\ 8: Three-way\ split\ based\ on \frac{Y}{L}\ (denoted\ as\ y)$

In this case we split the sample between low and intermediate when initial output is at 3586.22$ and between intermediate and high when the control variable is 9793$. Diving into the results, it seems that the three-way split in this case doesn't produce any fruitful results. The coefficient of investment share in GDP is positive and significant only in the high group. Also, the $\ln(n+g+\delta)_i$ coefficients are almost equal and meaningful in the first two groups. None of the other coefficients are helpful for any of the groups. Hence, the extreme low $\bar{R}^2$ rates.

After having going into the sample splits for each one of the control variables (estimating only the parameters of the basic variables of the model), it is needed to estimate the M-R-W sample as well as the splits when the variables interact. Therefore, we can compare with the research that is our base for this study. The comparison relies only to the unconstrained regressions, as mentioned before.

Dependent variable: $\ln(Y/L)_{I,1985} - \ln(Y/L)_{I,1960}$

Method: OLS	*M-R-W*	*Low Group (LR and Y/L interaction)*	*High Group (LR and Y/L interaction)*
Constant	1.73	4.36**	2.05
	(0.97)	(1.18)	(1.56)
$\ln(Y/L)_{I,1960}$	-0.16*	-0.39**	-0.16
	(0.08)	(0.10)	(0.12)
$\ln(I/Y)_i$	0.08	0.02	0.39**
	(0.08)	(0.12)	(0.14)
$\ln(n+g+\delta)_i$	-0.30	0.02	-0.35
	(0.24)	(0.28)	(0.37)
$\ln(School)_i$	0.26***	0.31**	0.18
	(0.08)	(0.08)	(0.12)
Adjusted R^2	0.16	0.13	0.27
Observations	98	42	42

Table 9: Results following M − R − W approach and 2 − way split for both variables

The constant quantities are all positive, but only the constant for the second group is statistically significant. Comparing with the research on which we based on, our estimations for the constant terms, seem to be a bit higher but the numerical indications remain the same. S.N Durlauf and P.A. Johnson (1995) find that the constant term of M-R-W is significant, whilst we find the constant for the low group is. As for the $\ln(\frac{Y}{L})_{i,1960}$ coefficients the authors find that all of them are meaningful for the analysis, whilst we find that the M-R-W and low group coefficients are important. In both analyses the numerical sings are negative in all terms. Proceeding to the investment share in GDP coefficients, the estimations in our analysis are quite lower than those in the base study. We find that only the coefficient in the high group is significant while the base study exhibits that the three terms are significant. This is quite a difference between the two panels. Regarding the estimations on the working-age population growth variable, both panels find that none of the coefficients are important for the analysis. Lastly, for the $\ln(SCHOOL)_i$, we find that the coefficient in the M-R-W estimation is significant when α=10% whilst in the base study it is significant when α=5%. For the low group both estimations are significant in significance level of 5%. Regarding the high group, both panels find that the numerical quantity isn't meaningful. Furthermore, the estimations seem to be almost the same in both panels.

Moving past the OLS estimations, we need to dig into the Wald specification tests. Through Wald tests we will examine whether the data exhibit multiple regimes. As stated before, the null hypothesis claims that all parameters are equal across the subgroups. Hence, this means no multiple regimes existence. Panel 10 presents our calculations on this matter:

Subsamples	*Unconstrained Regressions*
Two-way split based on $(Y/L)_{I,1960}$ $LR_{i,1960}$	 0.1014 0.0000
Three-way split based on $(Y/L)_{I,1960}$ $LR_{i,1960}$	 0.0954 0.4555
Four-way split based on both $(Y/L)_{I,1960}$ and $LR_{i,1960}$	0.0453

Table 10: *Significance level of Wald tests regarding the null hypothesis that the parameters are constant across the subgroups*

Our calculations show that for the output control variable in the two-way split we don't reject the null hypothesis, due to the significance level being higher than 5%. Therefore, there are no evidence of multiple regimes existence which is a different outcome of the one that the authors produce. In contradiction, the same calculations regarding the literacy rate variable exhibit presence of multiple regimes, which is compatible with what the main research finds. As for the three-way split section we find that for both of the variables, the null hypothesis isn't rejected. The difference with the main study is that we reject the multiple regimes hypothesis, whilst the initial outcome rejects it only for the literacy rate. For the last section, when the variables affect each another our results are compatible with what the base estimation panels suggest, leading to the rejection of the null hypothesis.

The previous analysis was for the equations containing only the basic variables. To have a clearer view, we add an extra set of variables and running the empirical process once more.

Dependent variable: $\ln(Y/L)_{I,1985} - \ln(Y/L)_{I,1960}$

Method:OLS	*Low Group (LR_{1960})*	*High Group (LR_{1960})*	*Low Group (Y/L_{1960})*	*High Group (Y/L_{1960})*
Constant	2.95**	2.27	1.30	2.96
	(1.35)	(2.63)	(1.48)	(1.64)
$\ln(Y/L)_{I,1960}$	-0.35**	-0.83**	-0.24**	-0.18
	(0.08)	(0.05)	(0.11)	(0.14)
$\ln(I/Y)_i$	0.14	-0.32	-0.009	0.25**
	(0.09)	(0.35)	(0.10)	(0.08)
$\ln(n+g+\delta)_i$	-0.06	-2.44**	-0.23	-0.07
	(0.35)	(0.72)	(0.40)	(0.23)
$\ln(School)_i$	0.03	0.40	0.02	0.10
	(0.10)	(0.25)	(0.12)	(0.10)
AFRICA	-0.29**	-0.15	-0.36**	-0.30**
	(0.12)	(0.41)	(0.12)	(0.15)
LATAMER	0.07	-0.31	-0.31	-0.38**
	(0.25)	(0.35)	(0.34)	(0.13)
G	5.95**	0.52	6.84**	-0.29
	(1.44)	(4.30)	(1.41)	(1.23)
PRIM70	0.66**	-0.41	0.51**	-0.19
	(0.28)	(0.28)	(0.25)	(0.11)
STTEAPRI70	-0.44	0.15	-0.33	0.13
	(0.40)	(0.85)	(0.43)	(0.40)
STTEASEC70	0.49	0.28	0.32	0.41
	(0.61)	(1.20)	(0.52)	(0.86)
Adjusted R^2	0.47	0.96	0.45	0.45
Observations	48	48	48	48

$Table\ 11: Two-way\ split\ based\ on\ LR\ and\ \frac{Y}{L}\ initial\ values, adding\ extra\ variables\ as\ stated\ in\ Barro$ (1991)

From the table 11 is clear that the numerical indications of the coefficients are consistent with the theory. In some cases, though, such as the coefficients in literacy high group and low output group, the negative signs of the investment variable don't quite affect the analysis since there are not significant and, therefore, they can be excluded. Also, all of the coefficients of $\ln(n + g + \delta)_i$ are affecting negatively the dependent variable. However, only the parameter in high literacy rate group is significant. Interestingly, the AFRICA variable affects in all cases the dependent variable in a negative way. The coefficients don't differ much, quantitatively, and are significant in all cases except the high literacy group. One may understand that this dummy variable is dominating over the LATAMER one, providing a fixed factor from which we can drain information. The most important feature in this table is the variable of government consumption in defense/military and education. The coefficients are high and significant in the low groups whereas are low and insignificant in the high groups. The low groups are mostly consisting of countries that are located in areas that exhibit more or less war conflicts, terrorism and political instability. Most of these countries, invest much on military equipment, military wages and defense systems. However, there are a few cases of countries that lack of data on military spending. Obviously, this happens due to the need of keeping sensitive information secret. For these cases, we have replaced the military spending with spending on law enforcement. The coefficients of the PRIM70 variables are showing that are positive and significant for the low groups, but negative and insignificant for the high groups. Furthermore, all the coefficients of STTEAPRI70 and STTEASEC70 don't add something useful in the specific analysis. At this point, we need to clarify that the majority of the countries that are included in the initial sample lack of data regarding the student-teacher ratio in primary and secondary school for the decade 1960-1969. To overcome this problem, we use the values of 1970, when the data recording started occurring in a great scale. We are able to do so, because most of the countries exhibit -according to UNESCO statistics[22]- a constant pattern for the 1960–1969-time spectrum. Hence, we don't expect important differences between the 1960 and 1970 values. Lastly, the $\bar{R}^2$ seems to be relatively high and equal for the first, third and fourth group and extremely high for the second group. Thus, we notice significant improvement in the overall fit for the first, third and fourth group, compared to the calculations without the extra set of variables.

Improvement in the fit is also being noticed in the following table which shows the augmented estimations of the literacy rate three-way split. The improvement, of course, is being reflected on the higher adjusted-R squared values of each group.

Dependent variable: $\ln(Y/L)_{I,1985} - \ln(Y/L)_{I,1960}$

Method:OLS	*Group A (LR split) (Low)*	*Group B (LR split) (Intermediate)*	*Group C (LR split) (High)*
Constant	6.71**	0.61	2.57
	(1.69)	(1.71)	(1.36)
$\ln(Y/L)_{I,1960}$	-0.41**	0.02	-0.23**
	(0.07)	(0.11)	(0.09)
$\ln(I/Y)_i$	0.02	0.11	0.24
	(0.08)	(0.11)	(0.20)
$\ln(n+g+\delta)_i$	1	-0.14	-0.45
	(0.64)	(0.63)	(0.45)
$\ln(School)_i$	0.14	-0.11	0.24
	(0.07)	(0.18)	(0.24)
AFRICA	-0.05	-0.91**	0
	(0.13)	(0.18)	(0)
LATAMER	0.09	-0.53**	-0.12
	(0.24)	(0.16)	(0.31)
G	3.45	1.13	0.52
	(2.40)	(2.19)	(1.44)
PRIM70	0.98**	-0.35**	-0.29
	(0.20)	(0.16)	(0.26)
STTEAPRI70	-1.48**	0.08	-0.46
	(0.41)	(0.47)	(0.44)
STTEASEC70	-0.12	-1.17**	1.54**
	(0.55)	(0.59)	(0.48)
Adjusted R^2	0.58	0.44	0.41

[22] http://uis.unesco.org/?URL_ID=2867&URL_DO=DO_TOPIC&URL_SECTION=201

Observations	32	32	32

Table 12: Augmented three – way split based on Literacy Rate control variable

In table 12, we notice that convergence is occurring in the Group A and Group C due the negative and significant coefficients of the initial output variable. The same coefficient for Group B is almost zero and it doesn't affect the dependent variable at all. Surprisingly, the rest of the basic variables aren't useful for this case. AFRICA and LATAMER exhibit significant coefficients only for the Group B. In the case of AFRICA, we exclude the dummy variable from the Group C analysis, hence the zero value, due to the dummy variable trap. The Group C doesn't include any country from the sub-Saharan region. The results related to the PRIM70 variable, are exhibiting some contradiction. The parameter value for the first group is significant and close to 1, while the parameter value for the second group is -0.35. As for the STTEAPRI70, only the coefficient of Group A is statistically significant and below zero. A differentiation exists in the case of STTEASEC70 where the Group B exhibits a negative value while the Group C exhibits positive value.

Following the same analysis and breaking the sample into three subgroups when initial output is the control variable, we derive the results presented in table 13:

Dependent variable: $\ln(Y/L)_{I,1985} - \ln(Y/L)_{I,1960}$

Method:OLS	*Group A (y split) (Low)*	*Group B (y split) (Intermediate)*	*Group C (y split) (High)*
Constant	-0.75	0.39	3.86**
	(4.64)	(1.78)	(1.53)
$\ln(Y/L)_{I,1960}$	-0.26	-0.21	-0.29**
	(0.24)	(0.21)	(0.10)
$\ln(I/Y)_i$	0.13	0.18**	0.16
	(0.17)	(0.08)	(0.12)
$\ln(n+g+\delta)_i$	-1.14	-0.93**	-0.13
	(1.28)	(0.32)	(0.31)
$\ln(School)_i$	0.01	0.29	0.15
	(0.16)	(0.15)	(0.13)
AFRICA	-0.13	-0.48**	-0.49**
	(0.21)	(0.15)	(0.13)
LATAMER	-1.14**	-0.01	-0.38**
	(0.46)	(0.18)	(0.15)

G	9.41**	4.46**	-2.46
	(3.42)	(1.85)	(1.80)
PRIM70	0.76	-0.16	-0.05
	(0.46)	(0.22)	(0.13)
STTEAPRI70	-1.07	0.53	-0.65**
	(0.56)	(0.52)	(0.26)
STTEASEC70	0.69	1.23	1.33**
	(1.02)	(0.86)	(0.44)
Adjusted R^2	0.51	0.35	0.64
Observations	32	32	32

$Table\ 13: Augmented\ three-way\ split\ based\ on \frac{Y}{L} (denoted\ as\ y)\ control\ variable$

Similarly in this case, the overall fit of the model has improved in a sharp way compared with the one without the extra variables. This group of variables contributes for a better explanation of the variation of the dependent term of the equation. As expected, the initial output coefficients produce negative numerical signs, however only the last one is statistically significant. The numerical quantity is very close between the subgroups. Something worth mentioning in this table is that all the coefficients from AFRICA and LATAMER are exhibiting negative signs. The coefficients of those dummy variables seem to affect the high group in a negative way. On the other hand, government expenditures on defense and education are extreme high in the first group and about the half in the second group. The gap between G and the rest of the variables with regards to quantity, is quite high, not only in this table, but in general. The rest of the variables (PRIM70, STTEAPRI70, STTEASEC70) don’t affect the Group A and B at all, but they play significant role in the Group C.

As always, we are more interested on seeing how the variables react when we let the control variables interact with one another. Quantitatively, we don’t expect much variation in the parameters, but we expect to see differences in the number of the variables affecting each subgroup in the sense that some crucial variables may be excluded.

Dependent variable: $ln(Y/L)_{I,1985} - ln(Y/L)_{I,1960}$

Method:OLS	Low Group (LR-Y/L interaction)	High Group (LR-Y/L interaction)
Constant	2.87**	3.30
	(1.35)	(1.85)
$ln(Y/L)_{I,1960}$	-0.37**	-0.18
	(0.08)	(0.14)
$ln(I/Y)_i$	0.05	0.30**
	(0.06)	(0.15)
$ln(n+g+\delta)_i$	-0.08	-0.04
	(0.40)	(0.34)
ln(School),	0.05	0.19
	(0.10)	(0.16)
AFRICA	-0.25	-0.38**
	(0.13)	(0.18)
LATAMER	0.32	-0.31
	(0.21)	(0.25)
G	7.50**	-1.98
	(1.43)	(1.53)
PRIM70	0.49	-0.19
	(0.26)	(0.16)
STTEAPRI70	-0.59	-0.11
	(0.38)	(0.42)
STTEASEC70	0.75	0.50
	(0.57)	(0.58)
Adjusted R^2	0.50	0.40
Observations	42	42

Table 14: *Augmented low literacy – low output and high literacy – high output splits*

From table 14, it is understandable that a big proportion of the variation of the dependent variable is being explained only by three parameters in the low group and by two in the high group. The initial output coefficient is negative and significant for the low group, while its negative but insignificant for the high group. This may be happening due to the heterogeneity within the high group. This means that the high group includes not only developed and rich countries, but also countries that are very close to the edge of being categorized in this group. AFRICA's coefficient remains negative for both of the groups but it is significant only in the high group. As usual, the public expenditure numerical quantity is high for the low group. In every one of the previous tables that we have included so far, each low group exhibits high numbers relating the G variable.

As usual, we carry out Wald specification tests to examine whether the data provide evidence of multiple regime existence.

Additional control Variables	*Significance value*
AFRICA, LATAMER, G, PRIM70, STTEAPRI70, STTEASEC70	0.0000

Table 15: Significance level of Wald test of Ho that the Solow model coefficients remain the same between the subsamples when adding the extra variables dictated by Barro (1991)

From panel 15 it is completely clear, due to the marginal significance value is far less than the 0.05 significance level, that the Solow model coefficients are still different between the two groups even if we add into the analysis the extra set of variables. This outcome is consistent with what the main article claims, with the only difference that in the base study the authors use an even more extended set which allows them to run Wald tests using different combinations of the extra variables each time. Still, the findings between our analysis and theirs, remain the same.

In the next sub-part that follows, we carry out the same empirical process from scratch but for an extended sample of countries and for a bigger time spectrum.

4.2 Results for 1960-2015 and 108 countries

In this section we extend the number of years and the sample of the countries. We expect that the expansion of the time period will produce interesting results due to the massive economic changes that took place. The sharp increase of trade, the globalization that led to spread of knowledge/technology and established a working environment where the workers can move easily between countries, the terrorist acts of 9/11 and the global economic recession of 2008[23] seem to have created a brand-new economic environment. For these reasons, we believe that the inclusion of the years until 2015 (where the data actually stop) would provide us with precious results. Regarding the sample we add 12 more countries that we could use data for the variables of the model. Namely, the countries we added are:

1. Albania
2. Barbados
3. Bulgaria
4. Cambodia
5. Cyprus
6. Gabon
7. Iran
8. Kuwait
9. Malta
10. Qatar
11. Cabo Verde
12. Botswana

Table 16 presents the OLS results of the two-way split process for each of the control variables:

Dependent variable: $\ln(Y/L)_{I,2015} - \ln(Y/L)_{I,1960}$

Method: OLS	*Low group (LR_{1960})*	*High Group (LR_{1960})*	*Low group (Y/L_{1960})*	*High Group (Y/L_{1960})*
Constant	4.33	1.64	-0.97	3.44**
	(5.10)	(2.90)	(6.32)	(1.64)
$\ln(Y/L)_{I,1960}$	-0.58**	-0.46**	-0.65**	-0.57**
	(0.15)	(0.13)	(0.24)	(0.16)
$\ln(I/Y)_i$	0.27	-0.09	-0.04	0.29
	(0.24)	(0.26)	(0.28)	(0.17)
$\ln(n+g+\delta)_i$	-1.02	-1.62**	-3.09	-1.50**
	(1.65)	(0.70)	(1.73)	(0.53)
$\ln(School)_i$	0.48**	0.75**	0.60**	0.52**
	(0.14)	(0.29)	(0.22)	(0.18)
Adjusted R^2	0.28	0.25	0.33	0.42
Observations	54	54	54	54

$Table\ 16: Two-way\ splits\ based\ on\ LR\ and\ \frac{Y}{L}\ 1960\ initial\ values$

[23] The global economic recession of 2008 revealed also important problems in the structure of society. Populist and extremist parties and groups made their appearance in many parts of the world which led to the increase of political instability.

The majority of the constant coefficients are relatively high and positive, with the last one of the rows to be significant. The most crucial part here, is that all groups reveal negative and significant coefficients on the initial output variable. Convergence seems to occur. These results differ from the ones that the 1960-1980 period produced where only two out of four groups exhibited significant coefficients. We can see an improvement in this sector. Regarding the coefficients of investment share in GDP, the difference between the 1960-1985 and 1960-2015 results is that the coefficient of the fourth group (High group based on initial output) is significant for the small time-frame and sample but it is not when we expand the data. All the numbers related to population growth have negative signs while only the high groups are meaningful. Last and most surprising, the share of working age population enrolled in secondary level of education, is important for all the group. The coefficients are between 0.48 and 0.75. We clearly see an improvement in three out of four groups due to the higher rates of adjusted-R squared. Though, the literacy rate high group captures a sharp decline in the $\bar{R}^2$ from 0.96 in the 1960-1985 period to 0.25 in the 1960-2015. The break points of the groups when we consider literacy rate as control variable is 47.05% and when the control variable is the initial output the break point is 5608.845$. We remind once more, that the 1960-2015 span of time includes bigger sample as well. Moving on, we split the sample in three equal subgroups for each one of the control variables.

$Dependent\ variable$: $\ln(Y/L)_{I,2015} - \ln(Y/L)_{I,1960}$

Method: OLS	*Group A (LR split) (Low)*	*Group B (LR split) (Intermediate)*	*Group C (LR split) (High)*
Constant	2.50	-2.17	4.89**
	(6.18)	(4.61)	(2.05)
ln(Y/L)$_{I,1960}$	-0.46**	-0.43	-0.54**
	(0.19)	(0.24)	(0.12)
ln(I/Y)$_i$	0.07	0.12	0.11
	(0.27)	(0.35)	(0.19)
ln(n+g+δ)$_i$	-1.11	-3.12**	-0.92
	(2.03)	(1.03)	(0.52)
ln(School)$_i$	0.37**	0.75**	0.97**
	(0.18)	(0.29)	(0.30)
Adjusted R^2	0.18	0.23	0.42
Observations	36	36	36

$Table\ 17: Three-way\ split\ based\ on\ Literacy\ Rate$

Following the previous pattern, panel 17 exhibits quite different quantitative values. Specifically, the initial output coefficient values, are close enough with the first and the third to be significant. The same stands for the 1960-1985 era, but the quantities are much different. Once more, the investment share doesn't play a role in the model at all. Group C in 1960-1985 exhibits a statistically significant coefficient. The quantities of the population growth seem to vary much but we notice a homogeneity regarding the numerical signs. However, only the Group B parameter is significant. All of the estimated parameters for the $\ln(School)_i$ are negative and significant. We also notice a steady scale-up from Group A to Group C, with numbers in absolute values. Adjusted-R squared exhibits a smooth scale-up as well. Between the low and intermediate group, break point is 21.85% and between the intermediate and high ones it is 71.05%.

Table 18 splits the sample into three equal groups based on the initial output as control variable.

$Dependent\ variable$: $\ln(Y/L)_{I,2015} - \ln(Y/L)_{I,1960}$

Method: OLS	*Group A (y split) (Low)*	*Group B (y split) (Intermediate)*	*Group C (y split) (High)*
Constant	-6.99	-3.45	-0.44
	(10.5)	(5.54)	(1.69)
$ln(Y/L)_{I,1960}$	-0.28	-0.01	-0.40**
	(0.42)	(0.31)	(0.17)
$ln(I/Y)_i$	0.02	0.47	-0.02
	(0.37)	(0.31)	(0.14)
$ln(n+g+\delta)_i$	-4.19	-2.46	-2**
	(3.07)	(1.43)	(0.40)
$ln(School)_i$	0.28	0.73**	0.25
	(0.27)	(0.29)	(0.23)
Adjusted R^2	0.03	0.46	0.40
Observations	36	36	36

$Table\ 18: Three-way\ split\ based\ on\ initial\ output\ \frac{Y}{L}\ (denoted\ as\ y)$

As one may notice, numerical indications don't are mostly the same between the parameters of the same variable, with an exception on the investment one where the

Group A and B are positive while the Group C exhibits negative coefficient. As for $\ln(\frac{Y}{L})_{i,1960}$ estimates, we spot presence of convergence within Group A and Group C, but only the latter is significant. Estimation of Group B is almost zero and insignificant. Working-age population growth affects negatively with only the coefficient of Group C to be significant. The estimations of $\ln(School)_i$ for the Group A and C are positive and almost equal. The Group B exhibits a higher coefficient, but still positive and significant. A general view of how the model fits in every case is being reflected from the last row of the table. Completely bad fit for Group A, but good enough for the other two groups. Comparing with the results of the 1960-1985 era, one notices an improvement for the Groups B and C but a deterioration for Group A. The break point between the low-intermediate and intermediate-high groups are 3031.06$ and 9472.305$ respectively.

Our interest is mainly focused on the estimations that produced when the control variables interact, which is something closer to the reality. We then derive the estimations presented in table 19:

$Dependent\ variable$: $\ln(Y/L)_{I,2015} - \ln(Y/L)_{I,1960}$

Method: OLS	*Low Group (LR and Y/L interaction)*	*High Group (LR and Y/L interaction)*
Constant	17.86**	5.90**
	(7.96)	(2.17)
$ln(Y/L)_{I,1960}$	-1.12**	-0.58**
	(0.24)	(0.14)
$ln(I/Y)_i$	0.35	0.26
	(0.37)	(0.20)
$ln(n+g+\delta)_i$	2.42	-0.76
	(2.34)	(0.67)
$ln(School)_i$	0.65**	0.98**
	(0.15)	(0.28)
Adjusted R^2	0.45	0.32
Observations	42	42

$Table$ 19: $Low\ and\ High\ subgroups\ whilst\ Literacy\ Rate\ and \frac{Y}{L}\ interact$

The panel 19 shows some extremely high significant levels of constant for both of the groups, but even more for the low group. Both of them, exhibit convergence which is faster in the low group and slower in the high. Investment per GDP exhibits positive effectiveness in both of the groups, but none of them are meaningful to be included. Furthermore, population growth is positive for the low group and negative for the high group. For both, though, it isn't significant. Enrollment in secondary school seem to have a great effect on both groups with the high group to have a higher coefficient. $\bar{R}^2$ is higher for the low group (0.45) and lower for the high group (0.32). Comparing with the estimations of the two groups in previous section, we notice a rise -in absolute values- for all the coefficients. The results in this table seem to be more rational and more convenable, since we have a good percent variation on the dependent variable that is being explained from the independent variables. The related Wald tests for these groups are giving the following results:

Subsamples	*Unconstrained Regressions*
Two-way split based on	
$(Y/L)_{I,1960}$	0.1738
$LR_{i,1960}$	0.5273
Three-way split based on	
$(Y/L)_{I,1960}$	0.0043
$LR_{i,1960}$	0.7208
Four-way split based on both $(Y/L)_{I,1960}$ and $LR_{i,1960}$	0.0064

$Table$ 20: *Significance level of Wald tests regarding the null hypothesis that the parameters are constant across the indicated subsamples*

The results give mixed information about the presence or not of multiple regimes in our data. In the first sector of table 20, where we examine the two-way split of the data, we do not reject the null hypothesis of the single regime model. As for the three-way split, we reject the null hypothesis when we break the sample based on initial output, but we do not reject it when using the literacy. The third panel rejects the single regime specification. And in this case, we exclude the low-literacy/high-output and high-literacy/low-output from the four-way split due to the small number of observations.

We now proceed the extract the results when we add the extra set of variables, as we did and in the previous section.

The results are presented in table 21:

Dependent variable: $\ln(Y/L)_{I,2015} - \ln(Y/L)_{I,1960}$

Method:OLS	*Low Group (LR_{1960})*	*High Group (LR_{1960})*	*Low Group (Y/L_{1960})*	*High Group (Y/L_{1960})*
Constant	4.70	5.07	2.14	5.87**
	(5.65)	(3.32)	(5.14)	(2.05)
$ln(Y/L)_{I,1960}$	-0.82**	-0.45**	-0.75**	-0.69**
	(0.12)	(0.16)	(0.16)	(0.18)
$ln(I/Y)_i$	0.23	-0.06	0.02	0.24
	(0.20)	(0.38)	(0.18)	(0.19)
$ln(n+g+\delta)_i$	-1.21	-0.56	-2.08	-1.16
	(1.84)	(0.68)	(1.65)	(0.67)
ln(School),	0.11	0.87**	0.38**	0.52
	(0.14)	(0.44)	(0.15)	(0.28)
AFRICA	-0.53**	-0.95**	-0.74**	-0.15
	(0.22)	(0.22)	(0.17)	(0.44)
LATAMER	-0.29	-0.38	-1.78**	-0.30
	(0.27)	(0.32)	(0.30)	(0.22)
G	5.35	-1.08	8.13**	-1.77
	(3.67)	(3.59)	(3.76)	(2.86)
PRIM70	0.92**	-0.26	0.20	-0.08
	(0.39)	(0.32)	(0.34)	(0.41)
STTEAPRI70	-1.16**	-0.22	-0.25	-0.58
	(0.39)	(0.91)	(0.45)	(0.77)
STTEASEC70	0.84	0.60	-0.02	0.04
	(0.95)	(0.71)	(0.77)	(1.08)
Adjusted R^2	0.45	0.33	0.66	0.40
Observations	54	54	54	54

$Table\ 21: Two-way\ split\ based\ on\ LR\ and\ \frac{Y}{L}\ initial\ values, adding\ extra\ variables\ as\ stated\ in\ Barro$

The general view of table 21 is that there is a relatively good fit of the model. The adjusted-R squared quantities are in the range between 33% and 66%. Constants in all cases exhibit high values, with the latest one to be also significant. On the one hand, the initial output coefficients are all negative in the range of -0.45 and -0.82, showing significance. On the other hand, the $\ln(\frac{I}{Y})_i$ coefficients are small and insignificant. In addition, estimations of $\ln(n+g+\delta)_i$ coefficients are negative and may be excluded from the model. The $\ln(School)_i$ seems to contribute in the model, significantly and positively for the high literacy group and low output group. Moreover, AFRICA affects negatively three out of four groups while LATAMER is significant only in one out of four. As expected, military expenditure is high and positive for the low groups, though only in the low output group is significant. It is addressing a negative impact on the high groups. Only the low literacy group is being affected by the PRIM70 and STTEASEC70, positively and negatively respectively. Overall, the fit of the model in each case is quite decent. One quite not desirable feature is that the $\bar{R}^2$rates are exhibiting a small decline compared to the rates in 1960-1985 table.

Table 22 is describing the three-way split based on literacy rate when we add the extra set of variables:

Dependent variable: $\ln(Y/L)_{I,2015} - \ln(Y/L)_{I,1960}$

Method:OLS	*Group A (LR split) (Low)*	*Group B (LR split) (Intermediate)*	*Group C (LR split) (High)*
Constant	3.10	0.79	4.02**
	(7.53)	(3)	(1.75)
ln(Y/L)I,1960	-0.73**	-0.37	-0.56**
	(0.23)	(0.20)	(0.13)
ln(I/Y)i	0.04	-0.03	0.12
	(0.24)	(0.30)	(0.19)
ln(n+g+δ)i	-1.55	-2.36**	-1.29**
	(2.43)	(0.76)	(0.57)
ln(School)i	0.14	0.75**	0.90**
	(0.19)	(0.26)	(0.35)
AFRICA	-0.32	-1.51**	0
	(0.23)	(0.33)	(0)
LATAMER	-1.12**	-1.07**	0.29
	(0.54)	(0.21)	(0.28)
G	3.20	-5.32	1.82
	(4.17)	(5.40)	(1.92)
PRIM70	1.13**	-0.87**	-0.06
	(0.44)	(0.23)	(0.28)
STTEAPRI70	-1.43	0.66	-2.02**
	(1.24)	(0.96)	(0.73)
STTEASEC70	-0.33	-1.17	1.97**
	(0.94)	(1.40)	(0.86)

Adjusted R^2	0.38	0.60	0.53
Observations	36	36	36

$Table\ 22: Augmented\ three - way\ split\ based\ on\ Literacy\ Rate\ \ control\ variable$

For the Group A, we notice evidence of convergence due to the negative and significant coefficient of initial output variable. The coefficients of LATAMER and PRIM70 are more or less equal, with the first affecting negatively and the second affecting positively the dependent variable. The $\bar{R}_A^2$ is estimated at 0.38, which is the lower numerical value amongst these groups. As for the Group B, there is no significance in the initial output coefficient. The rate of change of the working age population seems to highly affect the growth in a negative way. The same stands for the two dummy variables and for the PRIM70, however the latter has a smaller effect in growth. Only school enrollment may push the growth up. The adjusted-R squared is the highest one, and one of the highest so far in the analysis. Lastly, the Group C exhibits somehow different results. It includes a high and significant constant followed by negative and significant coefficients in $\ln(\frac{I}{Y})_{i,1960,}$ $\ln(n+g+\delta)_i$ and STTEAPRI70 whilst the $\ln(School)_i$ and STTEASEC70 have a positive effect. The adjusted-R squared stands at 0.53. In an effort to compare these results with the ones produced from the 1960-1985 period, we see an increase over 10% in the explanation of the growth in the Group B and C but a 20% decline in the Group A. Though, the quantities are higher in absolute values in the expanded time frame.

A more stable frame is exhibited in the table 23:

Dependent variable: $\ln(Y/L)_{I,2015} - \ln(Y/L)_{I,1960}$

Method:OLS	*Group A (y split) (Low)*	*Group B (y split) (Intermediate)*	*Group C (y split) (High)*
Constant	13.29	0.20	1.60
	(7.22)	(8.08)	(1.95)
ln(Y/L)$_{I,1960}$	-1.18**	-0.30	-0.58**
	(0.29)	(0.41)	(0.14)
ln(I/Y)$_i$	0.54**	0.60**	-0.14
	(0.21)	(0.23)	(0.17)
ln(n+g+δ)$_i$	0.37	-1.96	-1.69**
	(2.39)	(1.70)	(0.57)
ln(School)$_i$	0.37**	0.60	0.04
	(0.16)	(0.39)	(0.20)
AFRICA	-0.57**	-0.48	-0.47
	(0.25)	(0.25)	(0.32)
LATAMER	-3.51**	-0.14	-0.38**
	(0.61)	(0.26)	(0.17)
G	-0.16	8.17	-0.95
	(2.52)	(6.51)	(2.51)
PRIM70	0.31	-0.07	0.61**
	(0.42)	(0.54)	(0.25)
STTEAPRI70	0.10	-1.20	-1.34**
	(0.75)	(0.74)	(0.67)
STTEASEC70	-1.04	1.94	0.30
	(0.95)	(1.29)	(1.02)
Adjusted R^2	0.57	0.58	0.57
Observations	36	36	36

Table 23: Augmented three − way split based on output as control variable

Similar with the previous panel, Group A and C exhibit significant and negative coefficients on initial output, but Group B coefficient lacks significance. Also, investment is positive and significant in Group A and B, but negative and insignificant in Group C. Coefficients of $\ln(n+g+\delta)_i, PRIM70, STTEAPRI70$ are significant only in Group C. On the other side, only Group A exhibits significant coefficients in $\ln(School)_i, AFRICA$. $LATAMER$ is far away from zero and in the negative part and significant for Group A, whilst its smaller, negative and significant for Group C. All of the adjusted-R squared quantities are in the same level. Comparing the results with the estimations from the table in 1960-1985 we notice a small improvement in Group A and B but an also small decline in Group C.

Finally, we derive the estimations when the control variables are let to affect each other:

Dependent variable: $\ln(Y/L)_{I,2015} - \ln(Y/L)_{I,1960}$

Method:OLS	*Low Group (LR-Y/L interaction)*	*High Group (LR-Y/L interaction)*
Constant	14.09**	8
	(5.76)	(2.21)
$\ln(Y/L)_{I,1960}$	-0.97**	-0.57**
	(0.14)	(0.12)
$\ln(I/Y)_i$	0.08	0.30
	(0.20)	(0.21)
$\ln(n+g+\delta)_i$	2.02	-0.19
	(1.84)	(0.71)
$\ln(School)_i$	0.31	1.09**
	(0.16)	(0.44)
AFRICA	-0.63**	-0.96**
	(0.19)	(0.22)
LATAMER	-0.32	-0.02
	(0.29)	(0.31)
G	10.67**	-0.41
	(3.23)	(2.47)
PRIM70	0.17	-0.23
	(0.34)	(0.37)
STTEAPRI70	-0.46	-1.05
	(0.28)	(0.73)
STTEASEC70	0.73	0.91
	(0.81)	(1)
Adjusted R^2	0.68	0.39
Observations	54	54

Table 24: Augmented low literacy – low output and high literacy – high output splits

Noteworthy is that the constant of the low group is extremely high, opening the gap between the two groups in the first place. Convergence seems to be faster in the low group where is -0.97 and significant at α=5% whereas it is -0.57 and significant at the same level for the high group. The coefficient being higher in the low groups is something that comes straight from the theory and it is being verified from these results. School enrollment coefficient is just a bit above 1 in the high group and is significant. AFRICA affects both groups negatively, with bigger tense on the high group. Military and education expenditure is giving a great effect in the low group where the coefficient is 10.67 and statistically significant. None of the other variables is significant for any group. Again, the specific estimations are higher than those in the 1960-1985 time frame accompanied by higher adjusted-R squared quantities.

For the last part of this section, we need to carry out Wald test to see whether the Solow model parameters are remaining the same between the low and high group when we add on the extra variables. The Wald test provides us with the result presented in table 25:

Additional control Variables	*Significance value*
AFRICA, LATAMER, G, PRIM70, STTEAPRI70, STTEASEC70	0.0064

Table 25: Significance level of Wald test of Ho that the Solow model coefficients remain the same between the subsamples when adding the extra variables dictated by Barro

The Wald test provides us with more evidence to reject the null hypothesis in favor of the multiple regime alternative hypothesis. The findings remain consistent with what the authors found in their research.

For the last part of the empirical approach, we estimate in the following section all the coefficients for all the possible cases for the 1960-2015 time frame using as sample the 96 countries we used in the first place.

4.3 Results for 1960-2015 and 96 countries

In this section, we don't expect big difference on how (focusing mostly in the directions of the effects) the indicated variables affect growth, however we expect a decline in the adjusted-R squared values. Going straight into the estimations, we start with the basic variables and splitting the data properly, we derive estimations that we present in table 26:

$Dependent\ variable$: $\ln(Y/L)_{I,2015} - \ln(Y/L)_{I,1960}$

Method: OLS	*Low group (LR_{1960})*	*High Group (LR_{1960})*	*Low group (Y/L_{1960})*	*High Group (Y/L_{1960})*
Constant	1.91	1.33	-4.34	3.87
	(4.57)	(3.54)	(7.66)	(2.14)
$ln(Y/L)_{I,1960}$	-0.47**	-0.42**	-0.47	-0.42**
	(0.15)	(0.21)	(0.26)	(0.12)
$ln(I/Y)_i$	0.20	-0.18	0.004	0.32
	(0.24)	(0.25)	(0.31)	(0.20)
$ln(n+g+\delta)_i$	-1.48	-1.49	-3.73	-0.84
	(1.42)	(0.76)	(2.20)	(0.57)
$ln(School)_i$	0.38**	0.61	0.39**	0.57**
	(0.15)	(0.41)	(0.19)	(0.25)
Adjusted R^2	0.19	0.15	0.23	0.31
Observations	48	48	48	48

$Table\ 26: Two - way\ splits\ based\ on\ LR\ and\ \frac{Y}{L}\ 1960\ initial\ values$

Table 26, shows that convergence coefficients for the low literacy and output groups are the same, but only the first is significant. Both of the coefficients of the high groups are at the exact same level and both are significant. Lastly, the $\ln(School)_i$ affect positively three out of four groups. Therefore, we only have a handful of variables that are able to explain the variation of growth, hence the low values of the adjusted-R squared. The results are as expected. We notice that the same sample of countries but in a bigger time frame gives higher adjusted-R squared values in three out of four groups, but it reveals a sharp decline in the high literacy rate group. Also, the $\bar{R}^2$ values in the two-way split for 1960-2015 and 108 countries are higher than those for 1960-2015 and 96 countries. As for the break point when we split the sample based on literacy rate this is at level of 46.15% whilst the break point when we use initial output is at 5777.33$. Table 27, we present the results of the M-R-W approach for the 1960–2015-time spectrum:

$Dependent\ variable$: $\ln(Y/L)_{I,2015} - \ln(Y/L)_{I,1960}$

Method: OLS	*M-R-W*
Constant	0.82
	(2.06)
$ln(Y/L)_{I,1960}$	-0.42**
	(0.11)
$ln(I/Y)_i$	0.08
	(0.17)
$ln(n+g+\delta)_i$	-1.79**
	(0.48)
$ln(School)_i$	0.53**
	(0.15)
Adjusted R^2	0.28
Observations	98

$Table\ 27: M - R - W\ approach\ for\ 1960 - 2015\ time\ frame$

Table 27 exhibits negative and statistically significant coefficients for output per capita and working-age population growth variables, but positive and significant parameter for $\ln(School)_i$. Comparing with the M-R-W results we have extracted for 1960-1985 period, we clearly see an improvement regarding the fit of the model, as adjusted-R squared becomes higher. The numerical signs for the variables remain the same for both tables.

The three-way split of the control variables is presented in the tables 28 and 29.

$Dependent\ variable$: $\ln(Y/L)_{I,2015} - \ln(Y/L)_{I,1960}$

Method: OLS	*Group A (LR split) (Low)*	*Group B (LR split) (Intermediate)*	*Group C (LR split) (High)*
Constant	3.20	-3.63	5.03**
	(6.67)	(3.72)	(2.37)
$ln(Y/L)_{I,1960}$	-0.47**	-0.26	-0.52**
	(0.21)	(0.22)	(0.14)
$ln(I/Y)_i$	0.08	0.16	0.11
	(0.27)	(0.32)	(0.18)
$ln(n+g+\delta)_i$	-0.86	-2.95**	-0.73

	(2.17)	(0.73)	(0.58)
ln(School)$_i$	0.37	0.43	0.74**
	(0.20)	(0.29)	(0.25)
Adjusted R^2	0.14	0.07	0.33
Observations	32	32	32

$Table\ 28\!: Three - way\ split\ based\ on\ Literacy\ Rate$

Looking at table 28, one may notice high constant values in all cases, with the constant in Group C being the only significant one. Groups A and C include negative and significant values of convergence, whilst Group B includes negative but not significant one. From the rest of the estimates, we find a negative and significant value of working-age population growth in the intermediate group and a significantly positive value of $\ln(School)_i$ in the high literacy group. From these three groups, only the high group exhibits a decent value in adjusted-R squared. The other two are fairly small, doubting the good fit of the model, whereas the $\bar{R}^2$ from the sample table split for 108 countries give somehow higher values. However, in this case the Group B performs better but the other groups perform weaker, compared to the 1960-1985 period. The numerical break points between low-intermediate and intermediate-high groups are 21.85% and 70.25% respectively.

Table 29 is covering the results of the three-way split based on the initial output as control variable. A first look at the table shows that many of the variables can be excluded from the analysis since the majority of parameters are insignificant. Comparing the results from table 27 with those in table 28, we understand that the constant in M-R-W is much smaller than the constants in all three literacy rate-splited groups (in absolute values). Both tables exhibit negative values for $\ln(\frac{Y}{L})_{i,1960}$ with all being statistically significant, except the one in intermediate group of table 28. Furthermore, all of the coefficients regarding the investment share per GDP are positive, fairly small but insignificant for the analysis. We, also, notice a homogeneity regarding the numerical signs for both of the $\ln(n+g+\delta)_i$ and $\ln(School)_i$. However, the coefficient of working-age population growth is significant only in the M-R-W table and in the Group B of table 28. Coefficients of $\ln(School)_i$ are meaningful only in the M-R-W and Group C.

Dependent variable: $\ln(Y/L)_{I,2015} - \ln(Y/L)_{I,1960}$

Method: OLS	*Group A (y split) (Low)*	*Group B (y split) (Intermediate)*	*Group C (y split) (High)*
Constant	3.20	-0.15	-0.08
	(6.67)	(3)	(1.95)
$ln(Y/L)_{I,1960}$	-0.47**	0.02	-0.30
	(0.21)	(0.41)	(0.16)
$ln(I/Y)_i$	0.08	0.55	0.07
	(0.27)	(0.35)	(0.12)
$ln(n+g+\delta)_i$	-0.86	-0.97	-1.58**
	(2.17)	(1.30)	(0.33)
$ln(School)_i$	0.37	0.54**	0.17
	(0.20)	(0.24)	(0.19)
Adjusted R^2	0.14	0.23	0.29
Observations	32	32	32

Table 29: Three − way split based on output as control variable

Examining the results, we see an intense difference between the constants of each group. Moreover, Group A exhibits a negative and significant parameter regarding the initial output showing evidence of convergence within the group. Also, Group C exhibits negative coefficient, which is not statistically significant. The coefficient of Group B is almost zero, exhibiting no signs of convergence. Once more, the investment share of GDP exhibits small, positive and insignificant coefficients for all cases. Growth rate of working-age population is negative in all subsamples but significant only in Group C. Similar for the $\ln(School)_i$, only the coefficient for Group B is significant. Despite the weird nature of these results, they tend to be more satisfactorily since they explain a higher proportion of the variation of growth in each case compared to the results produced in 1960-1985. Comparing with the results for the 108 countries, we clearly see an explanation improvement for Group A, but a weaker status for the other two groups. The break points between Group A and B and between Group B and C are 3586.22$ and 9793.025$.

Regarding the case where we split the data based on both literacy rate and initial output, we can't extract unbiased results since the split produces unequal subsamples. Hence, the OLS estimation of the groups is not possible.

The results of the Wald specification tests in this section are:

Subsamples	*Unconstrained Regressions*
Two-way split based on $(Y/L)_{I,1960}$ $LR_{i,1960}$	 0.2181 0.4357
Three-way split based on $(Y/L)_{I,1960}$ $LR_{i,1960}$	 0.7056 0.5900
Four-way split based on both $(Y/L)_{I,1960}$ and $LR_{i,1960}$	-

Table 30: Significance level of Wald tests regarding the H_O hypothesis that the Solow model parameters are constant across the indicated subsamples

The results from the table 30 aren't consistent with what the base study finds. For all the possible cases we do not reject the null hypothesis, thus there are evidence of a single regime model existence.

For the last part of the empirical approach of this study is to add the extra variables and run again the OLS estimations.

$Dependent\ variable$: $\ln(Y/L)_{I,2015} - \ln(Y/L)_{I,1960}$

Method:OLS	*Low Group (LR_{1960})*	*High Group (LR_{1960})*	*Low Group (Y/L_{1960})*	*High Group (Y/L_{1960})*
Constant	2.50	3.02	-1.39	6.58**
	(4.82)	(2.95)	(4.81)	(2.76)
$ln(Y/L)_{I,1960}$	-0.70**	-0.35	-0.52**	-0.47**
	(0.18)	(0.17)	(0.20)	(0.14)
$ln(I/Y)_i$	0.18	-0.37	-0.001	0.28
	(0.20)	(0.24)	(0.19)	(0.17)
$ln(n+g+\delta)_i$	-1.65	-0.51	-2.38	-0.24
	(1.62)	(0.78)	(1.45)	(0.65)
$ln(School)_i$	0.09	0.35	0.08	0.63**
	(0.14)	(0.44)	(0.18)	(0.31)
AFRICA	-0.47**	-1**	-0.64**	-0.91**
	(0.17)	(0.25)	(0.18)	(0.23)
LATAMER	-0.19	-0.73**	-1.07**	-0.39
	(0.40)	(0.31)	(0.51)	(0.21)
G	5.79	0.96	10.74**	-3.04
	(4.06)	(2.90)	(4.80)	(2.45)
PRIM70	0.64	-0.25	0.43	-0.18
	(0.47)	(0.26)	(0.46)	(0.27)
STTEAPRI70	-1.23**	0.51	-0.65	-0.47
	(0.60)	(0.81)	(0.48)	(0.75)
STTEASEC70	1.31	0.45	0.54	0.42
	(1)	(0.82)	(0.97)	(1.04)
Adjusted R^2	0.39	0.38	0.52	0.39
Observations	48	48	48	48

$Table\ 31\colon Two-way\ split\ based\ on\ LR\ and \frac{Y}{L}\ initial\ values, adding\ extra\ variables\ as\ stated\ in\ Barro(1991)$

All of the coefficients are negative and significant at α=5% with the exception of the coefficient of high literacy rate group. Thus, signs of convergence exist in all cases. On the other hand, the estimations of the $\ln(\frac{I}{Y})_i$ aren't useful in the specific analysis. One and only significant value of the $\ln(School)_i$ estimations is included in the high initial output group, exhibiting a rather high estimation. The most important variable in this table, should be the AFRICA, which is negative and significant in all the subsamples. The numerical values range from -0.47 to -1. In an almost similar way the LATAMER, exhibits negative values in all the groups, however only the two out of four coefficients are significant. Military and educational expenses are significant and fairly high in low initial output group. High value, also, exhibits the low literacy rate group, however it is not significant. From the rest educational variables, only the coefficient from STTEAPRI70 in the low literacy group is significant. The performance of the model in the cases above are a bit weaker than the fit of the model when examining the sample of 108 countries. Surely, the adjusted-R squared are lower even from the 1960-1980 time frame.

For the last time we split the sample into three equal groups for each one of the control variables, estimating the coefficients of the Solow model parameters along with the extra variables. The results are presented in table 32:

Dependent variable: $\ln(Y/L)_{I,2015} - \ln(Y/L)_{I,1960}$

Method:OLS	*Group A (LR split) (Low)*	*Group B (LR split) (Intermediate)*	*Group C (LR split) (High)*
Constant	2.86	0.39	4.09
	(9.05)	(2.70)	(2.16)
$\ln(Y/L)_{I,1960}$	-0.70**	-0.24	-0.55**
	(0.31)	(0.17)	(0.13)
$\ln(I/Y)_i$	0.06	0.01	-0.03
	(0.25)	(0.28)	(0.25)
$\ln(n+g+\delta)_i$	-1.57	-1.95**	-1.03
	(2.84)	(0.63)	(0.56)
$\ln(School)_i$	0.16	0.45	0.65
	(0.21)	(0.26)	(0.58)
AFRICA	-0.33	-1.47**	0
	(0.27)	(0.34)	(0)
LATAMER	-1.08	-1.02**	-0.05
	(0.58)	(0.19)	(0.40)
G	4.46	-4.65	3.58
	(5.11)	(4.99)	(2.82)
PRIM70	0.96	-0.85**	-0.17
	(0.55)	(0.19)	(0.38)
STTEAPRI70	-1.36	0.44	-0.97
	(1.27)	(0.86)	(0.62)
STTEASEC70	-0.18	-1.60	1.94**
	(1.02)	(1.47)	(0.88)
Adjusted R^2	0.32	0.48	0.41

Observations	32	32	32

Table 32: Augmented three – way split based on Literacy Rate control variable

Table 32 shows that there are significant signs of convergence in two out of three groups. Estimations of $\ln(n+g+\delta)_i$ are negative and exceeding the threshold of -1, however only for the Group B is significant. AFRICA and LATAMER follow a similar way, although the variable pointing out the Sub-Saharan countries is being excluded from the Group C due to the ''dummy variable trap''. PRIM70 exhibits negative and statistically significant coefficient in the Group B, while STTEASEC70 exhibits a relatively high, positive and significant estimation in the Group C. We notice a small decrease in the explanation percentage when we compare with the 108 countries sample as well as the 1960-1985.

We complete the empirical approach with the three-way split based on initial output.

Dependent variable: $\ln(Y/L)_{I,2015} - \ln(Y/L)_{I,1960}$

Method:OLS	*Group A (y split) (Low)*	*Group B (y split) (Intermediate)*	*Group C (y split) (High)*
Constant	2.86	5.91	3.71
	(9.05)	(3.14)	(2.42)
$\ln(Y/L)_{I,1960}$	-0.70**	-0.30	-0.53**
	(0.31)	(0.38)	(0.13)
$\ln(I/Y)_i$	0.06	0.73**	0.0001
	(0.25)	(0.21)	(0.13)
$\ln(n+g+\delta)_i$	-1.57	0.16	-1.02**
	(2.84)	(1.07)	(0.49)
ln(School),	0.16	0.43	0.16
	(0.21)	(0.26)	(0.29)
AFRICA	-0.33	-0.65**	-0.96**
	(0.27)	(0.25)	(0.18)
LATAMER	-1.08	-0.28	-0.30
	(0.58)	(0.27)	(0.18)
G	4.46	5.08	-1.70
	(5.11)	(6.07)	(2.46)
PRIM70	0.96	-0.26	0.27
	(0.55)	(0.42)	(0.16)
STTEAPRI70	-1.36	-0.88	-1.77**
	(1.27)	(0.83)	(0.61)
STTEASEC70	-0.18	2.40**	1.15
	(1.02)	(1.22)	(0.85)
Adjusted R^2	0.32	0.48	0.61

Observations	32	32	32

Table 33: Augmented three − way split based on output as control variable

Not a major difference in table 33 compared to the previous results, as convergence still occurs in all groups, though in Group B the coefficient isn't statistically significant. The intermediate group exhibits positive effect of investment in the growth process whilst working-age population growth is negative and significant only in the high group. AFRICA is negative in all groups but significant only in intermediate and high groups. STTEAPRI70 is negative and significant for the high group whilst STTEASEC70 is positive and significant for the intermediate group. The adjusted-R squared values are smoothly scaling up. More or less the values are close to the values that produced in the 1960-1985 example as well as the 1960-2015 time frame with 108 as sample.

We are not able to estimate the coefficients when the initial output and literacy rate are affecting each other, because the indicated subsamples aren't equal. Hence, we can't run the Wald tests either.

5. Summary

To sum up, the main goal of this study is to verify the claim of multiple-regime existence. At first, we properly split the data into subgroups based on two control variables, the literacy rate at 1960 and the initial output at 1960. Before we proceed to the OLS estimations for the unconstrained equations, we carry out heteroscedasticity tests. White tests find evidence of heteroscedasticity presence in some cases. Thus, we fix this problem by selecting HAC-robust standard errors. Afterwards, we proceed on estimating the parameters of the Solow model variables. The results exhibit that the effects in our estimations have the same direction with the ones produced in the main study, however there may be differences in some cases regarding the numerical values. We then, proceed into Wald specification tests to examine whether the data exhibit different laws of motion.

Our results are not always consistent with this claim, exhibiting in some cases evidence of a single-regime specification model. The specific process is being carried out once for the basic variables and once when we add the extra set of variables for every time frame. We start examining the time frame of 1960-1985 that includes 96 countries (sample extracted from MRW research), the time frame 1960-2015 for the same 96 countries and finally the time frame 1960-2015 for 108 countries. From the results that we have derived, we understand that convergence occurs in the majority of the indicated subsamples. For a group that is categorized as ''low'' the convergence process takes place with higher rates of speed rather than the intermediate and high groups. Furthermore, it seems from the analysis that the investment share of GDP adds little meaning in the process, since the majority of the coefficients related to this variable are around zero (or at least very close to zero) and insignificant. The rest of the basic variables seem to affect -mostly- randomly the indicated subsamples in all time frames. However, when we add the additional variables, we see a decent increase in the adjusted-R squared rates in the striking majority of the subsamples. Therefore, we derive better fits of the model. A major problem that we faced throughout the analysis of this study is the inaccessibility in the data regarding some social-economic variables that Barro (1991) proposes. When it is feasible, we replace a variable that we don't have access to its data, with another variable being used as a proxy. In general, the literature about cross-country growth and the law of motions that affect the growth process keeps growing in a tremendous way, especially in the modern era where authorities, institutions and governments storage data in a great scale. Despite all the efforts made to establish a general theory that could explain the growth process, it is wise to keep in mind that every country and every geographical region has its own specific peculiarities. Our Wald specification tests exhibit mixed results regarding the 1960-1985 and 1960-2015 (108 countries) time frames; thus, we can't verify the presence of multiple regimes. For the 1960-2015 but with the original sample, we purely find evidence of a single regime model existence. Specifically, regarding the 1960-1985 time period, when the control variable is per capita output and we split into two subsamples, we find evidence of a single regime model. However, the two-way split under literacy rate provides us with evidence of multiple regimes existence. Also, the tree-way split provides with the rejection of the alternative hypothesis in favor of a single regime model. Lastly, when we let the two control variables to affect each other, the outcome strengthens the aspect of multiple regimes existence. In a similar way, the 1960-2015 time period for 108 countries also provides a mixed situation, where we do not reject the single regime model hypothesis for both variables in the two-way split. As for the three-way split, null hypothesis is rejected when control variable is output per capita, but it is not rejected when we use literacy rate. Interaction between the two control variables, leads to the rejection of the single regime model.

6. Bibliography-Internet Sources

Azariades Costas, Drazen Allan. (1990, May). Threshold externalities in economic development. *The Quarterly of Journal Econometrics*.

Barro J. Robert. (1991, May). Economic Growth in a Cross Section of Countries. *The Quarterly Journal of Economics, Vol.106*, pp. 407-443.

Barro R., Lee J.W. (n.d.). *Educational Attainment for Total Population, 1950-2015*. Retrieved from http://www.barrolee.com/

Baumol William J. (1986, December). Productivity Growth, Convergence and Welfare: What the Long-Run Data Show. *The American Economic Review, Vol.76*, pp. 1072-1085.

Beylunioglu F.C., Stengos T., Yazgan M.E. (2018, April). *researchgate.net.* Retrieved from https://www.researchgate.net/profile/M-Yazgan/publication/326303731_DETECTING_CONVERGENCE_CLUBS/links/5b7db8c04585151fd127c584/DETECTING-CONVERGENCE-CLUBS.pdf

Corden, W. M. (1984). Booming Sector and Dutch Disease Economics: Survey and Consolidation. *Oxford Economic Papers*, pp. 359-380.

Corrado L., Martin R., Weeks M. (2005). Identifying and interpreting regional convergence clusters across Europe. *The Economic Journal*.

Corrado L., Stengos T., Weeks M., Yazgan M. (2018, December 27th). *ArXiv.org.* Retrieved from https://arxiv.org/pdf/1812.09518.pdf

Dufrenot G., Mignon V., Naccache T. (2012). Testing catching-up between the developing countries:“growth resistance” and sometimes “growth tragedy”. *Bulletin of Economic Research 64*, pp. 470–508.

Durlauf S.N., Johnson P.A. (1995). Multiple regimes and cross-country growth behaviour. *Journal of applied econometrics, Vol.10*, pp. 365-384.

Durlauf, S. N., Johnson P.A., Temple J.R. (2005). Growth Econometrics. *Handbook of economic growth*, pp. 555–677.

Eviews 10 package, Retrieved from https://www.eviews.com/home.html

Feenstra, Robert C., Inklaar Robert and Timmer P. Marcel. (2015). The Next Generation of the Penn World Table. *American Economic Review*, pp. 3150-3182.

Fritsche, U. and Kuzin V. (2011). Analysing convergence in europe using the non-linear single. *Empirical Economics 41*, pp. 343–369.

Government expenditure on education. (n.d.). *World Data Bank*. Retrieved from https://data.worldbank.org/indicator/SE.XPD.TOTL.GD.ZS

Hausmann, R., Pritchett L., Rodrik D. (2005). Growth accelerations. *Journal of Economic*, pp. 303–329.

Havranek, Tomas; Horvath, Roman; Zeynalov, Ayaz. (2016). Natural Resources and Economic Growht: A meta analysis. *Econstor*.

Johnson, P. C. Papageorgiou. (2017). What remains of cross-country convergence? *Journal of economic literature*.

Mankiw N. Gregory, R. David, David N. Weil. (1992, May). A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*.

Military Expenditure. (n.d.). *World Data Bank*. Retrieved from https://data.worldbank.org/indicator/MS.MIL.XPND.GD.ZS

Oded Galor. (1996, January). Convergence?: Inferences from theoretical models. *Economic Journal*.

Omerovic S., Friedl H., Grun B. (2022). *researchgate.net.* Retrieved from https://www.researchgate.net/profile/Herwig-Friedl/publication/350055365_Modelling_Multiple_Regimes_in_Economic_Growth_by_Mixtures_of_Generalised_Nonlinear_Models/links/607fbf858ea909241e1221ec/Modelling-Multiple-Regimes-in-Economic-Growth-by-Mixtures-of-

Population ages 15-64, total. (n.d.). *World Data Bank*. Retrieved from https://data.worldbank.org/indicator/SP.POP.1564.TO

Pupil-teacher ratio, primary. (n.d.). *World Data Bank*. Retrieved from https://data.worldbank.org/indicator/SE.PRM.ENRL.TC.ZS

Pupil-teacher ratio, Secondary. (n.d.). *World Data Bank*. Retrieved from https://data.worldbank.org/indicator/SE.SEC.ENRL.TC.ZS

School enrollment primary. (n.d.). *World Data Bank*. Retrieved from https://data.worldbank.org/indicator/SE.PRM.NENR

Siebert, Horst; Stolpe, Michael. (2001). Technology and economic performance in the German economy. *Econstor*.

SIPRI Military Expenditure Database. (n.d.). Retrieved from https://www.sipri.org/databases/milex/sources-and-methods

Solow M. Robert . (1956, February). A Contribution to the Theory of Economic Growth. *The Quarterly Journal of Economics, Vol 70*, pp. 65-94.

Summers R., Heston A. (1988). A new set of international comparisons of real product and prices: estimates for 130 countries. *Review of Income and Wealth*, pp. 1-26.

Tol Richard S.J. (2009). The Economic Effects of Climate Change. *Journal of Economic Perspectives*, pp. 29-51.

UNESCO, uis.unesco.org, Retrieved from
http://uis.unesco.org/?URL_ID=2867&URL_DO=DO_TOPIC&URL_SECTION=201

6. Appendix A: Sample and subsamples

The first table presents the M-R-W sample and the country classification. The second table presents the subsamples that are produced throughout the regression tree analysis in the S.N. Durlauf and P.A. Johnson framework.

Sample of M-R-W	Non-Oil	Intermediate	OECD
Algeria	1	1	0
Angola	1	0	0
Benin	1	0	0
Botswana	1	1	0
Burkina Faso	1	0	0
Burundi	1	0	0
Cameroon	1	1	0
Central African Republic	1	0	0
Chad	1	0	0
People's Rep. of Congo	1	0	0
Egypt	1	0	0
Ethiopia	1	1	0
Ghana	1	0	0
Ivory Coast	1	1	0
Kenya	1	1	0
Liberia	1	0	0
Madagascar	1	1	0

Malawi	1	1	0
Mali	1	1	0
Mauritania	1	0	0
Mauritius	1	0	0
Morocco	1	1	0
Mozambique	1	0	0
Niger	1	0	0
Nigeria	1	1	0
Rwanda	1	0	0
Senegal	1	1	0
Sierra Leone	1	0	0
Somalia	1	0	0
South Africa	1	1	0
Sudan	1	0	0
Tanzania	1	1	0
Uganda	1	0	0
Zaire (Dem. Rep. Congo)	1	0	0
Zambia	1	1	0
Zimbabwe	1	1	0
Bangladesh	1	1	0
Burma	1	1	0
Hong Kong	1	1	0
India	1	1	0
Israel	1	1	0
Japan	1	1	1
Jordan	1	1	0
Rep. of Korea	1	1	0
Malaysia	1	1	0
Nepal	1	0	0
Pakistan	1	1	0
Philippines	1	1	0
Singapore	1	1	0

Sri Lanka	1	1	0
Syrian Arab Rep.	1	1	0
Thailand	1	1	0
Austria	1	1	1
Belgium	1	1	1
Denmark	1	1	1
Finland	1	1	1
Germany	1	1	1
Greece	1	1	1
Ireland	1	1	1
Italy	1	1	1
Netherlands	1	1	1
Norway	1	1	1
Portugal	1	1	1
Spain	1	1	1
Sweden	1	1	1
Switzerland	1	1	1
Turkey	1	1	1
United Kingdom	1	1	1
Canada	1	1	1
Costa Rica	1	1	0
Dominican Rep.	1	1	0
El Salvador	1	1	0
Guatemala	1	1	0
Haiti	1	1	0
Honduras	1	1	0
Jamaica	1	1	0
Mexico	1	1	0
Nicaragua	1	1	0
Panama	1	1	0
Trinidad & Tobago	1	1	0
United States	1	1	1

Argentina	1	1	0
Bolivia	1	1	0
Chile	1	1	0
Colombia	1	1	0
Ecuador	1	1	0
Paraguay	1	1	0
Peru	1	1	0
Uruguay	1	1	0
Venezuela	1	1	0
Australia	1	1	1
Indonesia	1	1	0
New Zealand	1	1	1
Papua New Guinea	1	0	0

Source: Table, pages 434,435,436 (Appendix) from M – R – W (1992).

Terminal node 1	**Terminal node 2**	**Terminal node 3**	**Terminal node 4**
Burkina Faso	Algeria	Madagascar	Austria
Burundi	Angola	South Africa	Belgium
Ethiopia	Benin	Hong Kong	Denmark
Malawi	Cameroon	Israel	Finland
Mali	Central African Rep.	Japan	France
Mauritania	Chad	Korea	Germany
Niger	People’s Rep. Congo	Malaysia	Italy
Rwanda	Egypt	Philippines	Netherlands
Sierra Leone	Ghana	Singapore	Norway
Tanzania	Ivory Coast	Sri Lanka	Sweden
Togo	Kenya	Thailand	Switzerland
Uganda	Liberia	Greece	United Kingdom
Zaire	Morocco	Ireland	Canada
Burma (Myanmar)	Mozambique	Portugal	Trinidad & Tobago
	Nigeria	Spain	U.S.A.

	Senegal	Costa Rica	Argentina
	Somalia	Dominican Rep.	Chile
	Sudan	El Salvador	Uruguay
	Tunisia	Jamaica	Venezuela
	Zambia	Mexico	Australia
	Zimbabwe	Nicaragua	New Zealand
	Bangladesh	Panama	
	India	Brazil	
	Jordan	Colombia	
	Nepal	Ecuador	
	Pakistan	Paraguay	
	Syria	Peru	
	Turkey		
	Guatemala		
	Haiti		
	Honduras		
	Bolivia		
	Indonesia		
	Papua New Guinea		

Source: Table IV, page 374 from Durlauf and Johnson (1995).

7. Appendix B: Confidence Intervals

In this Appendix we will present the confidence intervals of subsamples when the two control variables interact for all time frames.

- ***1960-1985 Low Group*** $(LR_{i,1960}$ *and* $(\frac{Y}{L})_{i,1960})$

Variable	Coefficient	90% CI Low	90% CI High	95% CI Low	95% CI High	99% CI Low	99% CI High
LOGY1960	-0.395221	-0.567970	-0.222472	-0.603018	-0.187423	-0.674800	-0.115642
LOGI	0.027744	-0.177090	0.232578	-0.218648	0.274135	-0.303761	0.359249
LOGNGD	0.020267	-0.466480	0.507013	-0.565233	0.605767	-0.767488	0.808022
LOGSCHOOL	0.317746	0.171670	0.463822	0.142033	0.493459	0.081335	0.554157
C	4.368317	2.359397	6.377236	1.951819	6.784815	1.117065	7.619569

Variable	Coefficient	90% CI Low	90% CI High	95% CI Low	95% CI High	99% CI Low	99% CI High
LOGY1960	-0.379234	-0.532846	-0.225623	-0.564447	-0.194022	-0.629906	-0.128563
LOGI	0.050810	-0.064046	0.165666	-0.087674	0.189294	-0.136618	0.238238
LOGNGD	-0.086105	-0.775152	0.602942	-0.916902	0.744692	-1.210529	1.038319
LOGSCHOOL	0.057824	-0.114492	0.230140	-0.149941	0.265589	-0.223370	0.339018
AFRICA	-0.254922	-0.488090	-0.021754	-0.536057	0.026213	-0.635418	0.125574
G	7.504109	5.044463	9.963755	4.538466	10.46975	3.490326	11.51789
LATAMER	0.326225	-0.036988	0.689437	-0.111708	0.764157	-0.266486	0.918935
PRIM70	0.491424	0.037650	0.945199	-0.055701	1.038550	-0.249070	1.231919
STTEAPRI70	-0.592530	-1.254291	0.069232	-1.390428	0.205369	-1.672428	0.487368
STTEASEC70	0.752204	-0.228110	1.732518	-0.429780	1.934188	-0.847526	2.351933
C	2.878419	0.565154	5.191684	0.089270	5.667568	-0.896491	6.653329

- ***1960-1985 High Group (*** $LR_{i,1960}$ ***and*** $(\frac{Y}{L})_{i,1960}$***)***

Variable	Coefficient	90% CI Low	90% CI High	95% CI Low	95% CI High	99% CI Low	99% CI High
LOGY1960	-0.165001	-0.368112	0.038110	-0.408936	0.078934	-0.491911	0.161910
LOGI	0.393707	0.150928	0.636487	0.102130	0.685284	0.002950	0.784465
LOGNGD	-0.350919	-0.982257	0.280419	-1.109153	0.407315	-1.367069	0.665231
LOGSCHOOL	0.182817	-0.032605	0.398238	-0.075904	0.441537	-0.163909	0.529542
C	2.051086	-0.594621	4.696793	-1.126397	5.228569	-2.207229	6.309401

Variable	Coefficient	90% CI Low	90% CI High	95% CI Low	95% CI High	99% CI Low	99% CI High
LOGY1960	-0.180100	-0.425506	0.065307	-0.475296	0.115096	-0.577268	0.217068
LOGI	0.304834	0.043791	0.565878	-0.009171	0.618839	-0.117641	0.727309
LOGNGD	-0.048730	-0.630950	0.533489	-0.749073	0.651613	-0.990999	0.893539
LOGSCHOOL	0.198547	-0.077219	0.474314	-0.133168	0.530262	-0.247756	0.644850
AFRICA	-0.383173	-0.696637	-0.069708	-0.760235	-0.006111	-0.890487	0.124142
G	-1.987828	-4.591494	0.615837	-5.119738	1.144081	-6.201623	2.225966
LATAMER	-0.312563	-0.746427	0.121301	-0.834451	0.209325	-1.014732	0.389606
PRIM70	-0.192661	-0.476211	0.090890	-0.533739	0.148418	-0.651561	0.266240
STTEAPRI70	-0.112684	-0.830443	0.605074	-0.976065	0.750697	-1.274311	1.048943
STTEASEC70	0.505539	-0.481952	1.493030	-0.682299	1.693377	-1.092625	2.103703
C	3.306704	0.169875	6.443534	-0.466540	7.079948	-1.769967	8.383376

- ***1960-2015 Low Group ($LR_{i,1960}$ and $(\frac{Y}{L})_{i,1960}$) for 108 countries***

Variable	Coefficient	90% CI Low	90% CI High	95% CI Low	95% CI High	99% CI Low	99% CI High
LOGY1960	-1.129016	-1.542792	-0.715240	-1.627076	-0.630956	-1.800259	-0.457773
LOGI	0.353940	-0.275518	0.983397	-0.403735	1.111614	-0.667190	1.375069
LOGNGD	2.423003	-1.567352	6.413359	-2.380165	7.226172	-4.050297	8.896304
LOGSCHOOL	0.654820	0.399255	0.910386	0.347197	0.962444	0.240232	1.069409
C	17.86173	4.323719	31.39974	1.566102	34.15736	-4.100126	39.82358

Variable	Coefficient	90% CI Low	90% CI High	95% CI Low	95% CI High	99% CI Low	99% CI High
LOGY1960	-0.972172	-1.229221	-0.715123	-1.282432	-0.661912	-1.393220	-0.551124
LOGI	0.080033	-0.272418	0.432484	-0.345379	0.505444	-0.497285	0.657351
LOGNGD	2.028699	-1.134268	5.191666	-1.789029	5.846427	-3.152266	7.209665
LOGSCHOOL	0.319577	0.035636	0.603518	-0.023143	0.662296	-0.145521	0.784675
AFRICA	-0.633348	-0.963877	-0.302820	-1.032299	-0.234397	-1.174757	-0.091940
G	10.67373	5.121128	16.22633	3.971692	17.37577	1.578523	19.76894
LATAMER	-0.328249	-0.829729	0.173231	-0.933539	0.277041	-1.149677	0.493179
PRIM70	0.177996	-0.412076	0.768067	-0.534225	0.890217	-0.788546	1.144538
STTEAPRI70	-0.469211	-0.956604	0.018183	-1.057499	0.119077	-1.267565	0.329144
STTEASEC70	0.730365	-0.674731	2.135461	-0.965598	2.426328	-1.571194	3.031924
C	14.09289	4.215759	23.97003	2.171108	26.01468	-2.085934	30.27172

- ***1960-2015 High Group ($LR_{i,1960}$ and $(\frac{Y}{L})_{i,1960}$) for 108 countries***

Variable	Coefficient	90% CI Low	90% CI High	95% CI Low	95% CI High	99% CI Low	99% CI High
LOGY1960	-0.583351	-0.833263	-0.333440	-0.883494	-0.283209	-0.985589	-0.181114
LOGI	0.267075	-0.086550	0.620700	-0.157627	0.691777	-0.302091	0.836241
LOGNGD	-0.769269	-1.904107	0.365568	-2.132205	0.593666	-2.595812	1.057273
LOGSCHOOL	0.938140	0.449752	1.426528	0.351588	1.524692	0.152070	1.724209
C	5.902673	2.232445	9.572902	1.494744	10.31060	-0.004628	11.80997

Variable	Coefficient	90% CI Low	90% CI High	95% CI Low	95% CI High	99% CI Low	99% CI High
LOGY1960	-0.576003	-0.787628	-0.364379	-0.830563	-0.321443	-0.918498	-0.233509
LOGI	0.301242	-0.059586	0.662071	-0.132793	0.735277	-0.282726	0.885210
LOGNGD	-0.198787	-1.414542	1.016968	-1.661200	1.263626	-2.166375	1.768802
LOGSCHOOL	1.093000	0.334187	1.851812	0.180236	2.005764	-0.135069	2.321069
AFRICA	-0.964491	-1.340114	-0.588868	-1.416322	-0.512660	-1.572402	-0.356580
G	-0.419824	-4.608499	3.768850	-5.458316	4.618667	-7.198810	6.359161
LATAMER	-0.020752	-0.555316	0.513813	-0.663771	0.622268	-0.885896	0.844392
PRIM70	-0.238377	-0.876547	0.399793	-1.006022	0.529268	-1.271197	0.794443
STTEAPRI70	-1.059767	-2.306593	0.187059	-2.559555	0.440020	-3.077641	0.958106
STTEASEC70	0.910013	-0.788208	2.608235	-1.132751	2.952778	-1.838403	3.658430
C	8.001328	4.254169	11.74849	3.493929	12.50873	1.936895	14.06576