THE IMPACT OF ETHNIC DIVERSITY ON ECONOMIC GROWTH:
THE CASE OF AFRICAN COUNTRIES

Adora Ofodile¹
Duke University
Durham, NC
April, 2001

¹ Adora C. Ofodile, Trinity College, Class of 2001, graduated in May, 2001 with a BS in Economics, a minor in Political Science, and a Markets And Management Certificate. She will be moving to New York to be an analyst at Prudential. She plans to attend business school in the near future.
Acknowledgements

I dedicate this paper to my father for the inspiration to research this topic after our conversation and I thank God for His grace that got me through it. I also thank Vladimir Pantyushin, my advisor, who was always encouraging and willing to help. Others that gave me feedback on this topic include Joseph Grieco, Michelle Connolly, and Cathleen McHugh.
Preface

A bi-annual survey\(^2\) done by the International Monetary Fund(IMF) in May 2000 revealed the following:

\[\text{Figure 4.1. Advanced and Developing Economies: Per Capita Income}^1\]

\(^1\text{Thousands of U.S. dollars at 1996 prices.}\)

East Asia, which includes China and has the largest population, and the newly industrialized economies are the only country groups that are rapidly converging with the industrial countries. The thickness of the bars reflects the population in each region.

\[\text{Average growth, 1970–98 (percent)}\]

\[\text{1970} \quad \text{1985} \quad \text{1998} \quad \text{Average growth, 1970–98 (percent)}\]

- Africa
- East Asia
- South Asia
- Middle East and Europe
- Western Hemisphere
- Newly industrialized economies
- Industrial countries
- Memorandum
- Heavily indebted poor countries

Africa as shown has had a pretty steady growth in population but with decreasing growth of per capita income. The East Asia region, which includes China, has the largest population but has increased per capita income over the past 30 years at a rate of 5.6%. This increase in per capita income over the past 30 years is reciprocated by the other regions of the World, the only divergence being the African region at a –0.2% growth rate.

What is the explanation for this pattern of growth? I present a hypothesis based largely on the work done by Robert Barro on the determinants of economic growth. Barro’s regressions look at about 80 countries and suggest that the explanations for growth are found in the initial levels of physical and human capital, as well as some other variables he calls control and environmental variables. These explanatory variables, however, do not seem to explain the African puzzle. I propose that an ethnic diversity variable should be added to the list of determinants of economic growth. Ethnic diversity refers to the different groups within a country. The differences can be found in religion, language, culture, etc. I run an estimate for some 56 countries and find that ethnic diversity has a negative effect on growth. In addition I look at the African region as a whole and then at Nigeria, a specific case of interest, in search for an explanation of the negative tendency of economic growth in Africa. The basic idea is that the greater the ethnic diversity, the more difficult it is for a country to achieve levels of economic growth. It is true that underdeveloped institutions and environments non-conducive to growth as well as cultural factors are some of the reasons for the backwardness in African economies. However, I am suggesting that ethnic diversity is one of the main underlying initial factors. I expect to find a negative relationship between ethnic diversity and growth rate. In my study, the ethnic diversity variable will be an ethnic fractionalization score assigned to each country based on the ratio of ethnic groups within a country to the entire population. My data supports this hypothesis with a negative correlation (-0.3) with growth.

Introduction

This paper has five sections. In Section One I discuss thoughts/works of other economists relevant to the subject. Some work has been done in the area of explaining growth in the African region. In this section I look at those of interest. I begin Section
Two by discussing some models on growth including the neoclassical growth model and then Barro’s model, which is an extension of the neoclassical model. I construct a model based largely on the work done by Barro on the determinants of economic growth, but I seek to supplement his analysis by adding an ethnic diversity variable to the list of explanatory variables for economic growth. In Section Three I provide a description of the data used to test my model. I compile data over a 15-year period, on a sample of some 56 countries randomly selected. In Section Four I perform an econometric analysis in order to determine the relationship between ethnic diversity and growth. Specifically I look at the changes in the GDP per capita growth rate with respect to ethnic diversity. I expect the lower the level of ethnic diversity the less conflicts and instances of corruption, and therefore higher levels of economic growth. I first look at the data and establish a model using the method of least squares to obtain the best linear unbiased estimate. Then I analyze the model based on some of the OLS assumptions to make an empirical conclusion about economic growth of countries and ethnic diversity. Another part of this section looks at the African region specifically. Africa provides a good opportunity to test the validity of my hypothesis and establish whether ethnic diversity serves as a determinant of economic growth. I apply the same regression to some 30 African countries establishing a regional relationship between growth and diversity. In this section, I present my personal account supported by literature to give a rough background of Nigeria. It gives the reader an idea of where my hypothesis stemmed from. Section Five is a conclusion of my findings. I find ethnic diversity to be an underlying determinant of growth in the explanation of Africa.

Section 1: Relevant Literature

Many economists have performed and analyzed models each looking for an intricate range of variables that explain or influence economic growth in Africa. Barro’s model (Barro 1995, 1997) identifies two types of variables that are determinants of economic growth, state variables, and control and environmental variables. His model is discussed in detail in the next section. Sachs and Warner (1997) perform a cross-sectional analysis over a 25-year period concluding that significant variables include international openness, and other trade-related variables such as landlockedness, which
decreases the openness of countries in terms of trade, therefore reducing the diffusion of
ideas and technology that help growth. Easterly and Levine (1997) study an ethno-
linguistic variable. Their other explanatory variables include measures of financial depth,
the black market premium, infrastructure, and other standard variables. Easterly (2001)
goes further in a World Bank paper to argue that the institutions are important
distinguishing the effects of high-quality versus low-quality institutions. Ethnic diversity
is a key explanatory variable for growth but the quality of institutions mitigate or
exacerbate the adverse effects of ethnic diversity on economic growth. Sala-I-
Martin (1997) performed over two million regressions using some 60 variables based on
significance and found that only 22 of his variables were significant, using the “extreme
bounds test” applied by Levin and Renelt (1992) to isolate the robust significance of
variables based on different specifications of the variables. Sala-I-Martin specifically
takes the robustly significant variables and analyzes how the betas varied. The 22
variables encompassed much of the other variables used by his colleagues. So which
economist is correct or closest to the explanation for economic growth in Africa?

A CREDIT research paper done by Bleaney and Nishiyama (2000) tested Barro’s
model against each other using the latter’s data set over the same time period to perform
non-nested tests. Their conclusion was that each model had something to offer to the big
picture. However, the Sachs and Warner model includes the trade-like variables, and
thus performed better than the other two models. Perhaps their out-performance was due
to the biased data set, Sachs and Warners data set was used to test the three models. The
idea though, is that all the models have a partial contribution to the question of economic
growth in Africa.

What is important here is that the growth process depends on a range of variables
that are considered depending on the researcher’s objectives. Barro’s model takes the
standard variables of the traditional growth into account. It includes capital and allows
for other endogenous variables like the trade measures suggested by Sachs and Warner
(1997). However, the model fails to encompass the one thing that was covered by
Easterly and Levine (1997): ethnic diversity. I intend to fortify Barro’s regressions by
including an ethnic diversity variable. This variable is a bit different from the ethnic
diversity variable used by Easterly and Levine (1997) and will be discussed in the next few sections.

Section 2: The Hypothesis

Economists have done a lot of work on explaining the determinants of long-term growth. Robert Solow (1956) is the father of the neoclassical growth model, which explains traditional growth based on a few assumptions. The Solow model first assumes the Cobb-Douglas production function. Total output is a function of physical capital, labor and technology, which is exogenous and therefore is constant. The second assumption is that population growth is constant and equal to labor growth. The other assumptions include a constant savings rate, a closed economy, and that investment is a function of how much the country wants, how much the country has, and depreciation. There are implications following the neoclassical growth model. There is conditional convergence, no long-run growth, a need for continuous exogenous increases in technology to get long-run growth, and empirically, we observe very slow convergence.

Paul Romer (1983) realized that the reason for no long run growth in the neoclassical model is the inability to explain technological improvements while avoiding the diminishing returns to capital. His endogenous growth model has only inputs, capital and technology. The implications of the model include no conditional convergence, the rich countries always remain rich therefore history matters, a county’s initial stock of capital is important. Long run growth is achieved due to constant returns to capital. However, this model does not explain the empirically observed conditional convergence. Romer’s learning-by-doing model rids of diminishing returns to capital thru the externality of investment. Knowledge creation is an unintended by-product of investment in capital and is a non-rival private good. The positive spillover from investment allows for the creation of technology without private incentives. The positive spillover, however, is not internalized; therefore the market equilibrium is not socially optimal. There are other models including Lucas’ two-sector human capital model and R&D models of endogenous growth that provide suggestions for achieving long-run growth. Empirically, conditional convergence is observed and the neoclassical model implies this characteristic of growth.
Robert Barro by extending the neoclassical model presents a slightly different analysis in a lecture he gave at the London School of economics. Accordingly, the initial levels of output help explain the differences in growth between countries. Given the graph below,

![Graph](image)

at A, the growth rate is zero. If a country’s level of output is below its steady-state level of output $K^*$, it needs endogenous increases in technology to achieve long run growth and thereby catch up. According to Barro, two and half percent of the distance between the current level and the steady state level of output is eliminated due to each year’s growth. Keep in mind however that growth is attributed to other factors therefore the convergence is conditional and does not necessarily imply that rich countries will grow at a slower pace that the poor countries. One can look at the case of Africa to understand this point. There are explanatory variables for growth patterns that affect the steady state output level. Barro’s econometric analysis identifies these explanatory variables. These determinants of growth include a high the number of males attaining the secondary level of education, high life expectancy, a low fertility rate, low government consumption, low level of political instability, good terms of trade, etc.

According to Barro:

$$D_y = f(y,y^*)$$. $D_y$ is the growth rate of per capita output, $y$ is the current level of per capita output, and $y^*$ is the long-run or steady state level of per capita output. The growth rate, $D_y$, is diminishing in $y$ for given $y^*$ and rising in $y^*$ for given $y$. 

The target value \(y^*\) depends on an array of choice and environmental variables. The private sector’s choices include savings rates, labor supply, and fertility rates, each of which depends on preferences and costs. The government’s choices involve spending in various categories, tax rates, the extent of distortions of markets and business decisions, maintenance of the rule of law and property rights, and the degree of political freedom. Also relevant for an open economy is the terms of trade, typically given to a small country by external conditions.

Using the equation above, Barro conducted the following econometric analysis:

<table>
<thead>
<tr>
<th>Explanatory Variables</th>
<th>(1)</th>
<th>(2)</th>
<th>Explanatory Variables</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log(GDP)</td>
<td>-.0254</td>
<td>-.0225</td>
<td>Terms of trade change</td>
<td>.137</td>
<td>.127</td>
</tr>
<tr>
<td></td>
<td>(.0031)</td>
<td>(.0032)</td>
<td></td>
<td>(.030)</td>
<td>(.030)</td>
</tr>
<tr>
<td>Male secondary and</td>
<td>.0118</td>
<td>.0098</td>
<td>Democracy index</td>
<td>.090</td>
<td>.094</td>
</tr>
<tr>
<td>higher schooling</td>
<td>(.0025)</td>
<td>(.0025)</td>
<td></td>
<td>(.027)</td>
<td>(.027)</td>
</tr>
<tr>
<td>Log(life expectancy)</td>
<td>.0423</td>
<td>.0418</td>
<td>Democracy index</td>
<td>-.088</td>
<td>-.091</td>
</tr>
<tr>
<td></td>
<td>(.0137)</td>
<td>(.0139)</td>
<td>squared</td>
<td>(.024)</td>
<td>(.024)</td>
</tr>
<tr>
<td>Log(GDP)*male</td>
<td>-.0062</td>
<td>-.0052</td>
<td>Inflation rate</td>
<td>-.043</td>
<td>-.039</td>
</tr>
<tr>
<td>schooling</td>
<td>(.0017)</td>
<td>(.0017)</td>
<td></td>
<td>(.008)</td>
<td>(.008)</td>
</tr>
<tr>
<td>Log (fertility rate)</td>
<td>-.0061</td>
<td>-.0135</td>
<td>Sub-Saharan African dummy</td>
<td>-.0042</td>
<td>-.0043</td>
</tr>
<tr>
<td></td>
<td>(.0053)</td>
<td>(.0053)</td>
<td></td>
<td>(.0043)</td>
<td>(.0043)</td>
</tr>
<tr>
<td>Government</td>
<td>-.136</td>
<td>-.115</td>
<td>Latin America dummy</td>
<td>-.0054</td>
<td>-.0032</td>
</tr>
<tr>
<td>consumption ratio</td>
<td>(.026)</td>
<td>(.027)</td>
<td></td>
<td>(.0032)</td>
<td></td>
</tr>
<tr>
<td>Rule of law index</td>
<td>.0293</td>
<td>.0262</td>
<td>East Asia dummy</td>
<td>.0050</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.0054)</td>
<td>(.0055)</td>
<td></td>
<td>(.0041)</td>
<td></td>
</tr>
</tbody>
</table>

R\(^2\): .58, .52, .42
No. of observations: 80, 87, 84

The growth rates of real per capita GDP spanning 1965-1975, 75-85, and 85-90 are the dependent variable. Barro has two sets of independent variables, one that he calls state variables and the other, control and environmental variables. The state variables are all at initial levels representing physical and human capital. The idea of how these variables affect growth rates goes back to the observation of conditional convergence implied by the neoclassical model. Given a diminishing returns to capital characteristic, a country grows faster the further away it is from its own steady state \(k^*\). Therefore the country’s initial level, \(k\) is important in determining growth rate. Countries have higher initial levels of physical capital or human capital will then grow at a slower rate. Keep in mind however that the effects on growth by these initial levels are not straight forward. For example, for a constant level of physical capital, a higher level of human capital
implies a higher growth rate. All other variables are called control and environmental variables. These variables affect the growth rate indirectly by affecting the steady state level of output per worker. The idea goes back to one of the implications of the neoclassical growth model, the continuous exogenous increases in technology necessary to get long run growth.

Using Barro’s regression, I introduce an ethnic diversity variable to explain the slow economic growth of the African region. I am hypothesizing that in searching for an explanation to the economic growth of countries in Africa the level of ethnic diversity is important. In Nigeria ethnic balancing affected the economy. This inherently led to the civil war, corruption, and to the stagnant growth of the economy. The search for equality would not be of concern in a country more homogenous in its population. Africa presents the perfect region to test my hypothesis because it is the most diverse continent with a huge amount of natural resources yet it has the slowest rate of growth. An ethnic diversity variable will therefore be added to the above list of the determinants of growth.

Section 3: The Data

I plan to specifically look at variables within the following categories:

- Initial economic position (gross domestic product – GDP per capita)
- Initial human capital (life expectancy and school enrollment)
- Economic indicator (savings rate)
- Political stability (Decolonization)
- Government’s presence in the economy (government spending)
- International openness (trade)
- Ethnic diversity variable

All data come from World Development Indicators (WDI) and are defined accordingly.

1) GDP per capita (WDI)

Initial GDP per capita, used by Barro, has a negative effect on growth. It enters the regression in logarithmic form to capture a rate of growth. This rate turns out to be the conditional convergence rate. The negative effect on growth rate is in accordance with the idea of conditional convergence. The speed at which a country is growing
depends on its initial starting point. A country grows faster the further away it is from its steady state.

2) **Life expectancy** (WDI)

   Barro measures life expectancy by taking the average over the four years prior to the beginning of the period being observed. I use 1970 statistics due to the limitations of available data. It enters the regression in logarithmic form, inherently representing health in the economy. Life expectancy has a positive effect on growth, the higher the life expectancy, the faster the rate of growth.

3) **School enrollment** (WDI)

   The level of male enrollment at the secondary level, measured in 1975, is another representation of the initial level of human capital in the society. This variable in Barro’s regression has a positive effect on the growth rate. The interaction between school enrollment and the initial level of GDP per capita is also considered.

4) **Savings rate** (WDI)

   Gross domestic savings (GDS) as a percentage of GDP is not used by Barro but is an economic variable that inherently measures how much is being put away for investment. In this case a higher savings rate raises the steady state level therefore implying a higher growth rate. My data does show a positive correlation.

5) **Political Instability**

   I tried to be careful in choosing a measure of political stability, making sure that my variable would not be highly correlated with any other independent variable. Normally, when independent variables $X_a$ and $X_b$ are highly correlated, the method of least squares suggests that one of the variables be omitted from the regression. In this region of the world, political stability is and continues to be a huge determinate of the amount of investment in the economy and is therefore an important variable to include.

   I base this variable on Matthew Krain’s hypothesis that “If a state has a recent history of decolonization, the probability of the onset of genocide or politicide and the
degree of its severity should be greater than for a state without a recent history of
decolonization” (Krain 1997). He presents the presence or absence of decolonization as a
dichotomous variable, with countries that have fewer years of independence having the
value 1 and the other countries 0. The data on decolonization comes from the work done

Looking at Strang’s data most African countries gained their independence
between the 60’s and mid 70’s, with exceptions like Djibouti being relieved from the
French colonial rule in 1977.

This variable is presented in Barro’s regression as the Rule of Law index and the
Democracy Index, which both had a positive effect on the growth rate. The primary
reason I am using this variable is to avoid a higher correlation between my independent
variables. I structure de-colonization as a dummy variable. Each country will be
assigned a zero or a one based on how recently the country gained its independence
(Krain 1997). I expect de-colonization to act like the other control and environmental
variables, (i.e., if a country has been independent for a short time, it should have a lower
steady state level of output per worker, therefore decreasing the growth rate, and vice
versa).

6) Government consumption (WDI)

General government consumption as a percentage of GDP is included in Barro’s
model. This variable will be the average over the 15-year period. It has a negative effect
on growth. The more the government consumes, the lower the level of investment in
capital, and therefore the tendency to a lower rate of growth (“crowding out” effect).

7) Trade (WDI)

Trade as a percentage of GDP enters the regression as an average over the time
period. This is a variation of another variable used by Barro, who found it to have a
positive effect on growth, the greater the openness, the greater the gains of the economy.
8) **Ethnic fractionalization score**

There are basically different ways one could go about collecting an ethnic fractionalization score. One is based on linguistic cleavages as did Easterly and Levine, another on religion, and yet another based on the different groups. I obtained my data from Matthew Krain’s work on the subject (Krain 1997). His work is an extension of the ethnic fractionalization index created by Taylor and Hudson (1972). Their work on ethnicity based on linguistic cleavages “misrepresent the true nature of ethnic cleavages in many societies,” and according to Krain, language differences do not always work with ethnic ones.

Krain’s measurement of ethnic fractionalization included “…first calculating the proportion of the population of each ethnic group to the total population of the country and then squaring it. Next, [he] sums the squared proportions for all groups and subtracts that number from 1 to come up with the fractionalization measure”. Countries with low scores, like Japan, which has the minimum score of .01 are interpreted as having “a huge majority (99%) of one group and a very small minority (1%) of another”. In Nigeria the relatively high score of .87, “indicates many groups with small or relatively equal percentages of the population”. Mentioned earlier were the Igbo, Hausa, Yoruba and Fulani, all make up significant proportions of the population.

The data used by Krain to calculate the ethnic fractionalization scores were collected from Dan Golenpaul Associates (1959); Kurtz(1970); Bacheller(1979); Connor(1979); Bauazizi and Weiner(1986); Morrison, Mitchell, and Paden(1989); and Phandis(1990). There were other sources used to control for source bias or error. [His] index has the same idea of the Taylor and Hudson. Taylor and Hudson’s measure in the case of Africa, Latin America and other countries differs significantly. However, Krain, and Taylor and Hudson’s measures are highly correlated (.84) for the time period that they have in common.

This variable could fall under either of the two categories mentioned above, the initial variables or the control and environmental variables. Under the initial capital category, ethnic diversity makes sense because ethnicity is not a variable that changes significantly over time, in the case of the African region. In the other category, a lower
ethnic fractionalization score would result in a higher steady state level of output per worker. Why?

A slew of economists and political scientists have sought to understand the mechanisms behind the interaction between ethnic groups and the effects of interaction. Less conflict between groups and less difference translate into less diversion of time to corruption and more devotion to economic growth for the country as a whole. This higher steady state level translates into a higher growth rate in accordance with conditional convergence.

Krain had a score for a substantial amount of countries. My data set began as some 190 countries was first randomly narrowed to 75, some of which were removed due to limited data and is now further limited to some 56 countries. Later, I concentrate specifically on the African region, which is limited to some thirty countries, due to the absence of data on the African countries during certain periods. As mentioned earlier, my data supports the expected negative correlation between the growth rate and the ethnic fractionalization scores of the countries. Below are two scatter plots of growth and ethnic diversity. One is for the World and the other for a collection of African countries. Overall, there is a negative relationship between the ethnic fractionalization score and growth. There is a greater tendency toward the negative slope among the African countries.
The goal of this project is to explore the relationship between various economic growth variables, especially the ethnic diversity variable and the growth rate of an African country. Through careful econometric examination of the eight independent variables, I expect to find that those variables that have a positive effect the steady state will have a positive effect on the growth rate and vice versa. The degree to which each variable affects the growth rate will be determined through regression analysis.

Section Four: Analysis

Part I: Justification for using OLS

Barro used a two-stage least squares method. I will perform a basic OLS regression on our data in order to obtain the best linear unbiased estimation based on the following assumptions.

1. Linear in parameters

The stochastic process \{ (x_{t1}, x_{t2}, \ldots, x_{tk}, y_t): t = 1,2,\ldots,n \} follows the linear model

\[ y_t = \beta_0 + \beta_1 x_{t1} + \ldots + \beta_k x_{tk} + u_t, \]

where \{ u_t: t = 1,2,\ldots,n \} is the sequence of errors or disturbances. Here, n is the number of observations (time periods).

I assume that our dependent variable is linearly related to the parameters. The growth rate of GDP is linearly related to the listed independent variables. I believe that this is the correct model and thus we do not violate Assumption #1.

2. Zero conditional mean

For each t, the expected value of the error \( u_t \), given the explanatory variables for all time periods, is zero. Mathematically,

\[ E(u_t | X) = 0, t = 1,2,\ldots,n. \]

If this equation is true, then the explanatory variables are strictly exogenous and OLS is unbiased. This assumption says that \( u_t \) must be uncorrelated with \( x_{sj} \) even when \( s \neq t \).
A related equation:

\[ E(u_t \mid x_{t1}, \ldots, x_{tk}) = E(u_t \mid x_t) = 0 \]

says that the \( x_{ij} \) are contemporaneously exogenous and implies that \( u_t \) and the
explanatory variables are contemporaneously uncorrelated: \( \text{Corr}(x_{ij}, u_t) = 0 \) for all \( j \).

Two main reasons for a possible failure of this assumption are omitted variables
and measurement error. When forming my model, I selected economic indicators that
I thought would have a significant effect on the growth rate of an African country. I
used the factors tested by Barro and just added a variable. If I have left out
explanatory variables that are also correlated with \( y \), then the OLS estimators will be
biased and Assumption #2 will be violated. Another way that Assumption #2 can be
violated is if I take the wrong transformations of our \( x \)'s and \( y \) variables. More
specifically, if I choose to use \( \text{level} \ (x_1) \) when the population model really has \( \log \ (x_1) \)
then Assumption #2 will be violated. However the model I select is that which I
believe to be correct through logical economic reasoning.

Assumption #2 assumes that the error, \( u_t \), is uncorrelated with past and future
values of \( x \). Thus, if the error in time \( t \) affects the value of \( x \) in time \( t+1, t+2, \) etc.
(some future time period), Assumption #2 will be violated. One variable that I would
include but cannot at this time due to the absence of comprehensive data that
encompasses the time period in question is corruption. Corruption can affect the
growth rate in two ways. It could be considered at initial levels where it would be
presented as a lag variable. This suggests that corruption in period one has a negative
effect on growth in period two. Another way corruption plays a role is on the other
end of the spectrum. Perhaps, slow growth in period one stimulates corruption in
period two. Either way this variable would be controlled for.

A related assumption is the asymptotic assumption of zero conditional mean
(Wooldridge p.352). This assumption states that for each \( t \), \( E(u_t \mid x_t) = 0 \). It is much
weaker than the aforementioned assumption because it puts no restrictions on how \( u_t \)
is related to the explanatory variables in other time periods.

3. **No perfect collinearity**

   *In the sample no independent variable is constant or a perfect linear combination
   of the others.*
Because my explanatory variables are all economic indicators of some sort, I predict that some of them, in the true population, may in fact be highly correlated with others. For example, political stability in theory is likely a function of ethnic diversity. Although it is highly likely that certain independent variables are highly correlated with others, it is most likely not the case that one is a perfect linear combination of another. If it is the case that one independent variable is a perfect linear combination of another, I will have to correct for the violation by omitting one of the variables. However to alleviate the correlation problem of political stability, I have chosen to use a measure for political stability according to the work done by David Strang (1991) that is not a combination of another independent variable. In addition, interaction terms tend to be correlated.

Looking at the data on randomly selected countries, the correlation of the variables is as follows:
Looking at the data on some 30 African countries, the correlation is as follows:

\[
\begin{array}{ccccccccc}
\text{LogGDP} & \text{LogLifeExp} & \text{School Enroll} & \text{LogGDP} & \text{Govt. Cons.} & \text{Trade} & \text{Ethnic Score} \\
\text{LogGDP} & 1 & & & & & \\
\text{LogLifeExp} & 0.74 & 1 & & & & \\
\text{School Enroll} & 0.81 & 0.83 & 1 & & & \\
\text{LogGDP} \ast \text{School Enroll} & 0.86 & 0.82 & 0.99 & 1 & & \\
\text{Savings} & 0.48 & 0.42 & 0.37 & 0.36 & 1 & \\
\text{Decol. Dummy} & -0.5 & -0.5 & -0.7 & -0.7 & -0.2 & 1 & \\
\text{Govt. Cons.} & 0.3 & 0.06 & 0.14 & 0.18 & 0.06 & -0 & 1 & \\
\text{Trade} & 0.14 & 0.04 & 0.05 & 0.05 & 0.27 & 0.13 & 0.2 & 1 & \\
\text{Ethnic Score} & -0.5 & -0.5 & -0.5 & -0.5 & -0.2 & 0.54 & -0.2 & -0 & 1 & \\
\end{array}
\]

4. Homoskedasticity

Conditional on \( X \), the variance of \( u_t \) is the same for all \( t \):

\[
\text{Var}(u_t \mid X) = \text{Var}(u_t) = \sigma^2, \quad t = 1, 2, \ldots, n.
\]

This assumption states that \( \text{Var}(u_t \mid X) \) cannot depend on \( X \) – it is sufficient that \( u_t \) and \( X \) are independent – and that \( \text{Var}(u_t) \) must be constant over time. Because my data is of the cross-sectional nature, it is highly unsusceptible to sudden shocks that could be temporary or permanent. Thus, it should not be difficult to assume that the variance over a certain period will be constant.
5. No serial correlation

| Conditional on $X$, the errors in two different time periods are uncorrelated: |
| $\text{Corr}(u_t, u_s | X) = 0$, for all $t \neq s$. |

If $X$ is nonrandom, then the equation is simplified to:

| $\text{Corr}(u_t, u_s) = 0$, for all $t \neq s$. |

I expect certain economic shocks to have a simultaneous effect on the dependent variable within and across time periods, therefore this assumption should not be a problem since I am looking at a cross section of data. If the error term in time $t$ is positive and I believe that the error term in time $t+1$ will also be positive, then this assumption would be violated. In my regression, if the growth rate is unusually high in one period, it will be averaged into the other time periods (with the given levels of GDP, life expectancy, savings rate, political stability, government spending, trade, and ethnic diversity). Assumption #5 would be violated if my data was of time-series nature.

6. Normality

| The errors $u_t$ are independent of $X$ and are independently and identically distributed as $\text{Normal}(0, \sigma^2)$. |

If my regression follows the violations of the previous assumptions, then this assumption will also be violated. If the variance of the error term is not constant over time periods and the error terms in adjacent periods are correlated, then the errors $u_t$ can be normal without the other assumptions, but will not be distributed $\text{normal}(0, \sigma^2)$ nor iid.

**Part II: The basic regression on growth of GDP per capita using the World sample**

Again, my intent is not to recreate Barro’s regression but merely to use his econometric model and to supplement it with the addition of an ethnic diversity variable. First I look at the regression on the 59 countries chosen to participate. Then I specifically look at the African region to further discuss my findings.
The 59 countries used in this regression were randomly selected. I started with some 190 countries, selected 75 by a random generator process, and the sample continued to shrink due to the unavailability of data.

The graph above of the ethnic fractionalization score represents the 59 countries in the regression. It shows quite a variation within the sample.

Below are the statistics from the regression. The independent variable estimates are all in agreement with Barro’s regression. I will not go into further detail with respect to these independent variables. What I am interested in is the estimate on the ethnic score variable. The variable for ethnic diversity has a negative effect on growth rate. The coefficient is \(-0.03\). The result suggests that a one-unit increase in the ethnic fractionalization score, has a negative influence on the rate of growth.

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Standard Error</th>
<th>t Stat</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.09024</td>
<td>0.085855</td>
<td>-1.05107</td>
</tr>
<tr>
<td>Log(GDP_)</td>
<td>-0.03491</td>
<td>0.007059</td>
<td>-4.94497</td>
</tr>
<tr>
<td>Log(Life Exp)</td>
<td>0.100857</td>
<td>0.022732</td>
<td>4.436699</td>
</tr>
<tr>
<td>School Enrollment1975</td>
<td>-0.0021</td>
<td>0.000911</td>
<td>-2.30022</td>
</tr>
<tr>
<td>Lod(GDP)*Sch. Enroll</td>
<td>0.000257</td>
<td>0.000114</td>
<td>2.254862</td>
</tr>
<tr>
<td>Savings</td>
<td>0.001231</td>
<td>0.000302</td>
<td>4.076978</td>
</tr>
<tr>
<td>Decol. Dummy</td>
<td>-0.00274</td>
<td>0.008254</td>
<td>-0.33209</td>
</tr>
<tr>
<td>Govt. Cons.</td>
<td>-0.00029</td>
<td>0.000494</td>
<td>-0.58184</td>
</tr>
<tr>
<td>Trade</td>
<td>8.74E-05</td>
<td>5.02E-05</td>
<td>1.741069</td>
</tr>
<tr>
<td>Ethnic Score</td>
<td>-0.03263</td>
<td>0.010501</td>
<td>-3.10747</td>
</tr>
</tbody>
</table>
**Regression Statistics**

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
<td>0.781929</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R Square</td>
<td>0.611412</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted R Square</td>
<td>0.535384</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.017409</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>56</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Part III: The basic regression on growth of GDP per capita using the Africa sample**

Looking specifically at 30 African countries to further illustrate my argument, there are higher scores of ethnic fractionalization and there is less variance within this isolated sample.

The average ethnic fractionalization score average is 0.5 for the African region.
Regression Statistics

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
<td>0.830265</td>
</tr>
<tr>
<td>R Square</td>
<td>0.689341</td>
</tr>
<tr>
<td>Adjusted R Square</td>
<td>0.549544</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.017511</td>
</tr>
<tr>
<td>Observations</td>
<td>30</td>
</tr>
</tbody>
</table>

For a cross section of African countries the ethnic diversity variable has a coefficient of -0.05. Intuitively, this means that a unit increase in the ethnic fractionalization score has a negative effect on the growth rate of GDP by that the above coefficient. A high ethnic score would lead to a big decrease in the steady state level output per worker, therefore reducing the growth rate of GDP dramatically.

Part IV: Nigeria, an illustration

In 1960, Nigeria gained its independence. Then there was a civil war from 1966 to 1970. The Igbo tribe inhabited the eastern region of the south. The level of education was high in this region and they were more advanced economically. At the time of the civil war in the late 60’s, they had more lawyers, doctors, and engineers. The Yoruba tribe occupied the western part of the south. Muslims (the Hausa), the Kunari, and the Fulani, who were less educated than the easterners, inhabited the northern area. “It was no accident that in Independence Year 1960, the North with over half of Nigeria’s 50-million population had 41 secondary schools against the South’s 842; that the North’s first university graduate qualified just nine years before independence” (Forsyth 1969). The northerners feared eastern/southern domination of the economy and did not favor integration. In an attempt to better their position, the northern Federal government responded with Northernization, an ethnic balancing. The unification attempt included changes such as giving a salary to any northerner going to school. The school entrance exam levels were raised for easterners. The policies were seen as an injustice. As a result more northerners came to hold high positions in the country. Due to the Northernization efforts, inter-group tensions escalated in the mid to late 60s. The Igbos who were displaced by the efforts were angry and in search of solutions. These conflicts led to war. The British would support the northerners because they were “obedient and
undemanding (Forsyth 1969) and thus allowing the British to extract resources. This helped their position in the civil war. The easterners were supported by Haiti and secretly by the French. The war ended with northerners winning the battle and maintaining a majority in the federal government. As a result of the war, the population of males in the east decreased. The loss of a significant portion of qualified labor in addition to other societal problems amalgamated to an increase in fraud. There was a lack of law and order.

During the creation of states, those that would be in authority were not hired according to qualifications but according to ethnic balancing. Money was being diverted to families or to Swiss banks. Money for repairs and most other public affairs were diverted to individuals. There were no good roads and there were no telephones, etc. Parents could not take care of their kids. It became the survival of the fittest, leading to killings and fear. Few Igbos penetrated the society. Their daughters started marrying the Hausa. Before the civil war, no Igbo woman would ever touch a Hausa man but it became a necessity. In 1979, the civilians took over from the military. From 1979 to 1983, the economy experienced a boom. There was more prospecting of crude oil, a greater awareness of resources. Many called Nigeria the giant of Africa. Four major hydroelectric dams were built, enough to even export some to the neighboring Niger Republic. Nigeria was exporting food, tin oil, coal, rubber, etc. Although money continued to be diverted, there was so much of it due to the boom. The economy was doing well.

The ruling politicians were then dethroned by a military that squandered all the savings. The officers would build lots of houses. Some would purchase a whole street. Many owned cars, although there were no good roads.

Today, the case is not so different. Money is too important and continues to be the motivation behind most services. The level of corruption is above and beyond that found in other economies. Why? Where did the chain reaction begin? In the case of Nigeria, the ethnic balancing led to civil war. The loss of physical and human capital, led to a change in the labor market and to the case of “the end justifies the means” and “survival of the fittest”, increased the corruption level and the amount of capital being diverted out of the country. A British newspaper once wrote that three Nigerians could
buy the whole country. This all implies a decrease in the steady state level of output per worker and therefore a lower growth rate. It is therefore important to include an ethnic diversity variable in the regression on the determinants of economic growth.

The diversity of groups within a country presents a scenario/case much like that seen in the global market. Countries participating in the global system pay a price for improved economic performance and improved standard of living. Most of the time the trade off involves political sovereignty. For this reason, some countries do not participate in integration. The same is the case for different groups within a country. Each group has a tendency toward self-satisfaction. This leads to a gradual reduction in the contribution they will make to the national economy. According to Borjas (1992, 1995), this tendency toward self-satisfaction stems from the potential positive spillover each group might gain from the capital within. This possible spillover is depicted in Nigeria. The easterners were more educated that the northerners so they had more “ethnic capital” (Borjas) within their cluster. There are certain trade offs involved with integration, the preceding being one of them, the positive spillover becomes a public good. Other trade-offs have to do with power and pride. In Nigeria, the northerners were less educated than the easterners who knew how to sustain the economy. However, the northerners were not interested in integration. They proceeded to occupy the top seats of the federal governments despite the lack of qualifications, not caring to learn to live with the differences in cross-country ethnic groups. The economic success of a country depends on the groups/clusters within the country. Each ethnic group is a key actor with the goal of sustained economic growth. The determinate of success is how these actors interact. In Nigeria, ethnic conflict mostly resulted from competition between ethnic groups that had sole interest in advancing their businesses.

Section Five: Conclusion

There are different growth models explaining the patterns of economic growth around the world. Africa, however, remains the puzzling case. Robert Barro’s model names a couple of determinants of growth including the initial levels of physical and human capital, and other control and environmental variables. I supplement Barro’s model by adding a variable that accounts for ethnic diversity. I first look at the effect of
ethnic diversity in the world, then I look specifically at the African region to draw an empirical conclusion from my data.

The hypothesis that an ethnic diversity variable should be added to the determinants of economic growth is supported by the regressions above. The negative tendency with respect to growth rate occurs for different reasons that have been discussed by economists and political scientists. Ethnic diversity creates social barriers that hinder the positive spillover encountered through knowledge creation as a result of investment. It makes the integration of individual actors more difficult, and therefore reduces the level of social capital that would otherwise be attained. In addition, there are issues of trust, and issues of exploiting the public good for selfish reasons, all of which hinder growth.

Nigeria is a country abundant in resources, a real case where social relations between groups had significant economic implications. Much literature has been written on the subject with the intention of exposing the structures and processes that have caused much of its deterioration. The lists blame inadequate leadership, colonialism, neo-colonialism, etc. as being responsible for its national crisis. Easterly (2001) suggests that institutions are also important in this case. Specifically, the quality of institutions is important. High quality institutions tend to reduce effects of ethnic diversity, but poor institutions tend to worsen the case.

After performing econometric analysis on some determinants of growth, my conclusion is that the problem and the explanation are rooted in ethnic diversity. One last grid that might be helpful is that of 21 high-income OECD countries. They have an average ethnic fractionalization score of 0.15. The difference in growth lies between the 0.15 average ethnic fractionalization score and the 0.5 average for the African region.
Bibliography


