Falling Behind or Catching up - Structural Break Story of Africa's Convergence

Fan Zhuo
Boston University

Aparna Dutta
Boston University

February 27, 2014

Abstract

In this empirical paper we show that the "iron-law" of convergence (2%) stills holds for the world. We detect a break in Africa's convergence rate and argue that Africa was not really converging before 2000. The world convergence rate before 2000 was mostly driven by Asian and Latin American countries. We show that the recent institutional and infrastructural development has led the African countries on the path of "catching up". We use LASSO to select the variables and use the double selection method to estimate the treatment effect in the partially linear model. We also compare LASSO variable selections with GBM and Random Forest variable selections.

Keywords: Structural break, LASSO, data mining, Africa, world convergence.

1 Introduction

There has been a surge of empirical literature on the topic of growth and specifically convergence in the past two decades. With data, these papers like to test whether real per capita income differences between rich and poor countries reduce in the long run. According to the neoclassical theory, as developed by Solow (1956), the poorer countries have capital-labor ratio much lower than the optimum, making their rate of return higher on fixed investment, which leads to a faster growth rate compared to rich countries - essentially the convergence hypothesis says that poor countries "catch up" with the per capita income levels of the rich countries.

In the seminal contribution, Baumol (1986) uses Maddison's data to empirically calculate the convergence rate. Later Barro (1991), Barro and Sala-i-Martin (1991 and 1992) find roughly 2% convergence rate and recently Barro (2012) estimates the convergence rate conditioned on a set of variables . Barro (2012) tries to test the iron-law of convergence (2%) rate for a panel of countries for five year intervals ending 2009. He particularly looks at cases when to include country fixed effects and suggests that unless the data spans for a century, panel models predicting convergence rate should not include fixed effects.

The available empirical evidence does not support the universal convergence hypothesis unidirectionally. There is a lot of evidence suggesting diverging productivity levels and real per capita incomes between the group of advanced industrialized economies on the one hand and the developing countries on the other. Pritchett (1997) and Jones (1997) show the diverging absolute convergence. Paul Collier (2008) in his book raises the question of how the countries are actually diverging in growth rates since end of 1980's. This book is one of the most important motivations for this paper. We empirically justify Collier's point. Khan (2011) excludes the Sub-Saharan African countries and shows beta-convergence. In this paper we analyze five group of countries - (1) World (all the countries taken together) (2) Africa (countries in the African continent) (3) Asia and other rich developed countries (countries in the Asian continent, and US, Canada, Europe, Australia and New Zealand) (4) Africa and other rich developed rich countries (countries in the African continent, and US, Canada, Europe, Australia and New Zealand) (5) World without Africa (all the countries except the countries in the African continent). We calculate the convergence rate in each case.

There are numerous papers that look into structural break in the macroeconomic variables and few specially in convergence rate. Attfield (2003) looks into structural break in output growth for EU. He found shift in the mean in 1993Q3. We use the Bai and Perron (1998, 2003) methodology to test for structural break in Africa's growth, per capita income and convergence gap. We divide the total time period based on the results from structural break and estimate convergence rate separately for each sub-period. Arbache and Page (2009) try to argue that post 1995 Africa has seen growth but this is a "fragile growth". We show that both the resourceful and non-resourceful countries in Africa converge post 1995.

We calculate the convergence rate for each subperiod using double selection methodlogy for estimation of treatment effect in a partially linear model. We start from a big list of probable covariates and use Least Absolute Shrinkage and Selection Operator (LASSO) to select these variables. Tibshirani (1996) proposed methodolgies of shrinkages in dimensions through LASSO. Belloni et al. (2011a, 2011b and 2012) have developed various techniques for high dimensional sparse models. Using the Barro-Lee data they apply LASSO and post-LASSO techniques to estimate the convergence rate at 3%. Barro and other authors mostly use economic reasoning to choose the instruments and variables to condition-on. We use Belloni et al.'s methodology in our paper. We also compare our results with GBM and Random Forest variable selection results.

Section (2) discusses the theory behind the convergence rate which we estimate in this paper. In section (3) we discuss the main empirical results of the paper. In the final section of this paper we discuss the reasons on why African countries are converging post 1995. The focus is on institutional and infrastructural variables; a within growth policy in the countries.

2 Conditional Convergence

In this section we briefly describe the theoretical neoclassical growth model. The production function is given as:

$$Y = A \cdot F(K, L), \tag{1}$$

where $F(\cdot)$ satisfies the constatnt returns to scale in capital, K and labor, L assumption. Let $y := \dfrac{Y}{L}$ be the output per worker and $k := \dfrac{K}{L}$ be the capital per worker. The relationship is then given by:

$$y = f(k) \tag{2}$$

Let us assume that $f(k)$ satisfies the usual neoclassical properties. Assuming the economy to be closed, savings equals investments. Initially we assume no government sector as well, however as an extension government purchases and taxes are allowed. The labor force is assumed to be fully employed and corresponds to L in equation (1). Given all assumtions, y and k can be interpreted as per capita quantities.

The labor force is related to the population in a fixed way. The population is assumed to grow at a rate n. in the baseline model n is constant over time within an economy but can differ across economies. Using the Malthusian framework, population growth rate can be endogenized. Barro-Sala-i-Martin (2004) allow for choices of fertility rate for a given mortality rate. An economy with lower n at a given y tends to grow faster. Child rearing costs and preferences can shift the path of n when considered in relation to y.

The gross savings rate (which inturn equals to the ratio of gross investment to GDP) is represented by s. Extending the baseline model where s is constant (Solow-Swan framework) to endogenized savings rate (Ramsey-Cass-Koopmans framework) - the representative household determines the optimal savings rate at each point in time based on an iso-elastic, time-additive utility function. s then varies over time within economies along the transition path to steady state. Depending on the preferences and other parameters, across economies, the path of s can shift when considered in relation to y. An economy with higher s at a given y tends to grow faster.

In the baseline model, the productivity factor A is assumed to rise over time due to exogenous technical progress. The usual way to represent the technical progress is to introduce labor-augmenting technology such that effective labor $\hat{L} := Le^{xt}$ and effective output and capital per capita is given by $\hat{y} := \dfrac{Y}{\hat{L}}$ and $\hat{k} := \dfrac{K}{\hat{L}}$. x is the rate of technological progress.

In extensions, differences in A (within or across economies) can be endogenized by quality of institutions, efficiency of tax system, clear definitions of property rights and successful R& D activities. Thus endogenous technical change leads to endogenizing of the productivity factor. Further it can be argued that countries with higher levels of school attainment and better institutional quality will have higher technological diffusion.

Borrowing from Barro (2011), we can summarize the above discussion in the equation below:

$$Dy_{it} = \Phi(\underset{(-)}{y_{it}}, \underset{(+)}{s_{it}}, \underset{(-)}{n_{it}}, \underset{(+)}{A_{it}}, ...) \tag{3}$$

where Dy_{it} is the growth rate of real per capita GDP for country i at time t. The negative sign under y_{it} reflects convergence, conditional on the other variables. This is the framework for conditional beta-convergence.

There are basically two assumptions we can consider before taking a closer look at the beta-convergence for the world. The first assumption is that the steady state is the same for all the countries and they converge to this steady state at different paces, which are lower for rich countries and higher for poor ones. This is called absolute beta-convergence. In conditional beta-convergence countries will still converge, however the long-run steady states vary across countries and depend on a series of factors, such as a country's instituions and education level.

The model we used to measure the conditional beta-convergence has the form:

$$\ln(\frac{y_{i,t}}{y_{i,t-1}}) = \alpha + \beta \ln(y_{i,t-1}) + \gamma' z_{i,t-1} + \varepsilon_{i,t}, \tag{4}$$

where $y_{i,t}$ is the GDP per capital in country i at time t, $z_{i,t}$ are the other factors that influence the GDP growth rate (as described by equation 3) and $\varepsilon_{i,t}$ is an error term. The test of the beta-convergence hypothesis is equivalent to show that β is negative and significant, conditional on different factors included in $z_{i,t}$.

3 Empirical Results

In this section we present the empirical results. Our first aim was to check if the 2% "iron-law" of convergence still holds for the world, when recent data is used. Next we wanted to study the African countries and understand whether they are "catching-up". The first subsection discusses the data and gives some descriptive graphs. The next subsection talks about structural change in Africa's real per capita GDP. We use LASSO to select the variables and compare this selection with other data mining techniques - the subsection on methodology discusses each of them. And finally the key results are presented in the last subsection.

3.1 Data

Using Penn World Tables, we looked into the real per capita GDP (chained PPPs in mil 2005 US dollars) for all the countries in the world from 1960 - 2010. For Africa, we had real per capita income data for 41 countries in 1960 - 1970 and 48 countries in 1970 - 2010. Figure 1 gives the average real per capita income of the African continent. The mean is simple average and is not weighted by population.

Looking at figure 1, we can say that post 1992 - 1993 there has been an upward trend in the real per capita GDP for Africa. In the period 1975 - 1992 the real per capita GDP actually went down.

In figure 2 we look at the yearly growth rate of real per capita GDP for the African continent. We also fit a polynomial to estimate the growth trend. We can draw a similar conclusion that post

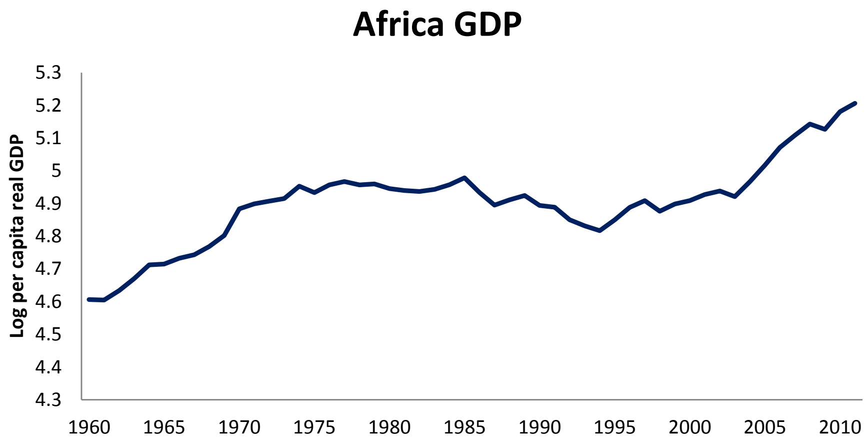

Figure 1: Real Per Capita GDP over the years

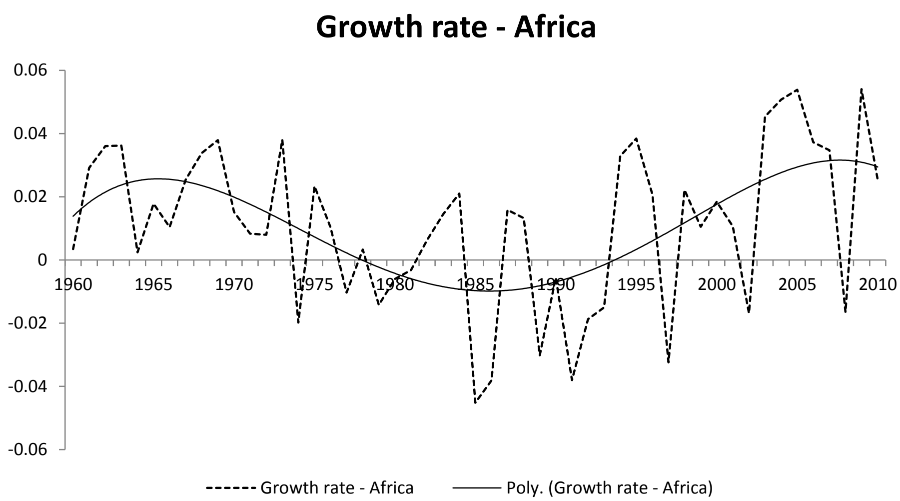

Figure 2: Growth in Real Per Capita GDP over the years

1993 Africa's real per capita growth has been positive and the trend can be clearly distinguished from previous years 1978 - 1992.

The basic convergence potential of countries is measured by the income gap that separates them from rich countries. Indexing US as the rich country, figure 3 gives the convergence gap of Asia and Africa with respect to US. The figure gives the regions real per capita GDP as percentage of US's real per capita GDP (essentially if we had to plot US's real per capita income it would be 100).

Clearly, Asia has been steadily closing the gap since the late 70's. Africa has only recently experienced a small closing in the difference of incomes. In the period post 2000, Africa looks like having a turn towards "catching-up".

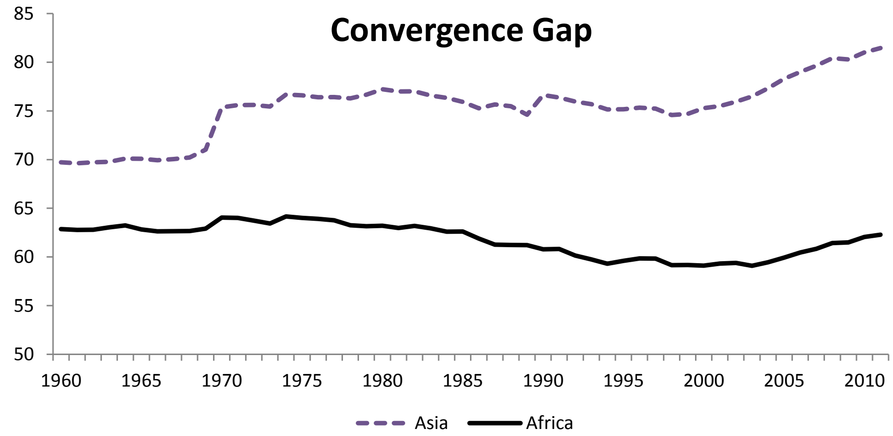

Figure 3: Convergence gaps for Asia and Africa in 1960-2010

There has been argument (Arbache-Page, 2009) that only the "resourceful" African countries have seen the recent trend towards higher real per capita GDP. We look into the "non-resourceful" (as classified by the World Bank) African countries. Figure 4 plots the average real per capita GDP of these countries. Very similar to figure 1, this graph also shows a rising trend in the real per capita post early 90's and definitely in the recent years.

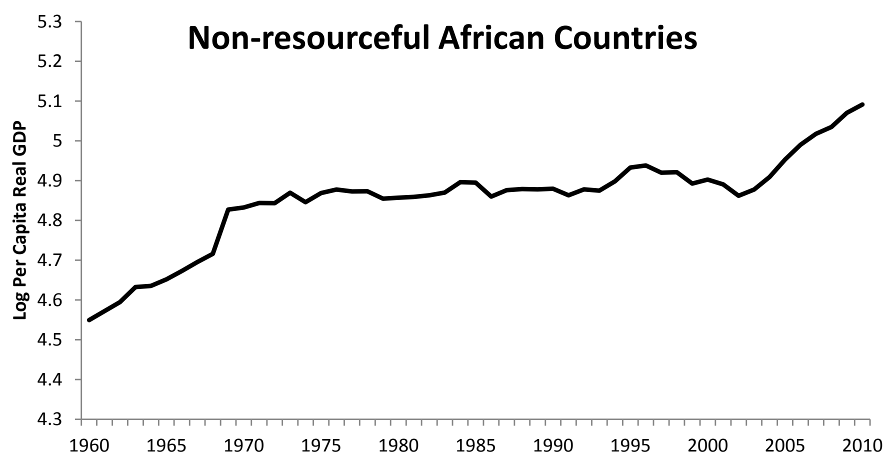

Figure 4: Real Per Capita GDP in 1960-2010

3.2 Structural Change

We apply the methods of Bai and Perron (1998, 2003) to estimate multiple break dates for Africa's income, growth and convergence gap over 1960-2011 without prior knowledge of when those breaks occur. When we look into figures 1-4 we can say that there has been a change in mean over time and

Table 1: Break date estimates

	Pre 1975	1975-2000	2001-2010
Real Per Capita GDP	1969	1986	2003
Growth of Real Per Capita GDP	1974		2003
Convergence Gap	1985	1993	2003
Non-Resourceful Africa GDP	1969		2003

Note: Break dates are estimted using the Bai and Perron (1998, 2003) methodology. Gauss code is made available by Bai and Perron. We report the break dates as chosen by the BIC criterion as explained in Bai and Perron (2003). We imposed a maximum number of 3 breaks; the sample runs from 1961 through 2010.

we want to test this hypothesis using structural break tests. Bai and Perron (1998, 2003) provide a least-squares based algorithm for estimating the break dates, significance tests and suggestions for how to interpret the various tests. We used the BIC to estimate the break dates. We refer the reader to Bai and Perron (1998, 2003) for further details. The column groupings in Table -1 (e.g . Pre-75, 75 to 2000, etc.) are based on the results (where the clustering seemed to occur) and are meant simply for ease of reading the table.

If we look at the break date estimates and compare it with the graphs in the previous section, we can say that the post 2003, there has been a clear growth in real per capita income; post 2003 the convergence gap (with US) seems to be decreasing, even though slightly. We look into a formal model to measure conditional convergence in the next two subsections - but based on the break date estimates we can conjecture that the convergence coefficient will be significant post 2003.

We also look into the real per capita GDP for the countries in Africa which are considered non-resourceful and found structural break in 2003 (among recent years). This evidence suggests that even the countires in Africa which are not necessarily resource abundant have seen a change in the real per capita GDP in the recent years. It is argued that Africa's growth is led by resourceful-countries which see high exports rather than a self-contained growth within Africa. In this paper we have tried to argue that the recent growth (diminishing convergence gap) is due to institutional and infrastructural dvelopment within the countries. We discuss this in section 4.

3.3 Results - absolute convergence, sigma-convergence and panel methods

Two countries exhibit convergence if the poorer country with lower initial income grows faster than the other (beta-convergence). Absolute convergence happens if countries' per capita income converges to a steady state value. The absolute beta-convergence is estimated by cross-country regression of per capita income growth on initial per capita income. The "iron-law" of convergence is 2% implying countries eliminate gaps in initial real per capita income at a rate of 2% per year. We have real income data for 142 countries in 1970, with years that increases to 165 countries in 2010. The dependent variable in the absolute convergence regression is each country's growth rate of real per capita GDP over eight 5-year intervals from 1970-1975 to 2005-2010. The only independent variable is initial real per capita GDP. We present the absolute convergence results in table 2. Based on the structural change results we grouped 1970-1995 and 1995-2010 and also estimated the absolute beta-convergence.

Table 2: Absolute beta-convergence

	Absolute Convergence	p-value
1970-1975	0.004	0.19
1975-1980	0.006	0.07
1980-1985	-0.01	0
1985-1990	0.011	0
1990-1995	0.002	0.6
1995-2000	0.001	0.73
2000-2005	-0.003	0.36
2005-2010	-0.006	0
1970-1995	0.002	0.42
1995-2010	-0.003	0.07

Absolute convergence implies a tendency towards the equalization of per capita incomes, i.e. "catch up" growth. For the period 1995 - 2010, the coefficient has the right sign (negative) and is significant at 10%. However the coefficient is very small and not as high as 2%. So the analysis does give us some evidence of economies converging over 1995 - 2010, but the rate is very small.

When the dispersion of real per capita GDP across a group of economies falls over time, there is said to be sigma-convergence. The progress in sigma-convergence is dependent on both differential rates of growth between poorer and richer countries, and also the size of the initial income gap. We measure the standard deviation of the countries' (all countries for which data is available) real per capita GDP over 5 year intervals and report the results in table 3.

Table 3: Sigma-convergence

	Standard deviation
1970	1.1
1975	1.15
1980	1.2
1985	1.16
1990	1.22
1995	1.28
2000	1.31
2005	1.32
2010	1.29

Beta-convergence is said to be necessary but not a sufficient condition for sigma-convergence. Beta-convergence implies the existence of a longer-term catch-up mechanism, i.e. forces which work towards the narrowing of income differences across countries. However there can be temporary shocks which might affect the short-run growth. Thus having beta-convergence might not be reflected in changes of dispersion of income levels. In our analysis we find the dispersion in 2000 is higher compared to 2010. So by just comparing year 2000 and year 2010 we might conclude that there is sigma-convergence in the world, however this might be too myopic conclusion.

Barro (2011) discusses panel data methodologies to estimate conditional beta-convergence. He concludes by saying that inclusion of country fixed effects produces much higher convergence rates and eliminates statistically significant effects of the institutional measures on economic growth. He further states that there are econometric issues associated with inclusion of country fixed effects in short panels. We start with a balanced panel of 66 countries and eight 5-year intervals (1970-75, 75-80..2005-2010) and estimate the conditional beta-convergence coefficient. We use similar covariates as used by Barro (2011). The data sources are discussed in the appendix.

In table 4 we find similar results as Barro (2011). The conditional beta-convergence coefficient without any country fixed effects is -0.017 implying convergence rate of 1.7%. With fixed effects the convergence rate is 6% which clearly is too high and as Barro suggests, it is overestimated. Variables like fertility rate, terms of trade, life expectancy at birth lose significance when fixed effects are considered.

Table 4: Growth-Rate Regressions for Cross-Country Panel

	No Fixed Effects	With Fixed Effects
Log Per Capita GDP	-0.0171^{**} (0.0033)	-0.0588^{**} (0.0064)
Female School Years	0.0001 (0.0012)	0.0039 (0.0037)
Fertility rate	-0.0088^{**} (0.002)	0.0014 (0.0035)
Life Expectancy at Birth	0.0011^{**} (0.0004)	0.0013 (0.0007)
Inflation rate	0.0000 (0)	0.0000 (0)
Political Status	-0.0233 (0.0122)	-0.0202 (0.0132)
Exchange Rate	0.0000 (0)	0.0000 (0)
Investment share	-0.0161 (0.0211)	0.0281 (0.0311)
Government expenditure share	-0.0233 (0.0221)	-0.019 (0.0316)
Openness ratio	0.0112^{*} (0.0052)	0.0066 (0.0104)
Terms of trade	-0.0188^{**} (0.0048)	-0.0151 (0.0052)
Political rights	0.0275^{*} (0.0133)	0.0206^{**} (0.0144)
Civil Liberty	0.0009 (0.0143)	-0.0039 (0.0173)

Note: *p<.05, **p<.01. Time effects are included in the regressions

Next we run cross-country regression for the African continent. Based on the structural test results we divide the period into five and three 5-year intervals, so 1970-75, 75-80,.., 90-95 as one set (henceforth referred as Pre-2000) and 1995-2000, 2000-2005 and 2005-2010 as another set (henceforth referred as Post-2000). For pre-2000 regression there were only 15 countries for whom balanced panel was available. Most of these countries were resourceful countries. In table 5, the 4% convergence rate pre-2000 is biased for the countries in the sample. However in post-2000 regression we are looking at 35 countries with balanced panel. An important feature about post-2000 regression result is also the significance of "average years of schooling" variable. In the next few sections we argue that the recent growth in Africa can be attributed to institutional and infrastructural development and this gives some evidence towards it.

We conclude this section by emphasizing on the results - firstly we find that the conditional convergence rate is around 2% for the world. Secondly, looking at absolute convergence and sigma-convergence we might say that only in recent years are the countries moving to the one steady state path. Thirdly, even though the world is converging at 2%, we need to analyze Africa's case a little more.

Table 5: Growth-Rate Regressions for Cross-Country Panel - Africa

	Pre-2000	Post-2000
Log Per Capita GDP	-0.0429^{*} (0.0171)	-0.0459^{**} (0.0094)
Female School Years	-0.0021 (0.0056)	0.0095^{**} (0.0032)
Fertility rate	-0.0136 (0.007)	-0.0035 (0.0064)
Life Expectancy at Birth	0.005^{**} (0.0018)	0.0018^{*} (0.0007)
Inflation rate	-0.0006^{*} (0.0002)	-0.0001 (0.0001)
Political Status	0.0264 (0.0344)	-0.0214 (0.0287)
Exchange Rate	0.0001 (0)	0 (0)
Investment share	-0.0776 (0.0487)	0.0622 (0.0563)
Government expenditure share	0.0153 (0.0486)	0.0841 (0.0669)
Openness ratio	-0.029 (0.0293)	0.0066 (0.0178)
Terms of trade	0.0006 (0.0185)	-0.1077^{*} (0.0496)
Political rights	0.0077 (0.0398)	0.0473 (0.035)
Civil Liberty	-0.0124 (0.0502)	-0.0363 (0.0446)

Note: *p<.05, **p<.01. Time effects are included in the regressions

3.4 Variable Selection Methodology

Our aim is to analyze Africa's conditional convergence rate pre-2000 and post-2000. Traditional panel methodologies do not give the option to select from a wide range of variables, in the last section we used only those as used by Barro (2011). Also with only 15 countries having a balanced panel - pre-2000 we might be analyzing a self-selected sample. In this section, we apply double selection approach as in Belloni et al (2011a). In this method we follow 3 steps (please refer to equation 4 for notations):

1. Selecting controls $z_{i,t-1}$'s that predict $y_{i,t-1}$.
2. Selecting controls $z_{i,t-1}$'s that predict $(\frac{y_{i,t}}{y_{i,t-1}})$.
3. Run Ordinary Least Squares (OLS): $(\frac{y_{i,t}}{y_{i,t-1}})$ on $y_{i,t-1}$ and the union of controls selected in the two previous steps.

This method was suggested in Leeb and Pötscher (2008) to control the omitted variable bias. In our model, this double selection approach is necessary as there are covariates in $Z_{i,t-1}$ that are correlated with $y_{i,t-1}$ and do not get selected in step 2. For selecting variables we apply LASSO methodology. In a selection and estimation problem like ours OLS is usually criticized on three grounds. Firstly, *almost sparsity* - in a situation when $p >> n$ there clearly is curse of dimensionality and the degrees of freedom is insufficient but also in the case where p is same as n we need a methodology which can zero out the effects of some of the variables at front. Secondly, *interpretation* - with a large number of availabe predictors, we would like determine a smaller subset but with strongest effects. Thirdly, the *prediction accuracy* - OLS estimator is best linear unbiased but when there are many correlated variables it has low bias and large variance. Prediction accuracy can be sometimes improved by shrinking or setting some of the coefficients to 0. By doing so the bias is slightly sacrificed for a much lower variance which overall improves the variance accuracy.

In a usual regression setting, suppose we have data (x^i, y_i), $i = 1, 2, ..., N$, where $x^i = (x_{i1}, ..., x_{ip})^T$ and y_i are the regressors and response for the ith observation. OLS estimates are obtained by minimizing the residual squared error. LASSO shrinks the coefficients of some of the covariates to 0. Formally, letting $\hat{\beta} = (\hat{\beta}_1, ..., \hat{\beta}_p)^T$, the lasso estimate $(\hat{\alpha}, \hat{\beta})$ is defined as:

$$(\hat{\alpha}, \hat{\beta}) = argmin \left[\sum_{i=1}^{N} \left(y_i - \alpha - \sum_j \beta_j x_{ij} \right)^2 \right] \text{ subject to} \sum_j |\beta_j| \leq t \tag{5}$$

Here $t \geq 0$ is the tuning parameter.

We use LASSO to select the variables in steps 1 and 2 and then run OLS in step 3. To make a comparison in terms of selection methodlogy - we also use Random Forest and Gradient Boosting Methodlogy to select the variables in steps 1 and 2.

Random Forest (RF) is an unsupervised learning mechanism. It is based on building a large collecion of decorrelated trees and then averaging them. Regression trees sub-divide or partition the data space (usually training-data) into smaller regions. They again partition the sub-divisons to get simple regions, in these final regions (terminal nodes or leaves) simple model is fit. The regression model predicts Y based on an average from all the final regions. The partitions are done using greedy algorithm (locally optimal strategy at each step with the hope of finding global optimum). In Random Forest before each split, $m \leq p$ of the input variables at random are selected as candidates for splitting. At each split in each tree, the improvement in the split-criterion is the importance measure attributed to the splitting variable, and is accumulated over all the trees in the forest separately for each variable. This is how variable importance is measured.

Gradient Boostng is a machine learning technique to solve regression problems. GBM builds the model stagewise and it generalizes them by allowing optimization of an arbitrary differentiable loss function. The goal is to find an approximation to $\hat{F}(x)$ to a function $F^*(x)$ that minimizes the expected value of a specified loss function. It starts by considering a constant function $F_0(x)$ and then expanding in a greedy way. It applies the steepest descent step to the minimization problem in order to choose the best f . It is an ensemble of weak prediction model, usually decision trees. As in Random Forest, the squared relative importance of variable x_l is the sum of such squared improvements over all internal nodes for which it was chosen as the splitting variable.

We use LASSO, RF and GBM to select the variables in step 1 and 2 and then use OLS to get the treatment and marginal effects.

3.5 Results - conditional convergence

Conditional convergence, allows each country to have a different level of per capita income towards which it is converging. This implies that each country is converging to its own steady state and that in the long run all the countries will have the same growth rates. We use the three steps as listed in the previous section to estimate the conditional convergence coefficient for the different groups of countries.

In the appendix (table 9) we provide the correlation matrix for a subset of variables (the data is for the world in year ??). The correlation among the variables is stronger for African countries. This provides evidence in why OLS without model selection might give a large variance (many correlated variables).

We next present the conditional convergence coefficients for each selection criterion we used.

Table 6: Conditional beta-convergence coefficient, Pre-2000

	LASSO	GBM	RF
World	-0.016^* (0.006)	-0.02^* (0.0057)	-0.017^* (0.0064)
Africa	-0.004 (0.0161)	-0.016 (0.0131)	-0.016 (0.0121)
Africa and Non-African Rich Countries	-0.004 (0.0075)	-0.013 (0.0073)	-0.006 (0.008)
World without Africa	-0.026^* (0.0064)	-0.027^* (0.0058)	-0.023^* (0.0063)
Asia and Non-African Rich Countries	-0.025^* (0.0082)	-0.028^* (0.0079)	-0.026^* (0.0082)

Note: *p<.05

In the table non-African rich countries mean Europe, US, Canada, Australia, New Zealand. From table (6) we can point a few things - firstly GBM gives the least variance for the beta-coefficient. Secondly, as already said in previous sections that the world is converging at a rate of around 2%. But different regions are converging at different rates. Specifically it looks like Africa was not converging before 2000. When we looked into the graphs of convergence gap between US - Asia and US - Africa we had concluded saying that the Asian countries have achieved to reduce the gap before 2000. The results in table (6) reconfirm this. Thirdly, re-emphasizing African countries do not look like converging before 2000. This is in line with the Bottom Billion book's argument saying that some countries were "trapped" in the late 80's period.

Next, we look into the post 2000 period and present the results in table (7).

Table 7: Conditional beta-convergence coefficient, Post-2000

	LASSO	GBM	RF
World	-0.014^{*} (0.0049)	-0.019^{*} (0.0051)	-0.018^{*} (0.0052)
Africa	-0.044^{*} (0.0161)	-0.058^{*} (0.0165)	-0.051^{*} (0.0205)
Africa and Non-African Rich Countries	-0.032^{*} (0.0088)	-0.031^{*} (0.0095)	-0.031^{*} (0.0105)
World without Africa	-0.009^{*} (0.0046)	-0.011^{*} (0.0063)	-0.014^{*} (0.0072)
Asia and Non-African Rich Countries	-0.013^{*} (0.006)	-0.015^{*} (0.0078)	-0.012^{*} (0.0075)

Note: *p<.05

Looking at table (7) we can say that the world still converges at a rate around 2%. But now the regional convergence story is quite different. Especially if we look into "World without Africa" - the convergence rate is even less than 1%. The convergence in the world in this period is driven by African countries.

4 Institutional and Infrastructural Change

We argue that the recent trend of convergence and growth in the African countries is due to core institutional and infrastructural development. We first look into the variables that were significant in the post-2000 estimation of convergence rate for Africa. We note here that the pre-2000 estimation only had the intercept and average growth of neighboring countries as significant.

From table (8) we can say that variabels like "female school years", "mortality rate" and "government expenditure share" provide some evidence towards the importance of "from-within" development.

(Next we provide some more analysis to show instittional and infrastructural development.)

Table 8: Double Selection Significant Estimates, Africa - Post-2000

Variable	Lower CI	Upper CI
Intercept	0.135132	0.361932
Log Per Capita GDP	-0.060947	-0.027612
Female Secondary School Years	0.021367	0.065483
Mortality Rate	-0.001141	-0.000187
French Language Speaking	-0.057556	-0.011187
British Colony	-0.057361	-0.005741
Exchange Rate	0.000004	0.000027
Government Expenditure Share	0.017651	0.298206

5 Conclusion

In this paper LASSO selection and double selection estimation methodologies are used to investigate the convergence properties of African counrties. We start with structural break tests on the growth and convergence rate, and find that there is a structural break in African growth and convergence rate. We divide the regimes as pre-2000 and post-2000 and estimate the conditional convergence rate. We find that Africa (even as a 'club' by itself) was not converging pre-2000 whereas post-2000 it is converging and the rate is 4%. We used LASSO to select from numerous variables. We also used GBM and Random Forest to compare the selections.

We attribute the post-2000 convergence to more institutional and infrastructural development within Africa. Before the late 1990's it was mainly foreign aid which was leading to some growth in the 'non-resourceful' African countries, but in recent years there has been development 'from-within' the countries.

As next steps it would be perhaps meaningful to disaggregate the African 'convergence-club' to smaller meaningful groups and see if the story still holds for each of these groups.

6 Appendix

6.1 Data description

6.2 Correlation among variables

Table 9: Glimpse of correlation structure in the data

	luF	lpF	lsF	yrschF	yrschpriF	yrschsecF	luM	lpM	lsM	yrschM	yrschpriM	yrschsecM
luF	-	-0.48	-0.77	**-0.91**	**-0.92**	-0.73	**0.95**	-0.18	-0.65	-0.84	-0.84	-0.64
lpF	-0.48	-	-0.15	0.11	0.35	-0.19	-0.47	**0.89**	-0.22	0.04	0.31	-0.24
lsF	-0.77	-0.15	-	**0.89**	0.76	**0.9**	-0.72	-0.42	**0.94**	0.87	0.71	0.85
yrschF	**-0.91**	0.11	**0.89**	-	**0.93**	**0.89**	-0.86	-0.18	0.79	**0.96**	0.88	0.81
yrschpriF	**-0.92**	0.35	0.76	**0.93**	-	0.67	-0.87	0.07	0.66	0.88	**0.96**	0.58
yrschsecF	-0.73	-0.19	**0.9**	**0.89**	0.67	-	-0.68	-0.44	0.83	0.88	0.6	**0.96**
luM	**0.95**	-0.47	-0.72	-0.86	-0.87	-0.68	-	-0.24	-0.66	-0.85	-0.87	-0.64
lpM	-0.18	**0.89**	-0.42	-0.18	0.07	-0.44	-0.24	-	-0.51	-0.24	0.07	-0.52
lsM	-0.65	-0.22	0.94	0.79	0.66	0.83	-0.66	-0.51	-	0.84	0.66	0.86
yrschM	-0.84	0.04	0.87	**0.96**	0.88	0.88	-0.85	-0.24	0.84	-	**0.89**	0.87
yrschpriM	-0.84	0.31	0.71	0.88	0.96	0.6	-0.87	0.07	0.66	**0.89**	-	0.56
yrschsecM	-0.64	-0.24	0.85	0.81	0.58	**0.96**	-0.64	-0.52	0.86	0.87	0.56	-

References

Barro, R. J. and X. Sala-i-Martin (1992): "Convergence," *Journal of Political Economy,* 100(2), 223-251.

Belloni A, V. Chernozhukov, and C. Hansen (2011b): "Inference on Treatment Effects After Selection Amongst High-Dimensional Controls With an Application to Abortion on Crime," available at *arXiv: 1201.0224.* [2369,2371,2374,2382,2383]

Leeb, H. and B. M. Pötscher (2008): "Can One Estimate The Unconditional Distribution Of Post-Model-Selection Estimators?," *Econometric Theory*, 24(02), 338-376.

Mankiw, N. G., D. Romer, and David N. Weil. (1992): "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107(2), 407-437.

Anselin L (1988) *Spatial Econometrics: Methods and Models*, 1988a

Armstrong H (1995a) *An appraisal of the evidence from cross-sectional analysis of the regional growth process within the European Union*, in Convergence and divergence among European Regions, ed. H. Armstrong and R. Vickerman. London: Pion.

Armstrong H (1995b) *Convergence among the regions of the European Union*, in Regional Science74: 143-52

Attfield CLF (2003) *Structural Breaks and Convergence in Output Growth in the EU*, Bristol Economics Discussion Papers 03/544, Department of Economics, University of Bristol, UK.

Barro, R. J. (1991) *Economic Growth in a Cross Section of Countries*, the Quarterly Journal of Economics (1991) 106 (2): 407-443.

Barro, R. J. (2012) *Convergence and Modernization Revisited*, National Bureau of Economic Research, Working Paper Series 18295, August 2012

Barro, R. J. and Sala-i-Martin, X. (1991) *Convergence across States and Regions*, Brooking Papers on Economic Activity, 1: 107-182.

Barro, R. J. and Sala-i-Martin, X. (1992) *Convergence*, Journal of Political Economy 100 (2), 223-251. (a)

Baumol, W. J. (1986) *Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show*, 1986: American Economic Review, 76 (5), 1072-1085.

Belloni A, V Chernozhukov, C. Hansen (2011a): *Estimation and Inference Methods for High-Dimensional Sparse Econometric Models*, in Advances in Economics and Econometrics, 10th World Congress of Econometric Society. Cambridge: Cambridge University Press. [2371,2374,2378,2382]

Belloni A, V Chernozhukov, C. Hansen (2011b): *Inference on Treatment Effects After Selection Amongst High- Dimensional Controls With an Application to Abortion on Crime*, available at arXiv: 1201.0224. [2369,2371,2374,2382,2383]

Belloni A, V Chernozhukov, C. Hansen (2012): *Sparse Models and Methods for Optimal Instruments with an Application to Eminent Domian*, Econometrica, Vol. 80, No. 6 (November, 2012), 2369â€“2429

Bräuninger M, Niebuhr A (2005). *Agglomeration Spatial Interaction and Convergence in the EU*, HWWA Discussion Papers 322, Hamburg Institute of International Economics (HWWA).

Paul Collier (2008) *The Bottom Billion: Why the Poorest Countries are Failing and What Can Be Done About It*, 978-0-19-537463-6, Paperback, 02 October 2008

Durlauf Steven N, Johnson Paul A (1995) *Multiple Regimes and Cross-Country Growth Behaviour*, Journal of Applied Econometrics, John Wiley & Sons, Ltd., vol. 10(4), pages 365-84, Oct.-Dec..

Elhorst JP (2003) *Specification and estimation of spatial panel data models*, International Regional Science Review 26(3):244-268

Jushan B (2010) *Common breaks in means and variances for panel data*, Journal of Econometrics, Volume 157, Issue 1, July 2010, Pages 78â€“92

Perron, P. and Qu, Z. (2006) *Estimating restricted structural change models*, in Journal of Econometrics, Volume 134, Issue 2, October 2006, Pages 373â€“399

Tibshirani, R. (1996) *Regression Shrinkage and Selection via the Lasso*, Journal of the Royal Statistical Society, Ser. B, 58, 267â€“288. [2370,2372,2380,2381]