Ethnic Diversity and Economic Growth

The application of ethnic diversity measures in economic empirical research.

Master Thesis: Erasmus University Rotterdam
Erasmus School of Economics
Department of Economics
January 2010

Supervisor: Dr. B. S. Y. Crutzen

Author: T. J. Bouwens

Exam Number: 314263

E-mail Address: Tom_Bouwens@hotmail.com

Abstract

Since Easterly and Levine published their paper: Africa's Growth Tragedy: "Policies and Ethnic Divisions"(1996), ethnic diversity has become the focus of many economic papers, discussing its negative effects on economic performance. However, ethnic diversity is a miscellaneous term and when measured by different individuals or institutions will yield different results. In part two of this thesis I find, conducting research on ethnic, linguistic and religious diversity data from several sources, that, even though institution might argue to measure the same, their outcomes are often very different. Resulting in correlations as low 0.64, which seems remarkably low considering they aim to measure the same phenomenon.

Conducting empirical research, on all 19 different diversity variables included in this thesis, shows that some ethnic fractionalization variables produce similar significant results in economic growth research. However, since the differences among the diversity variables are so large, results are always subject to doubt and are impossible to reliably quantify.

Keywords: Diversity (Ethnic, Religious, Linguistic), Economic Growth

Table of Contents

Introduction 3

1 Existing literature 4

2 Diversity measures 7
- **2.1 Ethnic measures** 7
- **2.2 Linguistic Measures** 8
- **2.3 Religious Measures** 9
- **2.4 Alternatives to Fractionalization** 11
  - 2.4.1 Polarization 11
  - 2.4.2 Group Counting 12
- **2.5 Other problems relating fractionalization** 13
  - 2.5.1 Endogeneity of diversity 13
  - 2.5.2 Spatial Distribution 14
- **2.6 Some Descriptive Data** 14

3 The Model 17
- **3.1 Economic Growth** 17
- **3.2 Initial Level of GDP** 17
- **3.3 Initial Level of Human Capital** 17
- **3.4 Fertility rate** 18
- **3.5 Government Consumption** 18
- **3.6 Democracy** 18
- **3.7 International Openness** 19
- **3.8 Investment ratio** 19
- **3.9 Inflation Rate** 19
- **3.10 Term of Trade and Rule of Law** 19

4 Results 21
- **4.1 Ethnic Fractionalization** 21
- **4.2 Linguistic Fractionalization** 23
- **4.3 Religious Fractionalization** 23
- **4.4 Other diversity variables** 25

Conclusion 26

References 27
- **Data Sources:** 29

Appendix 30

Introduction

Ever since Easterly and Levine's ground breaking research in 1996, the effect ethnic fractionalization has on economic performance has occupied the thoughts of many economists. By means of this thesis I, too, will try to make a contribution to this extensive discussion. Easterly and Levine found that many of Africa's growth "dramas" were subject to one reason: The population of Africa's nations is far more diverse than those of other –more successful- parts of the world. More recently, other scholars have confirmed these findings, reaching a consensus that a population's diversity is one of the key indicators in explaining economic performance. A major question that remains, however, is how best to quantify this diversity. Ethnicity is an ambiguous term and therefore hard to grasp in relation to economic science. The Encyclopedia Britannica Online defines ethnicity as follows:

"A Social group or category of the population that, in a larger society, is set apart and bound together by common ties of race, language, nationality, or culture."

Clearly, this leaves room for interpretation. For my own country, the Netherlands, I could come to several different sets of groups; based on language you might arrive at Dutch, Frisian, Arabic and Turkish while using race might lead to black, arabi and white; Nationality might lead to Dutch, Surinamese, Antilleans, Moroccan, Turkish and German or of course a mixture of these. The point therefore is: Ethnicity is a subjective term and, when measured by different individuals or institutions, will lead to different outcomes. The question this thesis tries to answer is whether these differences yield significantly different results when applied in economic research.

The structure of this thesis is as follows: first, I will continue by exploring existing literature in part one, after which part two will discuss different measures. Part three will present the empirical model, results of which will be presented in part four, followed by a conclusion.

1 Existing literature

This thesis is part of a larger discussion triggered by Easterly and Levine (1997) (E&L), who found that the ethnic diversity has a major influence on a nation's economic performance. The main argument for Easterly and Levine's findings is that: "…interest group polarization leads to rent-seeking behavior and reduces the consensus for public goods…" (1241). This in turn leads to low schooling, underdeveloped financial markets, distorted foreign markets, high government deficits and insufficient infrastructure, factors which, they show, are highly correlated to economic growth. In their study Easterly and Levine especially look at Africa, widely known as the worst performing continent[1] and with a history of ethnic struggles. Using a measure based on linguistic data, the Ethnic Linguistic Fractionalization index (ELF) by Taylor and Hudson (1972), they find, supports that Africa's so-called growth dramas may be largely explained by the extremely high ethnic fractionalization the continent faces.

Even though their results were widely accepted, Arcand et al (2000) where skeptic. They argued that E & L's dataset was too thin, and contained too few observations for Africa (27 out of a total 172). A second major critique was the relevance of fractionalization; as mentioned earlier, Easterly and Levine used an index called ELF, which measures ethnic diversity largely dependant on the number of ethnic groups, where Arcand et al argued that the relative sizes (polarization) of the groups are what really matters. Other scholars also found that when more traditional control variables where introduced, the effect of ethnic diversity faded (Guillaumont et al. 1998, Rodrik 1998, Sachs and Warner 1997). Convinced of his theory, William Easterly took these critiques to heart and in a later research together with Alesina, Devleeschauwer, Kurlat and Wacziarg (2003) confirmed his previous results. This time, using a larger sample for Africa and an extra measure for Polarization, which they found underperforming when compared to their fractionalization measures. Additionally they included separate measures for Linguistic and Religious fractionalization, finding that, though less effective, linguistics can be used as a proxy for ethnic diversity. They also found the relationship between religion and economic performance to be very different to that of ethnic and linguistic fractionalization. Having a greater religious fractionalization shows the country is more tolerant and free, terms that are found to have a positive effect on economic performance. Apart from this, religion is more endogenous, meaning that it is

[1] On average 0% economic growth between 1965 and 1990 (Easterly and Levine, 1997)

much more easy to hide or change your religion as it is to hide or change your skin color or accent.

Alesina and La Ferrara (2005) later managed, conducting more extensive research using the list from 2003[2] to quantify the difference in annual growth between a completely fractionalized (infinite number of groups) and a completely homogenic country to be 2 percent point. However, when Barro and Sala-i-Martin (2004) tested the ethnic, linguistic and religious fractionalization variables of Alesina et al. in a more extensive economic growth model, they found them to be insignificant. Largely, this can be explained, however, by the fact that many of the effects of diversity are channeled through some of the control variables used in economic growth research (Alesina et al. 2003).

More evidence of the link between ethnic diversity and the underprovision of public goods was given in a paper on American cities by Alesina, Baqir and Easterly. The scholars found that ethnic diversity reduces the supply of public goods by local administrations. This unwillingness of people to pay for services enjoyed by members of a different group was first uttered by Becker in 1957; he argued that it is human nature to dislike people of a different racial or ethnic group. Another more subtle view is that it is opportunist politicians that create cleavages among groups by targeting minorities and looking for scapegoats (Alesina and Gleaser 2004).

Another explanation for underprovision of public goods in ethnically diverse areas are differences in preferences making the public good not to everyone's liking. In a society with a more diverse population, preferences are bound to be further apart, making the compromise less satisfactory to a larger part of society, creating a lower willingness to pay.

For larger populations, Alesina and Spoloare (1997) show this creates a trade-off between the benefits of having a large population and the negative effects of a more diverse population[3]. A larger population makes the per capita costs of public goods lower, but since a larger population is bound to have more diverse preferences, it is harder for the government to find solutions, satisfactory to a sufficiently large part of the population. This often leads to so-called wars of attrition between the different interest groups (Alesina and Drazen 1991), where an instable and overall unsatisfactory situation is kept until one of the parties gives in and has to bear the bulk of the costs.

[2] An updated version of the dataset used by Alesina et al. (2003)

[3] Population size and the diversity of the population are found positively correlated in Alesina and Spoloare paper (1997).

Corruption is also aggravated by more ethnic diversity (Mouro 1995, Shleifer and Vishny 1993) as it leads to a poorer coordinated bribe-taking where different independent bribe-takers do not internalize the effect on the other bribe-takers' revenue, causing more and larger bribes per unit output and therefore lower output. This corruption also takes place on the labor market: In a research on Ghana, Collier and Garg (1998) found that the dominant ethnic groups gave an average 25% wage premium to public servants of their own group.

Research on the more positive aspects of ethnic fractionalization was conducted by Alesina and La Ferrara (2005): the scholars pressed upon a point not incorporated in earlier research: there is also a potential positive side to an ethnically diverse population. In an advanced economy where the labor input per worker is sufficiently high, heterogeneity of labor leads to higher outputs or at least dampens the negative effects. An earlier empirical study, by O'Reilly, Williams and Barsade (1997), supports this theory on a micro scale. They analyzed 32 project teams and found that, controlling for a lack of communication, the production of more ethnically diverse teams is higher.

Reviewing past literature: it seems scholars have found many implications of ethnic diversity for economic growth. However, there remains a great range of different diversity variables to pick from.

2 Diversity measures

Since the economic literature has reached common ground on the implications of ethnic fractionalization within countries, this thesis tries to take a closer look at the way this diversity is measured and assesses the effects different diversity measures have on economic growth. Even though some institutions might argue to quantify a similar variable, their results can be unexpectedly different. If we look at the United Kingdom, the CIA World Fact Book identifies seven distinct groups including: English, Scottish, Welsh and Irish, giving the UK a fractionalization index of 0.40. Alesina et al. (2003) - using data of the 2001 Encyclopedia Britannica- cluster these groups together yielding a fractionalization of 0.12, a significant difference on a zero to one scale. Regardless of who is right, it seems odd that such a discrepancy can even occur. However such differences are not unique, as the case of Morocco shows. The CIA World Fact Book finds Arab-Berber to be one group holding 99.1 percent of the population, the Encyclopedia Britannica (2009) however indicates them as separate groups and even includes another large group called the Moors. Again, despite of which one of the two options is correct, the point is: For institutions arguing to be measuring the same, they come up with very different results. This chapter will discuss several diversity measures available, based on ethnic, religious and linguistic differences, as well as some alternatives to the classic formula used to compute fractionalization and will end with some general problems faced when using diversity data.

2.1 Ethnic measures

In order to assess the effect of different ethic measures I will include measures by Alesina et al (2003) (EFAL), Fearon (2003) (FEAR) and two self-made measures based on data from the CIA World Factbook (WFB) (EFCIA) and the Encyclopedia Britannica (EB) (EFEB).

For their ethnic data Alesina et al. used the EB (2001 version) as a main source and completed this with data from the WFB, Levinson (1998), the minority at risk groups international and for some countries the national census, this way they tried to create a measure as disaggregated as possible.

Fearon's listing does not officially qualify as an ethnic measure as Fearon argues that: "the 'right list' of ethnic groups for a country depends on what the people in the country identify as the most socially relevant groupings." (2003, 198) He therefore created a miscellaneous list, which holds groups identified on religious and ethnic grounds, based on data of different sources like the EB, WFB, the minorities at risk database and an earlier study by Scarrit and Mozaffar (1999). As table one shows; Fearon's list is

strongly correlated with ethnic measures, more than with linguistic or religious measures[4], therefore I will treat it as such.

The data of the WFB and the EB, I included as detailed as possible, whenever both an umbrella group and a split-up in subgroups was given, I included the subgroups. I also included every possible group, irrelevant of its population share. As table one shows, correlation between the different measures is in some cases remarkably low, especially between the WFB and the EB; two institutions measuring the same can apparently come to very different lists.

Table 1: Correlation between ethnic fractionalization measures

	CIA World Fact Book	*Encyclopedia Britannica*	*Alesina et al.*
Encyclopedia Britannica	0.70	1	
Alesina et al.	0.70	0.85	1
Fearon	0.64	0.82	0.88

2.2 Linguistic Measures

For linguistic data I again turned to Alesina et al.'s (LFAL) fractionalization measure, which is based on EB (2001) data, a fractionalization measure offered by the Ethnologue Project (LFETH) and the earlier mentioned ELF (ELF). Especially the diversity measure by the Ethnologue project shows remarkably low correlation with the other two measures.

Table 2: Correlation between linguistic fractionalization measures.

	Ethnologue Project	*Alesina et al.*
Alesina et al.	0.68	1
ELF	0.70	0.90

The difference here can be explained by the fact that the data of the Ethnologue project is much more detailed than that of Alesina et al. and the ELF identifying for example 15 living languages for the Netherlands, where Alesina et al. only find four. The question is whether this detailed approach is relevant for economic research.

[4] See appendix two for a full overview of correlations between all variables.

The ELF is the original measure used by Easterly and Levine (1997) and created by Taylor and Hudson (1972). It was based on data collected in 1963 by the soviet scientists of the Atlas Narodov Mira, who used language to separate between ethnic groups. This is confirmed by the correlations in appendix two that show that ELF has more similarity with linguistic variables than it has with ethnic variables.

2.3 Religious Measures

For religious measures I again used those of Alesina et al. (RFAL) and created two measures based on WFB (RFCIA) and EB (RFEB) data. Alesina et al used the 2001 version of the Encyclopedia Britannica to form their measure. Interestingly, though I use the same source 8 years later, our measures only have a correlation of 0.81, as displayed in table three. It seems rather puzzling how this measure can change over such a short period of time. This leaves three different possibilities: Either Alesina et al. and I processed the data differently, religious fractionalization is endogenous and has changed over time or the EB has changed its selection criteria.

Table 3: Correlation between religious fractionalization measures

	CIA World Fact Book	*Encyclopedia Britannica*
Encyclopedia Britannica	0.75	1
Alesina et al.	0.77	0.81

For the first possibility: as with the ethnic data I again took the most disaggregate selection of groups as possible, to show the relevance of this look at the example of religious affiliation in Lebanon according to the EB:

"Religious affiliation (1995): Muslim 55.3%, of which Shii 34.0%, Sunni 21.3%; Christian 37.6%, of which Catholic 25.1% (Maronite 19.0%, Greek Catholic or Melchite 4.6%), Orthodox 11.7% (Greek Orthodox 6.0%, Armenian Apostolic 5.2%), Protestant 0.5%; Druze 7.1%."[5]

Had I included umbrella groups, I would have ended up with 3 religious groups (Muslim, Christian and Druze) with a fractionalization index of 0.55, but specifying as much as possible yields 10 groups with a fractionalization index of 0.79. Alesina et al. do not state how they selected their religious groups, but since their ethnic data was

[5] Encyclopedia Britannica Online, World Data Analyst, Lebanon.

specifically mentioned to be as disaggregated as possible, it seems logic for them to do the same with religion.
For the second option, even though I do think religious diversity can change over time, as Alesina et al. already found, it is relatively easy to change or hide your religious affiliation. However, I do not think worldwide major changes can occur over an eight-year time span. Most likely, therefore, seems the conclusion that the EB provided different data in 2001 and 2009[6]. This seems more likely when one looks at the WFB data, which has even lower correlation with the other two measures.

Another problem with religious terms is that they appear to be biased. Where the Christian faith is often split up with great detail (by both the WFB and the EB), even identifying smallest sects, African and Asian indigenous beliefs are usually clustered in one group, sometimes even in combination with unknown or others. To illustrate this, the WFB describes the religious structure of Trinidad and Tobago as follows;

Religions: Roman Catholic 26%, Hindu 22.5%, Anglican 7.8%, Baptist 7.2%, Pentecostal 6.8%, Muslim 5.8%, Seventh Day Adventist 4%, other Christian 5.8%, other 10.8%, unspecified 1.4%, none 1.9% (2000 census) [7]

Identifying 5 distinct Christian beliefs, the indigenous beliefs of Togo on the other hand are not specified at al;

Religions: Christian 29%, Muslim 20%, indigenous beliefs 51%[8]

Other than in most cultures, it is common in Japan to have two religions; over 80 percent of the Japanese population is Shintoist, of which a majority also practices Buddhism (70 percent of the total population). While this makes it impossible to calculate fractionalization using the classic Herfindahl index, Japan is excluded in all religious data.

Appendix two shows the correlations between all different fractionalization variables, while the linguistic and ethnic variables show great correlation, those based on religion are only slightly correlated to the others. This confirms the conclusion of Alesina et al.

[6] This was confirmed by the EB; "the preferred organization of data has changed over time as sources deemed superior or equally good have become available. Improvements are never ending."
[7] CIA World Factbook: Trinidad and Tobago
[8] CIA World Factbook: Togo – Year of data selection not indicated

that religious diversity is and functions very differently from ethnic and linguistic diversity.

2.4 Alternatives to Fractionalization

Fractionalization is usually measured using a Herfindahl index, as is the case with the ELF measure used by Easterly and Levine. The Herfindahl index calculates the chance that two randomly chosen inhabitants of a country belong to separate groups within society. The formula therefore looks as follows;

$$ELF = 1 - \sum_{i=1}^{n} s_i^2;$$

where i stands for a group and s_i for the share of the population this group holds. Lately many scholars have debated whether or not fractionalization is the best variable suited to explain economical performance, Arcand et al (2000), Alesina et al (2003) and Fearon (2003) use polarization as an alternative. Another alternative might be to actually count the number of groups present in society and add this number to the regression.

2.4.1 Polarization

The theory behind using polarization is quite straightforward: in the Herfindahl index much weight is given to the number of groups. However, the size of the different groups as well as the cultural distance between them is hardly considered. The Herfindahl based fractionalization is typically maximized when there is an infinite number of groups (fractionalization = 1); however, a society with only two equal sized groups (fractionalization = 0.5) might lead to more friction among the groups struggling for dominance. As a contemporary example we might look at the 2007 Belgium government formations, where political struggles between Flemish and Wallonian interest groups led to 194 days' lasting government formations.

The distance between the different groups refers to what Posner (2004) calls the depth and dimension of cleavage: how much do the different groups differ from each other.

Already in 1994 Esteban and Ray devised a formula for measuring polarization;

$$p_{(s,y)} = K \sum_{i=1}^{n} \sum_{j=1}^{n} s_i^{1+\alpha} s_j \left| y_i - y_j \right|;$$

where K is a constant, $\propto$ a constant between 0 an d 1.6, s_i and s_j are the population shares of two different ethnic groups and $\left| y_i - y_j \right|$ represents the depth of cleavage.

What is missing, however, in most empirics is a measurement for depth of cleavage: therefore, in most studies considering polarization this is presumed to be equal between all groups.

In their paper, Alesina et al. (2003) used this method in their study and compared results of both fractionalization and polarization in the growth model of E&L. They find results similar to fractionalization with an $\propto$ that leads to a value of polarization highly correlated with fractionalization, and worse when $\propto$ is chosen so that correlation is low.

Fearon (2003) did manage to include depth of cleavage by assessing the difference in mother tongues by a adding a factor r_{ij} to the standard Herfindahl index used to calculate fractionalization. This value is close to zero for two completely different languages and close to one for similar languages. The formula, which calculates what Fearon calls Cultural Diversity (CDIV), then becomes;

$$CDIV = 1 - \sum_{i=1}^{n}\sum_{j=1}^{n} s_i s_j r_{ij};$$

where s_i and s_j stand for the population shares of different groups in society. Even though Fearon manages to grasp the deph of cleavage, relative group sizes are not included in this measure. Table four shows the correlation between the cultural diversity measure of Fearon and ethnic fractionalization measures is relatively low, even with Fearon's own variable. In order to assess the added value of the cultural diversity variable, I will include it in my empirical research.

Table 4: Correlation between ethnic measures and Cultural diversity measure by James D. Fearon.

CIA World Fact Book	*0.68*
Encyclopedia Britannica	0.67
Alesina et al.	0.76
Fearon	0.79

2.4.2 Group Counting

A fairly simple and straightforward way to look at fractionalization is counting the number of groups, however it is still interesting to find out what adding such a variable to a growth model yields. Therefore, I will include this for the variables computed from EB (EGEB, RGEB) and WFB data (EGCIA, RGCIA). In order to get a number as accurate as possible I have considered the label "other" to be two groups, as this is the minimum

that suits the description. Labels similar to unknown, unspecified or synonyms of this I did not count, as they may well consist of unknown members of the described groups.
With pure group counting, every group is counted equal, though groups with a low population share are expected to have a smaller influence. By taking the inverse of the Herfindahl index, however, I retain an approximation of the relevant number of groups. As the formula below shows;

$$RG = \frac{1}{\sum_{i=1}^{n} s_i^2}$$

The following example can illustrate this: the EB describes the ethnic build-up of France as follows;

Ethnic composition (2000): French 76.9%; Algerian and Moroccan Berber 2.2%; Italian 1.9%; Portuguese 1.5%; Moroccan Arab 1.5%; Fleming 1.4%; Algerian Arab 1.3%; Basque 1.3%; Jewish 1.2%; German 1.2%; Vietnamese 1.0%; Catalan 0.5%; other 8.1%.[9]

Simple group counting would yield 14 groups, but computing the relevant number of groups gives 1.67 since there is a huge majority of ethnic French and many small minorities. The relevant number of groups is included for the data of the WFB (EICIA, RICIA) and EB (EIEB, RIEB).

2.5 Other problems relating fractionalization

This paragraph briefly stops at some difficulties involved in measuring fractionalization: is it stable over time and does the spatial distribution matter?

2.5.1 Endogeneity of diversity

The ELF measure, used by E&L, exists of data gathered in the early sixties and we can therefore assume that E&L thought ethnic fractionalization to be static over time. The question may however be raised if ethnic fractionalization is really exogenously given or might it be subject to changes over time. A good example of this is the Netherlands; during the creation of the ELF measure Dutch society was vertically divided in several groups based on religion or ideologies. During the sixties and seventies this

[9] Encyclopedia Britannica Online, data analyst, France.

"pillarisation" almost completely vanished, while at the same time migration of guest workers led to the creation of new groups. Therefore, one might argue that, over the course of time, the build-up of Dutch society has changed.

Alesina et al. (2003) also push upon this issue and find ethnic fractionalization to be endogenous to some extend but do not consider it a problem at a 20 to 30 year horizon. More so, since the number of cases in which a society has gone through substantial changes are very limited. Following the vision of Alesina et al. this thesis will also consider ethnic fractionalization to be exogenous, making it possible to apply fractionalization figures to several time periods.

2.5.2 Spatial Distribution

Daniel Posner (2004) points out the lack of spatial distribution in fractionalization measures, arguing that when different groups share a living environment, this leads to other outcomes than if they where in different regions of the country. If we take a further look at the example of US jurisdictions (Alesina, Baqir and Easterly, 1999), apparently people are less willing to pay for proper sewer, roads, schools etc, if this is also to the benefit of other ethnicities; Thus creating a situation where public good provision is at an inefficient low level.

The case of ethnic groups being separated in different parts of the country can be illustrated by the case of Ghana, where dominant Akan groups made macro economical policy such that cocoa, which was mainly produced by the Ashanti groups, was heavily taxed at its own benefit. This continuing struggle for the benefits of the country's most important export good led to cocoa prices falling from 89% of the world price in 1949 to 6% in 1983, and dropping cocoa exports from 19% of GDP in 1955 to 3% in 1983 (Bates 1981; Easterly and Levine 1997). In Ghana, therefore, population diversity led to bad macro economical policy. Even though I recognize the effects of spatial distribution, the sheer complexity of measuring it, forces me to exclude it in this thesis.

2.6 Some Descriptive Data

We have now gathered 18 different measures of diversity (appendix one gives an overview) covering religion, ethnicity and language based on five sources using four different methods. Table five shows an overview of them, with the average per world region. The first nine represent the conventional measures of fractionalization and the last are the nine less orthodox measures: polarization, groups and relevant groups.

Table 5: Averages Per World Region

	EFCIA	*EFEB*	*EFAL*	*FEAR*	*RFCIA*	*RFEB*	*RFAL*	*LFAL*	*LFETH*
Asia	0.42	0.45	0.39	0.45	0.41	0.51	0.38	0.46	0.60
Former USSR and Eastern Europe	0.31	0.32	0.37	0.38	0.41	0.53	0.47	0.35	0.37
Latin America	0.39	0.47	0.44	0.41	0.35	0.49	0.34	0.17	0.32
Middle East and Northern Africa	0.24	0.54	0.42	0.45	0.21	0.35	0.24	0.27	0.52
Sub Sahara Africa	0.52	0.69	0.69	0.72	0.55	0.60	0.56	0.68	0.71
Western World and Japan	0.30	0.30	0.22	0.24	0.43	0.51	0.45	0.22	0.28

	CDIV	EGCIA	EICIA	EGEB	EIEB	RGCIA	RICIA	RGEB	RIEB
Asia	0.32	6.63	2.38	8.28	2.42	5.89	2.08	6.89	2.53
Former USSR and Eastern Europe	0.29	6.19	1.55	7.42	1.57	5.54	1.93	7.19	2.52
Latin America	0.19	4.45	1.89	6.19	2.12	5.57	2.17	6.95	2.36
Middle East and Northern Africa	0.28	4.06	1.54	7.41	2.92	3.4	1.36	6.24	1.89
Sub Sahara Africa	0.43	6.56	2.98	9.05	4.27	4.4	2.66	6.19	3.09
Western World and Japan	0.19	5.75	1.61	8.05	1.64	6.65	2.42	9.45	2.82

Many scholars[10] before found Sub-Sahara Africa to be the most diversified region of the world; most of the measures above reinforce this observation. Only with group counting do other regions seem more diverse, especially considering religion. However, when we look at the relevance of these fractions, we find that it concerns mostly small minorities, resulting in a lower number of relevant groups compared to Sub-Sahara Africa. Partly, this might be explained by the biasness of religious data towards Christian beliefs, which was mentioned in paragraph 2.5. In line with these expectations is the fact that the Western World and Japan is the least diverse by all ethnic and linguistic fractionalization measures.

To be expected after the surprisingly low correlations we found before are the large differences in averages between the different measures. Especially when considering

[10] Alesina et al. (2003), Alesina and La Ferrara (2004), Easterly and Levine (1996),Fearon (2004), Scarritt and Mozaffar (1999)

the number of groups counted by the WFB and the EB, it becomes obvious that the EB distinguishes a much larger number of groups.

Taking into consideration the correlations found before as well as the descriptive statistics given above, there is reason to believe there might be implications of all these discrepancies for empirical research. The next chapters will therefore incorporate these different measures in an empirical growth model.

3 The Model

Robert Barro presented his growth model in his 1991 paper "Economic Growth in a Cross Section of Countries". In this study he found support for the neoclassical growth models by Solow-Swan and Ramsey who predicted a faster growth by countries coming from a lower initial GDP per capita. This phenomenon, known as absolute convergence, was disputed by other scholars who pointed at the lack of correlation between initial income and growth. Barro showed empirically that, when initial human capital is held constant, a faster growth of poor countries is found. He called this phenomenon conditional convergence. In this paper I will use a simplified version of Barro's model, which will however be sufficient to show the different effects of the various diversity measures.

This empirical model will cover three periods: the seventies, eighties and nineties, and includes a total of 147 countries (19 Asian, 17 from the middle East and North Africa, 26 of the former USSR and Eastern Europe, 21 Latin American or Caribbean, 43 Sub-Sahara African and 21 from the Western world and Japan). The next paragraphs will outline in short the control variables included in the model but not before first discussing the dependent variable: economic growth.

3.1 Economic Growth

As is common in economic research, economic performance is measured by the Gross Domestic Product (GDP) per capita growth. Following Barro (1991), this variable is included as the average growth over the ten-year period.

3.2 Initial Level of GDP

The Initial level of GDP (INITIAL) enters the regression as the log of GDP per capita in the starting year of each period (1970, 1980 and 1990). Barro and Sala-i-Martin (2004) calculated a coefficient of approximately -0.025, which in their model meant a 0.026 increase of growth with a one standard-deviation decline of initial GDP. This indicates the potentially large effect convergence has on economic growth rates. The coefficient in my model is significant and has a value of -0.037.

3.3 Initial Level of Human Capital

One of the main reasons vital for Barro to accept the theory of convergence, was the fact that he found that countries where not competing at a similar level. Where the

population of rich countries was generally well educated and enjoyed a healthy life that contrary to that of poor countries. This initial level of human capital is the main reason why absolute convergence does not take place (Barro 1991). I therefore included two variables as proxy for human capital; the percentage of men who attended secondary schooling (EDU) and the reciprocal life expectancy at birth (LIFE). Both are measured at the start of each ten-year period.

The first variable pictures the educational level of the entire population. The reason why female education is excluded is based on the large correlation this would have with the later included variable of fertility. Due to the availability of data I slightly diverge from the original model where the average years of male secondary and tertiary education was taken. From the reciprocal life expectancy at birth the inverse is taken so it illustrates the average chance of death per year. Though Barro found life expectancy at age one to perform better, again available data forced me to use life expectancy at birth. Both male education (-0.037) and the reciprocal life expectancy (-4.35) variables are significant and show the expected negative sign. Therefore, my model might confirm that a better-educated, healthy population creates more growth.

3.4 Fertility rate

A growing population spends part of its investment funds on providing new workers with new capital and part of its resources on raising children. For this reason a negative effect on economic growth is commonly expected (Barro 1996). The fertility rate enters as the log of the lifetime fertility (FERT) at the starting date of each period. As expected, fertility has a negative influence (-0.057) in the model, the value is significant.

3.5 Government Consumption

Through the effect of non-productivity advancing government spending and taxation, a negative effect is expected from government spending (GOV). Contrary to Barro and Sala-I-Martin, I am not correcting government spending for educational and defense expenditures, which may cause some correlation with male secondary education. Government spending enters my equation significantly and as expected, with a negative sign

3.6 Democracy

Following Barro and Sala-I-Martin, the variable of democracy (DEMO) comes in as a subjective measure created by the freedom house index, which rates the quality of

electoral rights on a one to seven scale. Subtracting the inverse of the variable from one, gave the value included. Both the average of the period and the square of the average are put into the regression. The first is expected to yield a positive effect and the second a negative, implicating a positive effect for nations going from a totalitarian regime to a more democratic government; however, negative effects for more advanced democracies, arguable due to the increasing complexity and sluggishness of the decision making process in highly developed democracies (Barro and Sala-i-Martin 2004). Other than with Barro, both my variables for democracy are insignificant.

3.7 International Openness

International openness (OPEN) is computed as the ratio of imports plus exports to GDP and enters the equation as the average over the ten-year period. Contrary to Barro and Sala-i-Martin's model, the variable that enters here is not corrected for population size and landmass, which makes it more sensitive to country size. This might be the cause for my measure is significant (at 5% level) where it was not with Barro and Sala-i-Martin. The sign however is positive as with Barro and Sala-i-Martin.

3.8 Investment ratio

Investment (INV) raises the output per worker, and, as a consequence, is estimated to have a positive affect on growth. Investment is included as the average of the ratio between real gross domestic investments to GDP over each of the three periods. The effect is found to be positive and significant.

3.9 Inflation Rate

Inflation (INFL) enters the growth equation as the average over each ten-year period, and is expected to have a negative affect on the international competitiveness of the country and, therefore, to slow economic growth. My model shows both a negative and significant effect.

3.10 Term of Trade and Rule of Law

Good regressors for "terms of trade" and "rule of law" were not available to me and are therefore neglected in this empirical study. Although I recognize the importance of these control variables, I don't expect their absence to be a major problem for this empirical study. The number of variables included should be sufficient to distill the effect of the different diversity figures.

A constant and dummies for the eighties and nineties are also included in the model. The dummies are negative and significant at 1%, which means that, although worldwide economic growth has stagnated since the seventies it has remained stable during the eighties and nineties.

Table six shows an overview of all variables with their coefficients, t-values and expected signs. It also shows a restricted version of the model that only includes initial GDP, male education, life expectancy plus the constant and dummy variables. This restricted model enables us to see the effects created by diversity when not controlled by other variables.

Table 6: Restricted and unrestricted growth model

Dependant variable: GDP per capita growth

	Unrestricted Model			*Restricted Model*			
Variable	*Coefficient*		*T Statistic*	*Coefficient*		*T Statistic*	*Expected Sign*
Constant	0.280	*	7.835	0.231	*	7.704	+
Log (Initial Income)	-0.037	*	-6.602	-0.026	*	-4.986	-
Male Education	-0.023	**	-2.381	-0.004		-0.522	+
1/Life expectancy at birth	-4.346	*	-4.844	-6.089	*	-7.391	-
Log (Fertility)	-0.057	*	-3.802				-
Democracy	-0.027		-1.174				+
Democracy^2	0.032		1.244				-
International Openness	0.009	**	1.971				+
Government spending	-0.034	**	-1.989				-
Investment share	0.053	**	2.502				+
Inflation	-0.003	*	-4.947				-
Dummy for Eighties	-0.021	*	-5.199	-0.023	*	-5.766	-
Dummy For Nineties	-0.021	*	-4.857	-0.024	*	-5.810	-
R-squared	0.3387			0.200			

*** Significant at 1 percent level**

**** Significant at 5 percent level**

4 Results

Before coming to the actual core of this thesis, discussing results of the empirical growth model, let us first take a look at the correlation of the different diversity measures with the control variables (appendix 2). In line with earlier studies[11] the ethnic and linguistic measures are correlated highly negatively to GDP per capita growth and education, as well as with democracy and investment. The religious fractionalization measures do not show any distinctive patterns of correlation with any of the other variables.

4.1 Ethnic Fractionalization

Equation one shows the basic model once more and equations 2 to 11 show the different ethnic fractionalization measures and the measures of Fearon (2003). The even numbered equations show the model without control variables, the odd numbered ones show the variable when added to the complete model.

Without control variables, the measure based on Encyclopedia Britannica data, and those of Alesina et al and Fearon have a significant effect. Not surprisingly, table one already showed these variables to be most tightly correlated.

When control variables are included, only the measure by Alesina et al. stays significant, though at a lower level as before. This confirms the findings of both Easterly and Levine (1996) and Alesina et al (2003) that the effect of ethnic diversity is largely channeled through some of the control variables. Bearing in mind the correlation between our ethnic variables and control variables like education, democracy and investment, it is most likely that these variables act as the channels.

More important when keeping this thesis' objective in mind are the differences among the different ethnic variables. Here, especially, EFCIA performs very differently from the others. It is highly insignificant with a t-statistic of 0.457 (appendix 3) and a positive sign, which seems contrary to the recent economic literature. Also the cultural diversity by Fearon underperforms when compared to the other three variables.

[11] Easterly and Levine 1996, Alesina et al. 2003

Table 7: Dependent variable: GDP per capita growth

	1		*2*		*3*		*4*		*5*		*6*		*7*		*8*		*9*		*10*		*11*	
C	0.280	*	0.238	*	0.294	*	0.228	*	0.277	*	0.229	*	0.274	*	0.230	*	0.278	*	0.228	*	0.279	*
INITIAL	-0.037		-0.029	*	-0.040	*	-0.025	*	-0.036	*	-0.003	*	-0.037	*	-0.027	*	-0.037	*	-0.026	*	-0.037	*
EDU	-0.024	**	-0.001		-0.024	**	-0.009		-0.024	**	-0.008		-0.022	**	-0.007		-0.023	**	-0.004		-0.023	**
LIFE	-4.346	*	-6.092	*	-4.513	*	-5.650	*	-4.267	*	-5.167	*	-3.976	*	-5.361	*	-4.133	*	-5.702	*	-4.283	*
FERT	-0.057	*			-0.060	*			-0.057	*			-0.050	*			-0.054	*			-0.056	*
DEMO	-0.027				-0.036				-0.029				-0.024				-0.024				-0.026	
DEMO^2	0.032				0.046				0.034				0.029				0.029				0.030	
OPEN	0.009	**			0.009	***			0.01	**			0.010	**			0.010	**			0.009	***
GOV	-0.034	**			-0.035	***			-0.037	**			-0.035	**			-0.034	**			-0.035	**
INV	0.053	**			0.057	**			0.053	**			0.047	**			0.047	**			0.051	**
INFL	-0.003	*			-0.002	**			-0.003	*			-0.003	*			-0.003	*			-0.003	*
EFCIA			0.003		0.009																	
EFEB							-0.017	**	-0.002													
EFAL											-0.027	*	-0.015	**								
FEAR															-0.020	*	-0.010					
CDIV																			-0.013		-0.005	
DUM80	-0.021	*	0.023	*	-0.023	*	-0.023	*	-0.022	*	-0.022	*	-0.02	*	-0.022	*	-0.021	*	-0.022	*	-0.021	*
DUM90	-0.021	*	0.023	*	-0.023	*	-0.023	*	-0.022	*	-0.021	*	-0.02	*	-0.021	*	-0.021	*	-0.023	*	-0.021	*
R-squared	0.339		0.207		0.322		0.212		0.342		0.233		0.348		0.215		0.343		0.204		0.337	

***** **Significant at 1 percent level**
****** **Significant at 5 percent level**
******* **Significant at 10 percent level**

The main question now is whether the coefficients of the restricted models for EFEB, EFAL and FEAR are statistically different from each other. Conducting a Wald coefficient restriction test shows that the three coefficients are not significantly different.
The variable EFAL keeps having a significant effect even when control variables are included, and rises the R-squared with about 0.009.

4.2 Linguistic Fractionalization

Equations 12 to 17 in table eight show the output when including linguistic fractionalization; here the odd numbered equations show the restricted model and the even numbered equations show the complete model. Linguistic fractionalization is used mostly as a proxy for ethnic fractionalization, not surprisingly therefore the linguistic variables underperform when compared to the "real" ethnic variables. All linguistic variables come out insignificant in the restricted as well as the full model.
Most surprisingly might be the insignificance of the ELF measure, also used in Easterly and Levine's research that triggered the discussion regarding this topic offering the first proof that there actually is a link between diversity and economic performance.
Also notable is that, although the variables based on data from the Ethnologue Project and the Encyclopedia Britannica showed a fairly poor correlation: they perform equally poorly, both having p-values of around 0.2 (appendix 4).

4.3 Religious Fractionalization

Equations 18 to 23 in table eight show the model's results when religious fractionalization measures are included. Just as with the linguistic variables, none of the religious variables is significant either in the restricted or the unrestricted model. This is in line with the outcomes of Alesina et al. (2003) who found their religious variable performing least in explaining economic performance. Its endogenous character largely explains this: It is relatively easy to hide your religion (as opposed to race or language) added by the fact that religious diversity may be seen as an indicator for tolerance. Since tolerance can have a positive effect on economic performance this may even out the negative effect of diversity.

Table 8: Dependent variable: GDP per capita growth

	12		*13*		*14*		*15*		*16*		*17*		*18*		*19*		*20*		*21*		*22*		*23*	
C	0.231	*	0.280	*	0.245	*	0.289	*	0.234	*	0.275	*	0.218	*	0.261	*	0.234	*	0.284	*	0.233	*	0.284	*
INITIAL	-0.027	*	-0.037	*	-0.028	*	-0.039	*	-0.033	*	-0.038	*	-0.024	*	-0.036	*	-0.027	*	-0.038	*	-0.026	*	-0.038	*
EDU	-0.005		-0.023	**	-0.007		-0.024	**	0.014		-0.014		-0.003		-0.020	***	-0.005		-0.024	**	-0.004		-0.024	**
LIFE	-5.831	*	-4.385	*	-6.208	*	-4.410	*	-5.274	*	-3.994	*	-5.623	*	-3.905	*	-6.105	*	-4.340	*	-6.080	*	-4.344	*
FERT			-0.057	*			-0.055	*			-0.061	*			-0.052	*			-0.058	*			-0.059	*
DEMO			-0.028				-0.027				-0.017				-0.015				-0.027				-0.027	
DEMO^2			0.033				0.032				0.019				0.022				0.033				0.033	
OPEN			0.009	***			0.009	***			0.008	***			0.008				0.010	**			0.010	**
GOV			-0.034	**			-0.036	**			-0.018				-0.031	***			-0.034	***			-0.034	**
INV			0.054	**			0.049	**			0.042	**			0.055	**			0.052	**			0.050	**
INFL			-0.003	*			-0.003	*			-0.003	**			-0.003	*			-0.003	*			-0.003	*
LFETH	-0.006		0.002																					
LFAL					-0.009		-0.008																	
ELF									-0.009		-0.003													
RFCIA													-0.007		-0.008									
RFCIA																	-0.003		-0.005					
RFAL																					-0.003		-0.004	
Dum80	-0.023	*	-0.021	*	-0.023	*	-0.021	*	-0.021	*	-0.019	*	-0.023	*	-0.019	*	-0.024	*	-0.021	*	-0.023	*	-0.021	*
Dum90	-0.023	*	-0.021	*	-0.024	*	-0.021	*	-0.020	*	-0.018	*	-0.022	*	-0.019	*	-0.024	*	-0.021	*	-0.024	*	-0.022	*
R-squared	0.200		0.339		0.208		0.342		0.232		0.363		0.186		0.324		0.200		0.339		0.200		0.340	

*** Significant at 1 percent level**
**** Significant at 5 percent level**
***** Significant at 10 percent level**

4.4 Other diversity variables

Appendix five and six include two tables containing the results for other measures assessed in this thesis. Simply counting the number of groups present in a country (regardless of size) seems to be irrelevant in explaining economic performance. Taking the relevant number of groups also underperforms compared to previously discussed variables for religion and ethnic diversity.

The results from this empirical research show a few major aspects: First, it confirms findings of earlier papers that ethnic diversity negatively affects economic growth; the results for linguistic and especially religious fractionalization, on the other hand, are much less conclusive. It also shows that the traditional fractionalization variables perform better than the alternatives offered.
However, more importantly, regarding the objective of this thesis is that the results show that the low correlations in chapter two do not implicitly lead to different outcomes in the growth model.

Conclusion

Ethnic diversity is a rather miscellaneous term and can be measured with very different results at different occasions. This was confirmed in part two of this thesis where I showed that correlations among different measures are surprisingly low, considering that they attempt to measure the same aspects. Linguistic and religious diversity have tighter descriptions, yet here also, the differences among various measures are large. The linguistic fractionalization variables of the Ethnologue Project and those of Alesina et al., for example, only have a joint correlation of 0.68, astonishingly low for two renowned institutions aiming to measure the same.

Added to an empirical growth model, two things become immediately clear; first of all, fractionalization measures perform better than the offered alternatives and secondly, only ethnic variables yield any significant effect. Even though the correlations of the ethnic variables were low, those based on Encyclopedia Britannica data had similar effects in an uncontrolled model. When control variables were included, only those of Alesina et al. (2003) remained significant.

This supports earlier research by both Alesina et al (2003) and Easterly and Levine (1996) who found that the effect of ethnic diversity is largely channeled through traditional control variables. This also explains why, in a more extensive model, Barro and Sala-i-Martin (2004) found an insignificant coefficient for the variable of Alesina et al. (2003).

The question that now remains is whether this matters, part of this touches on the essence of science. A basic demand of science is that results have to be reproducible and factors to be stable. Clearly, this is not the case with ethnic diversity, the definition of ethnicity is simply too vague. This is what makes the application of ethnic diversity; in economic research so difficult, there is no right way to measure it and therefore there is also no absolutely correct variable. It is never more than an indication and always subject to debate and doubt.

For research application purposes, this means it is impossible to give absolute, quantified results regarding the effects of ethnic diversity. One has to bear in mind there is always a certain amount of uncertainty in the term and that, just as the variable, results are an indication. However, this does not mean that ethnic diversity is unsuited as a variable for economic research, it is just not possible to quantify results reliably. As long as scholars keep this in mind, there might be numerous applications for explaining economic variables such as the “channel” variables in this thesis.

I would like to make a last remark regarding future research: it seems rather odd that we treat diversity similar over all countries, despite their sizes. Still effects in Sri Lanka or Luxemburg might be very different from those in the US or Russia, this might be subject of later research. This brings me back to the most important message this thesis carries; diversity variables are complex and results should be approached with restraint.

References

Alesina, A., R. Baqir and W. Easterly (1999), "Public Goods and Ethnic Divisions" Quarterly Journal of Economics 114(4), 1243-1284.

Alesina, A., A. Devleschauwer, W. Easterly, S. Kurlat and R. Wacziarg (2003), "Fractionalization", Journal of Economic Growth, 8, 155-94.

Alesina, A., A. Drazen (1991) "Why are Stabilizations Delayed?" The American Economic Review, 81 (5), 1170–1188.

Alesina, A. and E. Glaeser (2006), Fighting poverty in the US and Europe: a world of difference, New York, Oxford University Press.

Alesina, A, E. La Ferrara (2005), "Ethnic Diversity and Economic Growth", Discussion Paper, 2028, Harvard University.

Alesina, A. and E. Spolaore (1997), "On the Number and Size of Nations," Quarterly Journal of Economics, 112 (4), 1027-56.

Arcand, J., P. Guillaumont and S. Guillaumont Jeanneney, (2000), "How to Make a Tragedy; On the Alleged Effect of Ethnicity on Growth", Journal of International Development, 12, 925-938.

Bates, R. H. (1981), "Markets and States in Tropical Africa: the Political Basis of Agricultural Policies", Berkeley, CA: University of California Press.

Barro, R.J., X. Sala-i-Martin (2004), "Economic Growth", Second Edition, Cambridge, The MIT Press.

Barro, R.J. (1996), "Determinants of Economic Growth: A Cross-Country Empirical Study", NBER Working Paper 5698.

Becker, G. (1957), "The Economics of Discrimination" Chicago, University of Chicago Press.

Easterly, W., R. Levine (1997), "Africa's Growth Tragedy: "Policies and Ethnic Divisions", quarterly journal of economics" 111(4), 1203-1250).

Esteban, J. and D. Ray (1994), ``On the Measurement of Polarization'' Econometrica 62(4), 819-851.

Fearon, J. (2003) "Ethnic and cultural diversity by country", Journal of Economic Growth, 8 (2), 195–222.

Guillaumont, P., S. Guillaumont Jeanneney, J. F. Brun (1999) "How instability lowers African Growth" *Journal of African Economies*, March, 87-107

Levinson, D. (1998) "Ethnic Groups Worldwide, A Ready Reference Handbook" Phoenix, Oryx Press.

Mauro, P. (1995) "Corruption and Growth", Quarterly Journal of Economics, vol. 110, no. 3, August, 681-712.

Mozaffar, S. and J. R. Scaritt (1999), "The Specification of Ethnic Cleavages and Ethnopolitical groups for the Analysis of Democratic Competition in Contemporary Africa", Nationalism and Ethnic Politics, 5:1, 82 — 117.

O'Reilly, C. L., K. Y. Williams, and S. G. Barsade (1997) "Demography and Group Performance." Unpublished.

Posner, D. N. (2004), "Measuring Ethnic Fractionalization in Africa", American Journal of Political Science, Vol. 48, No 4, pp 849 – 863.

Rodrik, D. (1998) "Trade policy and economic performance in Sub-Saharan Africa" NBER Working Paper, No 6562.

Sachs, J., A. Warner (1997) "Sources of slow growth in African economies" Journal of African Economies 6, 335-376.

Taylor, C.L, M.C. Hudson (1972), " World Hand Book Of Political and Economic Indicators", Yale University Press.

Deutsch, K.W. (1953) "Nationalism and Social Communication" New York: Wiley.

Data Sources:

Central Intelligence Agency (2009) "The World Factbook"
https://www.cia.gov/library/publications/the-world-factbook/

Encyclopedia Britannica (2009) "World Data Analyst"
http://search.eb.com/analyst

Ethnologue Project (2009) "Statistical Summary"
http://www.ethnologue.com/ethno_docs/distribution.asp?by=country

Heston, A., Robert Summers and Bettina Aten, (August 2009) Penn World Table Version 6.3, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.
http://pwt.econ.upenn.edu/

International Monetary Fund (2009) "Principal Global Indicators"
http://www.principalglobalindicators.org/

World Bank (2009) "World Development Indicators Online"
http://web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/0,,contentMDK:20398986~menuPK:64133163~pagePK:64133150~piPK:64133175~theSitePK:239419,00.html

Appendix

Appendix 1: Data Sources and Definitions

GROWTH	Average annual GDP (constant prices) growth over ten-year period. Source: Penn World Table Version 6.3
INITIAL	Log of initial GDP (constant prices) per capita of starting year of period, measured at start of the ten-year period. Source: Penn World Table Version 6.3
EDU	Percentage of male citizens who enjoyed secondary schooling, measured at start of the ten-year period. Source: World Development Indicators Online
LIFE	Inverse of life expectancy at birth, measured at start of the ten-year period. Source: World Development Indicators Online
FERT	Log of average fertility rate, average over ten-year period. Source: World Development Indicators Online
GOV	Government consumption as part of GDP Source: Penn World Table Version 6.3
DEMO	One minus the inverse of the "political rights score" computed by the Freedom House. Average over ten-year period. Data about seventies taken from 1972 onwards. Source: Freedom House: Freedom in the world country rating,
OPEN	Imports plus exports as a ratio of GDP (constant prices). Average over ten-year period. Source Penn World Table Version 6.3
INV	Investment as a ratio of GDP (constant prices). Average over ten-year period. Source: Penn World Table Version 6.3

INFL	Price inflation, average of ten-year period. Source: World Development Indicators Online and IMF: Principal Global Indicators
DUM80	Dummy for the eighties.
DUM90	Dummy of the nineties
EFCIA	Ethnic Fractionalization computed as disaggregated as possible based on data of the CIA World Fact Book.
EFEB	Ethnic Fractionalization computed as disaggregated as possible based on data of the Encyclopedia Britannica World Data Analyst
EFAL	Ethnic Fractionalization computed by Alesina et al. (2003).
FEAR	Fractionalization computed by Fearon (2003)
CDIV	Cultural Diversity computed by Fearon (2003)
LFETH	Linguistic Fractionalization computed by the Ethnologue Project
LFAL	Linguistic Fractionalization computed by Alesina et al. (2003)
ELF	Ethnic Linguistic Fractionalization computed by Taylor and Hudson (1972)
RFCIA	Religious Fractionalization computed as disaggregated as possible based on data of the CIA World Fact Book.
RFEB	Religious Fractionalization computed as disaggregated as possible based on data of the Encyclopedia Britannica World Data Analyst
RFAL	Religious Fractionalization computed by Alesina et al. (2003)
EGCIA	Number of ethnic groups based on CIA World Fact Book Data.
EGEB	Number of ethnic groups based on Encyclopedia Britannica World Data Analyst data.
EICIA	Relevant number of ethnic groups, computed as

	the inverse of a Herfindahl index, based on CIA World Factbook Data.
EIEB	Relevant number of ethnic groups, computed as the inverse of a Herfindahl index, based on Encyclopedia Britannica World Data Analyst data.
RGCIA	Number of religious groups based on CIA World Fact Book Data.
RGEB	Number of religious groups based on Encyclopedia Britannica World Data Analyst data.
RICIA	Relevant number of religious groups, computed as the inverse of a Herfindahl index, based on CIA World Factbook Data.
RIEB	Relevant number of religious groups, computed as the inverse of a Herfindahl index, based on Encyclopedia Britannica World Data Analyst data.

Appendix 2: Correlations

	GROWTH	*INITIAL*	*EDU*	*LIFE*	*FERT*	*DEMO*	*OPEN*	*GOV*	*INV*	*INFL*	*EFCIA*	*EFEB*	*EFAL*	*FEAR*	*CDIV*	*LFAL*	*LFETH*	*ELF*	*RFCIA*	*RFEB*
INITIAL	0.062	1																		
EDU	0.167	0.773	1																	
LIFE	-0.266	-0.768	-0.789	1																
FERT	-0.278	-0.726	-0.814	0.747	1															
DEMO	0.219	0.604	0.614	-0.625	-0.695	1														
OPEN	0.077	0.091	0.058	-0.129	0.018	-0.009	1													
GOV	-0.088	-0.193	-0.114	0.085	0.157	-0.152	0.103	1												
INV	0.307	0.483	0.450	-0.560	-0.526	0.440	0.182	-0.063	1											
INFL	-0.153	-0.036	-0.024	0.024	0.048	-0.007	-0.083	0.172	0.071	1										
EFCIA	-0.116	-0.289	-0.275	0.390	0.300	-0.141	-0.161	-0.117	-0.385	0.082	1									
EFEB	-0.282	-0.457	-0.483	0.547	0.561	-0.407	-0.031	0.080	-0.464	0.057	0.698	1								
EFAL	-0.296	-0.517	-0.499	0.576	0.537	-0.384	-0.106	0.060	-0.471	0.060	0.727	0.882	1							
FEAR	-0.319	-0.451	-0.497	0.550	0.537	-0.423	-0.068	0.097	-0.548	0.029	0.596	0.841	0.888	1						
CDIV	-0.262	-0.340	-0.369	0.481	0.387	-0.313	-0.121	-0.119	-0.402	-0.020	0.661	0.726	0.808	0.815	1					
LFAL	-0.206	-0.545	-0.506	0.588	0.473	-0.384	-0.062	0.009	-0.440	-0.104	0.604	0.763	0.754	0.706	0.705	1				
LFETH	-0.186	-0.467	-0.426	0.560	0.503	-0.379	0.020	0.013	-0.464	-0.077	0.606	0.808	0.787	0.762	0.733	0.876	*1*			
ELF	-0.172	-0.534	-0.508	0.585	0.463	-0.342	-0.093	-0.062	-0.421	-0.010	0.688	0.805	0.817	0.762	0.807	0.881	*0.841*	*1*		
RFCIA	-0.047	-0.020	0.014	0.017	-0.078	0.062	-0.042	0.029	-0.129	-0.014	0.130	0.098	0.200	0.223	0.133	0.282	*0.172*	*0.258*	*1*	
RFEB	0.026	0.028	0.036	-0.072	-0.119	0.117	0.072	0.057	-0.032	-0.032	0.005	-0.013	0.040	0.141	0.066	0.151	*0.088*	*0.151*	*0.797*	*1*
RFAL	-0.020	-0.029	-0.024	0.020	-0.048	-0.007	0.051	0.045	-0.165	-0.050	0.131	0.109	0.181	0.261	0.154	0.288	*0.210*	*0.252*	*0.862*	*0.819*

Appendix 3: Dependent variable GDP per capita growth

	1		*2*		*3*		*4*		*5*		*6*		*7*		*8*		*9*		*10*		*11*	
C	0.280	*	0.238	*	0.294	*	0.228	*	0.277	*	0.229	*	0.274	*	0.230	*	0.278	*	0.228	*	0.279	*
	7.835		*7.745*		*7.920*		*7.495*		*7.597*		*7.789*		*7.692*		*7.702*		*7.779*		*7.522*		*7.697*	
INITIAL	-0.037		-0.029	*	-0.040	*	-0.025	*	-0.036	*	-0.003	*	-0.037	*	-0.027	*	-0.037	*	-0.026	*	-0.037	*
	-6.602		*-5.256*		*-6.763*		*-4.608*		*-6.279*		*-5.125*		*-6.549*		*-5.095*		*-6.565*		*-4.949*		*-6.459*	
EDU	-0.024	**	-0.001		-0.024	**	-0.009		-0.024	**	-0.008		-0.022	**	-0.007		-0.023	**	-0.004		-0.023	**
	-2.381		*-0.018*		*-2.320*		*-1.044*		*-2.390*		*-0.997*		*-2.269*		*-0.856*		*-2.368*		*-0.441*		*-2.176*	
LIFE	-4.346	*	-6.092	*	-4.513	*	-5.650	*	-4.267	*	-5.167	*	-3.976	*	-5.361	*	-4.133	*	-5.702	*	-4.283	*
	-4.844		*-7.053*		*-4.732*		*-6.600*		*-4.613*		*-6.144*		*-4.372*		*-6.331*		*-4.545*		*-6.618*		*-4.603*	
FERT	-0.057	*			-0.060	*			-0.057	*			-0.050	*			-0.054	*			-0.056	*
	-3.802				*-3.691*				*-3.590*				*-3.296*				*-3.520*				*-3.624*	
DEMO	-0.027				-0.036				-0.029				-0.024				-0.024				-0.026	
	-1.174				*-1.468*				*-1.244*				*-1.065*				*-1.039*				*-1.103*	
DEMO^2	0.032				0.046				0.034				0.029				0.029				0.030	
	1.244				*1.616*				*1.287*				*1.132*				*1.137*				*1.152*	
OPEN	0.009	**			0.009	***			0.010	**			0.010	**			0.010	**			0.009	***
	1.971				*1.764*				*2.046*				*2.073*				*2.081*				*1.901*	
GOV	-0.034	**			-0.035	***			-0.037	**			-0.035	**			-0.034	**			-0.035	**
	-1.989				*-1.880*				*-2.119*				*-2.081*				*-1.972*				*-2.027*	
INV	0.053	**			0.057	**			0.053	**			0.047	**			0.047	**			0.051	**
	2.502				*2.556*				*2.468*				*2.223*				*2.166*				*2.395*	
INFL	-0.003	*			-0.002	**			-0.003	*			-0.003	*			-0.003	*			-0.003	*
	-4.947				*-2.113*				*-4.829*				*-4.867*				*-4.908*				*-4.813*	
EFCIA			0.003		0.009																	
			0.457		*1.315*																	
EFEB							-0.017	**	-0.002													
							-2.235		*-0.293*													
EFAL											-0.027	*	-0.015	**								
											-3.947		*-2.105*									
FEAR															-0.020	*	-0.010					

								-2.886	*-1.422*		
CDIV										-0.013	-0.005
										-1.608	*-0.654*
DUM80	-0.021 *	0.023 *	-0.023 *	-0.023 *	-0.022 *	-0.022 *	-0.020 *	-0.022 *	-0.021 *	-0.022 *	-0.021 *
	-5.199	*-5.442*	*-5.165*	*-5.595*	*-5.185*	*-5.453*	*-5.043*	*-5.472*	*-5.122*	*-5.433*	*-4.958*
DUM90	-0.021 *	0.023 *	-0.023 *	-0.023 *	-0.022 *	-0.021 *	-0.020 *	-0.021 *	-0.021 *	-0.023 *	-0.021 *
	-4.857	*-5.187*	*-4.746*	*-5.346*	*-4.752*	*-5.121*	*-4.517*	*-5.180*	*-4.701*	*5.450*	*-4.682*
R-squared	0.339	0.207	0.322	0.212	0.342	0.233	0.348	0.215	0.343	0.204	0.337

***	Significant at 1 percent level**
****	Significant at 5 percent level**
*****	Significant at 10 percent level**
T-statistics in italic

Appendix 4: Dependent variable: GPD per capita growth

	12		*13*		*14*		*15*		*16*		*17*		*18*		*19*		*20*		*21*		*22*		*23*	
C	0.231	*	0.280	*	0.245	*	0.289	*	0.234	*	0.275	*	0.218	*	0.261	*	0.234	*	0.284	*	0.233	*	0.284	*
	7.621		*7.831*		*7.788*		*7.796*		*7.735*		*7.605*		*6.767*		*6.810*		*7.603*		*7.726*		*7.688*		*7.787*	
INITIAL	-0.027	*	-0.037	*	-0.028	*	-0.039	*	-0.033	*	-0.038	*	-0.024	*	-0.036	*	-0.027	*	-0.038	*	-0.026	*	-0.038	*
	-5.033		*-6.602*		*-5.220*		*-6.696*		*-5.752*		*-6.421*		*-4.311*		*-5.984*		*-4.983*		*-6.573*		*-4.997*		*-6.624*	
EDU	-0.005		-0.023	**	-0.007		-0.024	**	0.014		-0.014		-0.003		-0.020	***	-0.005		-0.024	**	-0.004		-0.024	**
	-0.533		*-2.365*		*-0.814*		*-2.401*		*1.523*		*-1.392*		*-0.359*		*-1.955*		*-0.529*		*-2.374*		*-0.502*		*-2.394*	
LIFE	-5.831	*	-4.385	*	-6.208	*	-4.410	*	-5.274	*	-3.994	*	-5.623	*	-3.905	*	-6.105	*	-4.340	*	-6.080	*	-4.344	*
	-6.897		*-4.842*		*-6.733*		*-4.428*		*-6.394*		*-4.563*		*-6.486*		*-4.183*		*-7.372*		*-4.806*		*-7.369*		*-4.838*	
FERT			-0.057	*			-0.055	*			-0.061	*			-0.052	*			-0.058	*			-0.059	*
			-3.810				*-3.624*				*-3.956*				*-3.291*				*-3.838*				*-3.849*	
DEMO			-0.028				-0.027				-0.017				-0.015				-0.027				-0.027	
			-1.203				*-1.175*				*-0.735*				*-0.602*				*-1.183*				*-1.197*	
DEMO^2			0.033				0.032				0.019				0.022				0.033				0.033	
			1.272				*1.233*				*0.721*				*0.793*				*1.262*				*1.265*	
OPEN			0.009	***			0.009	***			0.008	***			0.008				0.010	**			0.010	**
			1.960				*1.910*				*1.834*				*1.522*				*1.996*				*2.039*	
GOV			-0.034	**			-0.036	**			-0.018				-0.031	***			-0.034	***			-0.034	**
			-1.976				*-2.082*				*-1.089*				*-1.767*				*-1.961*				*-1.998*	
INV			0.054	**			0.049	**			0.042	**			0.055	**			0.052	**			0.050	**
			2.522				*2.288*				*1.984*				*2.477*				*2.394*				*2.326*	
INFL			-0.003	*			-0.003	*			-0.003	**			-0.003	*			-0.003	*			-0.003	*
			-4.949				*-4.939*				*-5.280*				*-4.877*				*-4.915*				*-4.925*	
LFETH	-0.006		0.002																					
	-1.186		*0.344*																					
LFAL					-0.009		-0.008																	
					-1.335		*-1.177*																	
ELF									-0.009		-0.003													
									-1.391		*-0.468*													
RFCIA													-0.007		-0.008									
													-0.933		*-1.177*									
RFEB																	-0.003		-0.005					
																	0.397		*-0.674*					
RFAL																					-0.003		-0.004	
																					-0.041		*-0.621*	

Dum80	-0.023	*	-0.021	*	-0.023	*	-0.021	*	-0.021	*	-0.019	*	-0.023	*	-0.019	*	-0.024	*	-0.021	*	-0.023	*	-0.021	*
	-5.691		*-5.201*		*-5.544*		*-4.951*		*-5.415*		*-4.366*		*-5.363*		*-4.593*		*-5.769*		*-5.215*		*-5.759*		*-5.208*	
Dum90	-0.023	*	-0.021	*	-0.024	*	-0.021	*	-0.020	*	-0.018	*	-0.022	*	-0.019	*	-0.024	*	-0.021	*	-0.024	*	-0.022	*
	-5.467		*-4.862*		*-5.482*		*-4.628*		*-4.844*		*-4.879*		*-5.142*		*-4.047*		*-5.734*		*-4.821*		*-5.790*		*-4.882*	
R-squared	0.200		0.339		0.208		0.342		0.232		0.363		0.186		0.324		0.200		0.339		0.200		0.340	

* **Significant at 1 percent level**
** **Significant at 5 percent level**
*** **Significant at 10 percent level**
T-statistics in italic

Appendix 5: Dependant variable: GPD per Capital growth

	24		*25*		*26*		*27*		*28*		*29*		*30*		*31*	
C	0.239	*	0.296	*	0.230	*	0.278	*	0.240	*	0.283	*	0.226	*	0.281	*
	7.850		*8.097*		*7.520*		*7.723*		*7.645*		*7.662*		*7.315*		*7.616*	
INITIAL	-0.029	*	-0.041	*	-0.026	*	-0.037	*	-0.029	*	-0.039	*	-0.025	*	-0.037	*
	-5.316		*-6.927*		*-4.813*		*-6.435*		*-5.028*		*-6.574*		*-4.582*		*-6.326*	
EDU	-0.002		-0.022	**	-0.005		-0.024	**	-0.002		-0.024	**	-0.007		-0.024	**
	-0.225		*-2.238*		*-0.582*		*-2.448*		*-0.236*		*-2.255*		*-0.811*		*-2.411*	
LIFE	-5.983	*	-4.433	*	-6.120	*	-4.385	*	-6.301	*	-4.426	*	-5.771	*	-4.360	*
	-7.326		*-4.905*		*-7.282*		*4.813*		*-7.207*		*-4.661*		*-6.638*		*-4.667*	
FERT			-0.056	*			-0.058	*			-0.060	*			-0.059	*
			-3.677				*-3.855*				*-3.722*				*-3.805*	
DEMO			-0.044	***			-0.029				-0.025				-0.030	
			-1.857				*-1.239*				*-1.015*				*-1.278*	
DEMO^2			0.053	**			0.034				0.033				0.035	
			1.969				*1.311*				*1.190*				*1.328*	
OPEN			0.009	***			0.010	**			0.008	***			0.010	**
			1.930				*2.115*				*1.684*				*2.037*	
GOV			-0.037	**			-0.037	**			-0.028				-0.037	**
			-2.132				*-2.154*				*-1.557*				*-2.095*	
INV			0.053	**			0.053	**			0.064	*			0.054	**
			2.463				*2.506*				*2.855*				*2.492*	
INFL			-0.002	*			-0.003	*			-0.003	*			-0.003	*
			-2.106				*-4.877*				*-4.734*				*-4.839*	
EGCIA	0.000		0.000													
	-0.046		*-0.053*													
EGEB					0.000		0.000									
					0.340		*0.696*									
EICIA									0.000		0.002					
									0.363		*1.489*					
EIEB													-0.001		0.000	
													-1.385		*0.252*	
Dum80	-0.024	*	-0.023	*	-0.024	*	-0.022	*	-0.023	*	-0.022	*	-0.024	*	-0.022	*
	-5.854		*-5.525*		*-5.806*		*-5.270*		*-5.281*		*-4.952*		*-5.647*		*-5.230*	
Dum90	-0.023	*	-0.023	*	-0.024	*	-4.912	*	-0.024	*	-0.022	*	-0.023	*	-0.022	*
	-5.682		*5.181*		*-5.794*		*-0.004*		*-5.303*		*-4.543*		*-5.484*		*-4.830*	
R-squared	0.209		0.322		0.201		0.344		0.207		0.355		0.205		0.342	

*** Significant at 1 percent level**
**** Significant at 5 percent level**
***** Significant at 10 percent level**
T-statistics in italic

Appendix 6: Dependant variable: GPD per Capital growth

	32		*33*		*34*		*35*		*36*		*37*		*38*		*39*	
C	0.215	*	0.262	*	0.227	*	0.271	*	0.217	*	0.259	*	0.234	*	0.283	*
	6.789		*6.902*		*7.475*		*7.479*		*6.757*		*6.820*		*7.642*		*7.786*	
INITIAL	-0.024	*	-0.035	*	-0.027	*	-0.037	*	-0.024	*	-0.035	*	-0.027	*	-0.038	*
	-4.413		*-6.041*		*-5.040*		*-6.582*		*-4.252*		*-5.970*		*-4.991*		*-6.588*	
EDU	-0.005		-0.022	**	-0.006		-0.025	**	-0.004		-0.020	**	-0.005		-0.023	**
	-0.557		*-2.176*		*-0.690*		*-2.488*		*-0.416*		*-1.973*		*-0.532*		*-2.354*	
LIFE	-5.729	*	-4.109	*	-6.024	*	-4.223	*	-5.662	*	-3.938	*	-6.110	*	-4.350	*
	-6.761		*-4.480*		*-7.272*		*-4.675*		*-6.543*		*-4.224*		*-7.376*		*-4.816*	
FERT			-0.052	*			-0.057	*			-0.050	*			-0.058	*
			-3.377				*-3.771*				*-3.194*				*-3.818*	
DEMO			-0.022				-0.026				-0.015				-0.027	
			-0.938				*-1.140*				*-0.610*				*-1.179*	
DEMO^2			0.029				0.030				0.023				0.033	
			1.112				*1.147*				*0.843*				*1.260*	
OPEN			0.010	**			0.010	**			0.008				0.009	***
			2.034				*2.097*				*1.550*				*1.961*	
GOV			-0.033	***			-0.034	**			-0.032	***			-0.034	**
			-1.916				*-2.007*				*-1.795*				*-1.985*	
INV			0.051	**			0.056	*			0.057	*			0.053	**
			2.364				*2.605*				*2.580*				*2.459*	
INFL			-0.003	*			-0.003	*			-0.003	*			-0.003	*
			-4.864				*-4.907*				*-4.888*				*-4.949*	
RGCIA	0.000		0.000													
	0.727		*0.118*													
RGEB					0.001		0.001									
					1.182		*1.466*									
RICIA									-0.001		-0.002					
									-0.923		*-1.364*					
RGEB													-0.001		-0.001	
													-0.441		*-0.756*	
Dum80	-0.024	*	-0.021	*	-0.023	*	-0.021	*	-0.023	*	-0.019	*	-0.024	*	-0.021	*
	-5.731		*-5.152*		*-5.697*		*-5.132*		*-5.378*		*-4.587*		*-5.771*		*-5.205*	
Dum90	-0.024	*	-0.021	*	-0.024	*	-0.021	*	-0.023	*	-0.019	*	-0.024	*	-0.021	*
	-5.593		*-4.663*		*-5.653*		*-4.697*		*-5.159*		*-4.039*		*-5.738*		*-4.807*	
R-squared	0.190		0.324		0.203		0.342		0.186		0.325		0.200		0.339	

***** **Significant at 1 percent level**
****** **Significant at 5 percent level**
******* **Significant at 10 percent level**

T-statistics in italic