Chapter 3

Measuring Ethnic Diversity

3.1 Introduction

There is a fast growing literature on ethnicity and its role in the economic development of a country or the incidence of conflicts. To advance the research in this area, current approaches try to improve data sources, to increase its coverage, and to construct indices to better measure its complexity. Because ethnicity is not a clear cut concept it contains various aspects. Therefore, better indices in this regard do not mean more accurate indices but rather those that reflect the different aspects more adequately. Starting with the ethno-linguistic fractionalization index (ELF) by Taylor and Hudson (1972), an index on polarization (Garcia-Montalvo and Reynal-Querol, 2002), the reduction to politically relevant groups (Posner, 2004a) or the role of regional segregation of ethnicity (Alesina and Zhuravskaya, 2011) have been studied more intensively.

All these indices, however, are based on pre-defined groups within a country or principal region. This gives rise to an important problem. All calculations rely on a rather arbitrary definition of groups that do not necessarily share a comparable line of differentiation. Fearon (2003) summarizes the absence of a clear-cut definition of ethnic groups and states, maintaining “that in many cases there is no single right answer to the question ‘What are the ethnic groups in this country?’.” (Fearon, 2003, p. 197). To be less arbitrary, a common differentiator, be it on the grounds of ethnicity, language, religion, or any other characteristic need to exist. So, an assessment of distances between groups “is such an

126Ethnic fractionalization is supposed to negatively affect corruption (Mauro, 1995), economic growth (Alesina et al., 2003; Easterly and Levine, 1997), public goods provision (Alesina et al., 1999), communal participation (Alesina and La Ferrara, 2000), general quality of government (Alesina and Zhuravskaya, 2011; La Porta et al., 1999) and democracy (Akedde, 2010). Collier (1998) initiated a new, and now broad strand of literature exploring ethnicity’s impacts on the incidence, onset or severity of conflicts that was furthered by the introduction of an index of polarization (Garcia-Montalvo and Reynal-Querol, 2003, 2005a, 2008).

127For a broad overview of the literature on conflict, see Blattman and Miguel (2010). A good description of concepts and measures of ethnicity is found in Brown and Langer (2010). A new approach to better study ethnic distribution at the micro-economic level is to geo-reference ethnic groups (Weidmann et al., 2010).

128For a similar line of critique, see Lind (2007).
absolutely fundamental concept in the measurement of dissimilarity that it must play an
esential role in any meaningful theory of diversity or classification” (Weitzman, 1992, p.
365). This, however, requires more detailed information on the groups so that they
show a comparable level of distinction in any of the characteristics. Nearly all authors
treat these attributes equally irrespective of the differences between the groups, i.e., how
big the distance is. This is mainly because data on the different similarity levels are
either hardly available, or quite complex. Thereby, it is obvious that two groups whose
respective members speak two completely different languages, follow different religions
and have different physiognomic attributes, are more distant than two groups that share
similarities in their languages, follow the same religion and have a similar appearance.
This underlines the key difference between the diversity concept and the fragmentation
and polarization indices. For many economic problems, it is not the pure number of
groups that is of interest, but rather how difficult coordination or instrumentalization
between the various groups is. In more diverse countries, agreement on public goods (e.g.,
infrastucture or social security systems) is more difficult (Alesina et al., 1999), the level of
generalized trust lower (Bjørnskov, 2008) and the incidence of conflicts higher (Collier and
Hoeffler, 2002). The main aim of this essay is to fill this gap and to offer an index taking
these aspects into account. The global data set offers the possibility to construct an index
covering the degree of diversity between groups within countries, as well as the cultural or
ethnic (dis)similarity between countries. A measure of cultural affinity which extends the
rather crude measure of genetic distance should affect international trade flows. Assessing
this new multi-faceted index is thus the base to further expand current research on the
implication of ethnicity with a new aspect of cultural distance, i.e., its diversity.

The remainder of this chapter is structured as follows. Section 3.2 briefly summarizes
the current discussion surrounding the conceptual und measurement problems. In section
3.3, the theoretical background of the new similarity parameters is outlined. Section 3.4
introduces the data sources used. Section 3.5 discusses the operationalization of the new
distance adjusted ethno-linguistic fractionalization index (DELF), and compares it with
existing measures. Section 3.6 outlines the resulting new diversity values for a range
of countries. In a second step, a (dis)similarity measure between countries, based on
comparable premises, is set up and discussed. Finally, section 3.7 summarizes the key
findings, concludes and gives an outlook for further research.

129 For a good, yet methodological-technical discussion of the prerequisites to measure diversity, see Bossert
et al. (2003) and Nehring and Puppe (2002). Both rely on the earlier concept developed by Weitzman
130 To be precise, ethnic fragmentation or diversity per se is not the cause of the various (negative)
socio-economic outcomes. However, both settings offer more possibilities to exploit these distinctions.
3.2 Different aspects of ethnicity and its measurement

Alesina et al. (2003) describe ethnicity as a “rather vague and amorphous concept” (Alesina et al., 2003, p. 160) that makes any measurement hard to grasp.\footnote{Brown and Langer (2010) offer a broad summary of the recent discussion surrounding the definitions of ethnicity as well as its measurement problems.} To better operationalize ethnicity, this essay follows Chandra and Wilkinson (2008). According to them, ethnic structure comprises a set of ethnic identities that includes all phenotypical attributes (skin pigmentation or body figure), as well as religion, language and the traditions one was raised in. This is very much in line with Barrett et al. (2001), whose data is used later on in this chapter.\footnote{They include language, ethnic origin, skin pigmentation, race, culture or religion, and nationality as characteristics to describe ethnicity.} Following these authors, ethnicity is defined in this chapter along language, ethno-racial (ethnic origin, skin pigmentation and race) and religious aspects.

Defining the characteristics of ethnicity in detail, which is already more diligent than most papers in this field, is not sufficient for what this essay strives for. Within each of the defining criteria a (dis)similarity level between two distinct groups must be assignable. Information on the degree of (dis)similarity is the crucial starting point in any assessment of diversity (Bossert et al., 2003). Despite the reluctance of many authors to define the characteristics of ethnicity, a more thorough examination of similarity differences has not been discussed at all. Distance between groups neither influenced the decision of how to draw the line between groups, nor the interpretation of the fractionalization found. Taking language groups as an example, one could divide groups based on mere dialects, different languages or even different language families. Depending on the level of similarity between groups, different group setups would then emerge.\footnote{For a discussion on how different levels of aggregation of linguistic fragmentation affect the outcomes in the analysis of ethnic conflicts, see Desmet et al. (2012).} In this case, the amount of common vocabulary would define their distance.

Based on the defined number of ethnic groups, the question of its mathematical operationalization arises.\footnote{Ginsburgh and Weber (2011, Ch. 6) offer a good overview of the different classes of indices used, their historical development and recent applications. Desmet et al. (2009) compare the effect of most of these different indices on the level of redistribution.} The most common measure for ethnicity is its fractionalization, known as the ethno-linguistic fractionalization index (ELF). It is calculated as an Herfindahl-Hirschman concentration index:

\[
ELF = 1 - \sum_{i=1}^{K} p_i^2, \quad i = 1, \ldots, K
\]

where \(K\) is the number of groups \(i\) and \(p_i\) their relative group sizes. Its value moves between zero and one and represents the probability that two randomly selected individuals from a population come from different groups. A higher value thus indicates a more fragmented country, i.e., a country with a higher number of distinct ethnic groups. A value close to
one indicates high fragmentation within countries. After the introduction of the ELF by Taylor and Hudson (1972), based on the data of the Atlas Narodov Mira (Bruk, 1964), several additional indices were developed. The second most prominent of these is the measure of polarization introduced by Garcia-Montalvo and Reynal-Querol (2002).\(^\text{135}\) It shows a completely different aspect of a country’s ethnic setup, and underlines that for each economic problem under analysis, the adequate index needs to be applied. Assessing the variation away from an even 50/50 split of two groups, Garcia-Montalvo and Reynal-Querol (2002) find that this index is a much better predictor of conflict incidence than the ELF measure. It apparently better measures the ethnic constellations responsible for an uprising. The polarization index (POL) is defined as:

\[
POL = 1 - \sum_{i=1}^{K} \left( \frac{0.5 - p_i}{0.5} \right)^2 \cdot p_i, \quad i = 1, \ldots, K
\]  

\(^{135}\) Their approach goes back to earlier work of Esteban and Ray (1994).

\(p_i\) are again the relative group sizes of groups \(i\). The POL index is also tending towards zero for very homogeneous countries, i.e., with only one group. However, with increasing group numbers, ELF and POL show clearly different courses. Figure 3.1 shows these differences based on equally sized groups. While ELF is an increasing function of the number of groups, POL reaches its maximum with two equally sized groups and decreases afterwards. This clearly underlines that the indices do in fact measure two different things although they are based on the same data.

**Figure 3.1:** ELF and POL values depending on the number of equally sized groups

Bossert et al. (2011) introduce a more flexible version of the ELF, the generalized ethno-linguistic fractionalization index (GELF). The technical side of the index brings two important improvements. Firstly, it does not rely on pre-defined groups but takes the individual
and its specific characteristics as a starting point. Based on the specific characteristics, a mutual similarity matrix between individuals takes the distance between them into account. Hereby the groups emerge ‘endogenously’ from the matrix. The similarity value between two individuals \( i \) and \( j \) for all \( i, j \in \{1, ..., N\} \) is given through \( s_{ij} \), with:

\[
1 \geq s_{ij} \geq 0 \tag{3.3}
\]

\[
s_{ii} = 1 \tag{3.4}
\]

\[
s_{ij} = s_{ji} \tag{3.5}
\]

A similarity value of one indicates perfect similarity, whereas a value of zero would indicate two individuals that do not share any characteristics. For a society with \( N \) individuals, all \( \{s_{ij}\} \) are contained in a \( N \times N \) matrix, labeled similarity matrix \( S_N \), which is the main building block of the GELF. Based on this matrix, the corresponding GELF value for a country with \( N \) individuals is given through:

\[
G(S_N) = 1 - \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} s_{ij} \tag{3.6}
\]

GELF is then the expected dissimilarity between two randomly drawn individuals. As data on individuals are seldom available, the transfer to group-specific data on the smallest aggregation level is needed. The adaptations are, however, rather small. In a society with \( N \) individuals, \( K \) groups exist with respective populations of \( m_k \) individuals for all \( k \in \{1, ..., K\} \). It holds that \( \sum_{k=1}^{K} m_k = N \) and \( p_k = m_k/N \) is the respective relative group size. The individuals in each group are all perfectly similar, i.e., their mutual individual similarity values would be one. By grouping all individuals together that share similarity values of one, groups emerge ‘endogenously’. The similarity between two groups, \( k \) and \( l \), is denoted as \( \hat{s}_{kl} \) and is equivalent to the individual similarity value \( s_{ij} \) for any \( i \in m_k \) and \( j \in m_l \). In rearranging Equation (3.6), it follows that:

\[
G(S_n) = 1 - \frac{1}{N^2} \sum_{k=1}^{K} \sum_{l=1}^{K} m_k m_l \hat{s}_{kl}
\]

\[
= 1 - \frac{K}{N} \sum_{k=1}^{K} m_k \frac{m_l}{N} \hat{s}_{kl}
\]

\[
= 1 - \sum_{k=1}^{K} \sum_{l=1}^{K} p_k p_l \hat{s}_{kl} = DELF \tag{3.7}
\]

The relation between the \( DELF \) and the ELF index is quite obvious. The ELF is based on groups that either have a similarity value of one, given both belong to the identical group, and zero otherwise. Thus, the products are always zero if two different groups

\[\text{Philipp Kolo - 9783653023954} \]
\[\text{Downloaded from PubFactory at 12/22/2018 04:23:20AM} \]
\[\text{via free access} \]
are matched. A value of one is only assigned if the groups are matched with themselves, leading to a value of \((p_k \cdot p_k \cdot 1) = p_k^2\) and \((p_k \cdot p_l \cdot 0) = 0\), respectively. The sum over all \(K\) groups then directly leads to Equation (3.1), where the ELF is specified.\(^{137}\) The important improvement in this approach is that it does not rely on pre-defined groups, thus avoiding to treat groups as equal that actually have very large distances between them.\(^{138}\)

Finally, de Groot (2009) assessed the ethnic affinity between African nations.\(^{139}\) In doing so, he also draws on the articles of Fearon (2003) and an earlier version of Bossert et al. (2011), and is closest to the approach of this essay. De Groot (2009), however, only offers data on ethnic affinity between countries and limits his assessment to Africa. This essay consequently extends the work of all three studies.

### 3.3 Calculation of the distance values

For the calculation of the distance values, this essay draws on Fearon (2003). His approach is adapted for three ethnicity characteristics: language, ethno-racial and religious identification. Taking a broader set of characteristics and similarity measures into account offers a more multifaceted picture.\(^{140}\)

#### 3.3.1 Language classification

Language is probably the most researched and operationalized characteristic.\(^{141}\) As is the case with a family tree, languages can be ordered in accordance with their mutual relatedness. The distance between the branches gives a measure of their degree of (dis)similarity. This is well analyzed and operationalized by the *Ethnologue* project (Lewis, 2009). To uniquely identify each language, it assigns each one with a three letter code. The de-

\(^{137}\)Note that due to the construction of Equation (3.7), DELF values take into account mutual similarity values between groups that are not fully identical and will therefore always be lower than the ELF values. The DELF delivers the same result as a monolingual weighted index proposed by Greenberg (1956) and used by Fearon (2003) in his calculation of ‘cultural fractionalization’. Further attributes of the new index and its relation to the other indices (ELF and POL) are discussed in Garcia-Montalvo and Reynal-Querol (2005a, 2008) and Esteban and Ray (2011). In the latter, the index is labeled as the ‘Greenberg-Gini’ index.

\(^{138}\)The superior theoretical explanatory power of such an index is also discussed in Ginsburgh and Weber (2011).

\(^{139}\)The ethnic linguistic affinity (ELA) of de Groot (2009) measures, in contrast to the ELF, the amount of characteristics shared between two countries and thus follows an inverse logic. Because it is the most widely propagated, this essay follows the logic of the ELF, where higher values denote more fragmented countries.

\(^{140}\)Ginsburgh (2005) and Ginsburgh and Weber (2011, Ch. 3) offer an introduction into alternative methods to assess the distances between groups, especially genetic and cultural distances. Genetic distance can be traced back to Cavalli-Sforza and Feldmann (1981). In contrast, Hofstede (2000) assesses differences between cultures and nations along four dimensions: power distance, individualism, masculinity and uncertainty avoidance. Comparable, but slightly different approaches, use answers from the *World Value Survey* (Desmet et al., 2011) or the voting behavior in the Eurovision Song Contest (Felbermayr and Toubal, 2010) to construct cultural differences between nations.

\(^{141}\)Ginsburgh and Weber (2011, Ch. 3) offer a good overview of the different approaches to assess the distances between languages.
cision and categorization as a separate language (instead of a dialect) not only follows pure linguistic and lexical similarities, but also considers how a mutual understanding in communication is possible.

This essay relies on a very closely related approach used in the World Christian Encyclopedia (Barrett et al., 2001). A wide congruency of both sources exists, as the World Christian Encyclopedia (henceforth WCE) is one of the sources for the Ethnologue data. Here, a seven character code is assigned to each distinct language. A distinct language is defined as “the mother tongue of a distinct, uniform speech community with its own identity” (Barrett et al., 2001, V.II, p. 245). It comprises all dialects that share at least 85% of their vocabulary and grammar to ensure adequate communication.\textsuperscript{142} In total, 6,656 distinct languages are contained in the data analyzed. Two persons speaking one language are treated as completely similar ($s_{ij} = 1$).\textsuperscript{143} The more characters of the assigned code two languages share, the more similar they are. The structure is depicted in Table 3.1.

<table>
<thead>
<tr>
<th>Glossocode</th>
<th>Description</th>
<th>Minimal similarity level</th>
<th>Number of distinct groups</th>
<th>$\bar{s}_{kl}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Macrozone</td>
<td>0%</td>
<td>10</td>
<td>0.01</td>
</tr>
<tr>
<td>01</td>
<td>Glosso-zone</td>
<td>5%</td>
<td>100</td>
<td>0.06</td>
</tr>
<tr>
<td>01-A</td>
<td>Glosso-set</td>
<td>30%</td>
<td>594</td>
<td>0.35</td>
</tr>
<tr>
<td>01-AA</td>
<td>Glosso-chain</td>
<td>50%</td>
<td>1,213</td>
<td>0.59</td>
</tr>
<tr>
<td>01-AAA</td>
<td>Glosso-net</td>
<td>70%</td>
<td>2,388</td>
<td>0.82</td>
</tr>
<tr>
<td>01-AAAA</td>
<td>Glosso-cluster</td>
<td>80%</td>
<td>4,241</td>
<td>0.94</td>
</tr>
<tr>
<td>01-AAAA-a</td>
<td>Language</td>
<td>85%</td>
<td>6,656</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 3.1: Language similarity classification according to Barrett et al. (2001)

The Afghan Persian (58-AACC-b) and Southern Pathan (58-ABDA-b) group share the first three digits and thus belong to one Glosso-set, sharing between 30% and 50% of their vocabulary and grammar. Subsequently, both groups are assigned a similarity value $\bar{s}_{kl}$. The assigned values are normalized on a scale between zero and one, and are matched to demonstrate the same decreasing slope as the lexical similarity levels. Belonging to one language group and thus sharing 85% lexical similarity corresponds to the highest $\bar{s}_{kl}$ with $\bar{s}_{kl} = 1$.\textsuperscript{144} In the case of the example $\bar{s}_{kl}$ takes a value of 0.35.

### 3.3.2 Ethno-racial distance

Fragmentation that is derived from a biological taxonomy of species is mainly based on genealogical relatedness between different people in modern humanity. The long evolu-

\textsuperscript{142}The same threshold is used by the Ethnologue project (Lewis, 2009), which is one of the main sources for the assignment of language similarity levels. The second source is Dalby and Williams (1999). The data and classification can also be found online under: http://www.linguasphere.info.

\textsuperscript{143}For a different way taking language differences into account, see Desmet et al. (2012). Depending on the similarity level defined (e.g., dialects vs. languages), different numbers of groups and thus different levels of fragmentation, eventually emerge. This follows on from the discussion in the introduction that the (arbitrary) group definition significantly impacts ELF levels.

\textsuperscript{144}For a discussion on alternative similarity values, see Appendix C.1.2.
tionary process is described by Ahlerup and Olsson (2007) as ‘genetic drift’. This means that the human species developed quite differently in various parts of the world, with one being able to map a genealogical tree based on the genetic congruence of the resulting races. Cavalli-Sforza and Feldmann (1981) created these phyloographic trees by mapping the differences in special sections of the human DNA. Cavalli-Sforza et al. (1993) assessed dyadic distances between 42 world populations computed from 120 alleles in the human genome.\textsuperscript{145}

This was certainly a pioneering piece of work but also demonstrates some limitations. The first one is the small number of groups (42) for the global classification. For Europe, Spolaore and Wacziarg (2009) only refer to four different genetic groups in their analysis of innovation and development diffusion across countries.\textsuperscript{146} It is quite obvious that this might not be sufficient to describe the diversity of Europe. The second caveat is brought forward by Giuliano et al. (2006), who discuss in detail the use of genetic distance data and conclude that it is a proxy for geographical distances, rather than a proxy for cultural distances.\textsuperscript{147} The genes used to assess the genetic distance in Cavalli-Sforza et al. (1993) are only in a very limited way responsible for the phenotypical or anthropometric differences. The part of the DNA used is located on neutral points only subject to random drift, and less to evolutionary selection.\textsuperscript{148} However, to assess the distance between two human beings, with respect to their ease or willingness to cooperate, phenotypical or anthropometric markers should be relevant.\textsuperscript{149}

In order to combine these views and caveats, this essay follows an ethno-racial taxonomy outlined by Barrett et al. (2001). Each unique group is assigned a six character code based on differences of race, skin pigmentation and ethnic origin.\textsuperscript{150} Although those characteristics are closely linked in their development, their role for mutual understanding differs and is treated as cumulative in the subsequent analysis.\textsuperscript{151}

\textsuperscript{145}Due to the special location of the DNA compared, differences are caused only by a constant random drift. This allows one to calculate when two populations split up genetically during the course of the peopling of the world.

\textsuperscript{146}For Europe, a more precise split of genetically different groups is available, but it is not possible to combine this with the global structures, because these data are based on a different set of genes. Ashraf and Galor (2011) use an extended version of genetic distance data covering 53 ethnic groups and their mutual heterozygosity based on Ramachandran et al. (2005).

\textsuperscript{147}Ramachandran et al. (2005) confirm this hypothesis in an analysis of their extended set of 53 populations. They show that correlation values between different measures of genetic distance and the geographical distance from Ethiopia is at least 0.76.

\textsuperscript{148}However, evolutionary selection is strongly driven by the appearance of species (e.g., mating) or their better adaptability to the surroundings; that is mainly due to differences in their physical shape.

\textsuperscript{149}Caselli and Coleman (2008), for example, attribute the emergence of the conflict in Rwanda to the possible distinction between Hutus and Tutsis according to their body sizes.

\textsuperscript{150}This also includes some major similarities between languages to define distinct cultural groups, which is due to the very closely linked development of genetical and language evolution (Cavalli-Sforza et al., 1988).

\textsuperscript{151}This approach is also followed by de Groot (2009).
Analogous to the pure language case, the different levels of ethno-racial classification are summarized in Table 3.2. The broadest classification is along racial lines, with five different races existing. The next level adds a geographical marker (e.g., African or European) to the race distinction. The major culture area adds an additional physiological characteristic, mainly driven by skin pigmentation. The first three characters of the code are thus driven by phenotypical differences. Local races are characterized as a “culture area, local breeding population/reproductive isolate and genetically distinct population” (Barrett et al., 2001, V.II, p. 19). To differentiate between larger ethno-racial families and to characterize distinct ethnic groups or ‘microraces’, a final character is assigned as an identifier. On the global scale, the data contains 393 such ethno-racial families.

<table>
<thead>
<tr>
<th>E-L-Code</th>
<th>Description</th>
<th>Similarity level</th>
<th>Number of distinct groups</th>
<th>$\bar{s}_{kl}^E$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Race</td>
<td>1</td>
<td>5</td>
<td>0.01</td>
</tr>
<tr>
<td>AU</td>
<td>Geographical race</td>
<td>2</td>
<td>13</td>
<td>0.21</td>
</tr>
<tr>
<td>AUG</td>
<td>Major culture area</td>
<td>3</td>
<td>18</td>
<td>0.59</td>
</tr>
<tr>
<td>AUG-03</td>
<td>Local race</td>
<td>4</td>
<td>72</td>
<td>0.88</td>
</tr>
<tr>
<td>AUG-03-b</td>
<td>Ethno-racial family</td>
<td>5</td>
<td>393</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 3.2: Ethno-racial group and similarity classification according to Barrett et al. (2001)

For the ethno-racial classification, Barrett et al. (2001) do not clearly develop a similarity measure, instead measuring the distance on integer values. The different similarity levels ($\bar{s}_{kl}^E$) are calculated with the same decrease in slope of the similarity values being found as that of the language characteristic.

Taking the same two groups in Afghanistan and comparing their ethno-racial classification, allows one to derive their similarity value of this characteristic. Accordingly, the Persians (CNT-24-f) and Southern Pathans (CNT-24-a) belong to one ethno-racial family and are eventually assigned a mutual similarity value $\bar{s}_{kl}^E$ of 0.88.

### 3.3.3 Religious classification

Religion is undoubtedly a major factor in shaping cultural habits and practices. The existence of different religions is often seen as an important reason for conflicts or general misunderstandings between different groups. Religious identification is in a certain way, an especially potent, but easily implemented instrument to expand ones political

152 Whenever it is not the unique contribution of Barrett et al. (2001), the ethno-racial classification closely follows the *Encyclopædia Britannica.*

153 Barrett et al. (2001) caution that these racial classification only act as a mere indicator as there “exist almost imperceptible gradations of genetic character from one group of people to the next” (Barrett et al., 2001, V.II, p. 15). In general, this allows for mixtures between the outlined races.

154 Therefore the values of $\bar{s}_{kl}^E$ clearly differ, because only five levels are assigned for the ethno-racial classification, instead of seven, as is the case for language.

155 See, for example, Garcia-Montalvo and Reynal-Querol (2003) for the increased incidence of conflicts and de Groot (2009) for its spillover effects between neighboring states. For a more general discussion on the effect of religious beliefs on economic growth, see Barro and McCleary (2003).
power through mobilizing one’s followers. Religious inspiration may then be used to trade loyal following in this life, for rewards in an afterlife. The commonalty of religion, however, can also be a major driver of trust, enhancing trade between nations with the same denomination (Guiso et al., 2009). This underlines the importance of this specific characteristic in assessing the differences between groups.

The major problem with religion is the assessment of their differences. How to treat the differences between different denominations, i.e., between Catholics and Protestants, or between Shias and Sunnis, is quite hard to answer. One could try to pursue the same method as that of language and race to assess mutual commonalities. For religion, one could rely on shared festivities, common holy books, common saints/prophets, traditions or values (e.g., mercy). However, there is no known source offering a discussion of this, let alone a structured assessment of the religions of the world. The WCE lists 14 major religions in the data: Agnostics, Buddhists, Chinese folk-religionists, Christians, Confucianists, Daoists, Ethnoreligionists, Hindus, Jews, Muslims, New religionists, Sikhs, Spiritists and Zoroastrians. This essay follows the approach that Bossert et al. (2011) applied in their study. For their partition along ethnic lines, they apply a purely categorical assessment, i.e., the mutual similarity values are either one or zero. This approach should be adjusted as better data become available.

### 3.3.4 Other socioeconomic aspects

An interesting idea championed by Bossert et al. (2011) is that for the distance people feel between each other, not only does their ethnicity play a role, but also their similarities in other dimensions. Bossert et al. (2011) use educational and income similarities in addition to ethnic diversity, arguing that these variables are relevant for a ‘felt’ distance between individuals or groups. Bossert et al. (2011) conclude that in states where one finds economic homogeneity, ethnic diversity might be less important than in economically more heterogeneous states, where both show comparable levels of ELF.

As for this essay, one faces two problems. Most socioeconomic variables are not available to the same level of granularity as the data used here, and data might not be matched to the ethnic groups. The more serious problem is that most economic literature finds a significant impact of ethnicity on various socioeconomic variables. Additionally, in many countries, the wealth or education stratification is closely linked to ethnic descent. Thus, with a high certainty there exists endogeneity of these socioeconomic variables with regard to ethnicity. As this cannot be ruled out – and there is no adequate data to match the level of detail for ethnicity employed hereafter – further analysis into this aspect is not pursued.

---

156 Guiso et al. (2009) use the same approach but with a slightly smaller amount of denominations.

157 In this regard, Bjørnskov (2008) points toward social trust and income inequalities. Another interesting approach for the US is that of Lind (2007). He tries to assess the inter-group distance through measuring differences in stated preferences on policy questions.

158 The same might be true for religion and languages, or even dialects.
3.4 Data description and comparison with other sources

There are various sources for religious, ethnic and language data that are widely used in the literature. Besides the wide range of ethno-linguistic groups in the *Atlas Narodov Mira* (Bruk, 1964), Alesina et al. (2003) mainly use data from the *Encyclopædia Britannica* (Encyclopædia Britannica, 2007) and from the *CIA World Fact Book* (CIA, 2011) for their data on ethnicity. For languages, the *Ethnologue* project (Lewis, 2009) offers very detailed data of nearly 7,000 languages. Finally, *L’Etat des Religions dans le Monde* (Clévenot, 1987) offers very exhaustive data on religious affiliation for a wide range of countries.\(^{159}\)

All these sources have their advantages and are certainly applicable for the intention of the respective authors. They, however, lack an important aspect, which is relevant for the analysis here. To build the similarity matrix based on all three traits (language, ethno-racial, religion), each group needs to be defined in accordance with all three of them. This is not possible with the above sources as the groups found in the sources vary depending on the defining criteria.

The source offering the required data is the *World Christian Encyclopaedia* (Barrett et al., 2001).\(^{160}\) It contains data for over 12,000 groups in 210 countries, classified according to language, ethno-racial group and religion.\(^{161}\) The data are based on various sources including official reports, national censuses, statistical questionnaires, field surveys and interviews, as well as several other published and unpublished sources. The level of detail and the vast coverage of countries is a strong advantage of this source. The data on languages and ethno-racial affiliation are widely used.\(^{162}\) Due to the Christian background of the publishing institutions, one could argue (at least for the data on religion), that the numbers might be biased. Their very detailed assessment of Christian denomination, however, is an indication of a real interest to survey Christianity, drawing an unbiased picture of their faith.\(^{163}\) The high granularity of data might still raise some questions.

---

\(^{159}\)Akdede (2010) gives a good overview of the data sources used in a broad set of influential articles and discusses their differences.

\(^{160}\)For all calculations the online version, *The World Christian Database* (Johnson, 2010), is used. It reflects the data in the printed version of Barrett et al. (2001) but includes significant updates and refers to the 2005 – 2010 time period.

\(^{161}\)In total, over 13,500 groups for 239 countries are included in the data. Groups that differ only through dialects or, in some cases, geographical specifics, like, for example, the Bedouin tribes in Algeria, were excluded. Additionally, very small islands and constituencies with an unclear legal status (e.g., Western Sahara) were excluded.


\(^{163}\)Additionally, Barrett et al. (2001) explicitly mention the United Nations’ Universal Declaration of Human Rights in their preface, which grants the freedom to choose one’s religion, including not having a religion at all. De Groot (2009) uses a similar, unorthodox evangelical source, the Joshua Project (2007). He also concludes that the “religious fervency with which this organization collects data works in our advantage” (de Groot, 2009, p. 14). Collier and Hoefller (2002, 2004) and Collier et al. (2004) used it for their index on religious fractionalization. However, Garcia-Montalvo and Reynal-Querol (2005a) discuss some bias towards Christianity at the expense of Animist cults in Latin American countries. Although there is no evidence of a general bias in religious affiliation, it can’t be ruled out completely.
### Table 3.3: Descriptive statistics of ethnic groups by geographical area

<table>
<thead>
<tr>
<th></th>
<th>World</th>
<th>Western Countries&lt;sup&gt;a&lt;/sup&gt;</th>
<th>MENA</th>
<th>Latin America&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Asia&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Eastern Europe</th>
<th>SSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of countries&lt;sup&gt;d&lt;/sup&gt;</td>
<td>210</td>
<td>33</td>
<td>21</td>
<td>38</td>
<td>40</td>
<td>29</td>
<td>49</td>
</tr>
<tr>
<td>Fraction of total</td>
<td>16%</td>
<td>10%</td>
<td>18%</td>
<td>19%</td>
<td>14%</td>
<td>23%</td>
<td></td>
</tr>
<tr>
<td>Number of groups</td>
<td>12,432</td>
<td>1,716</td>
<td>625</td>
<td>1,405</td>
<td>4,143</td>
<td>1,019</td>
<td>3,524</td>
</tr>
<tr>
<td>Fraction of total</td>
<td>14%</td>
<td>5%</td>
<td>11%</td>
<td>33%</td>
<td>8%</td>
<td>28%</td>
<td></td>
</tr>
<tr>
<td>Average groups per country</td>
<td>59</td>
<td>52</td>
<td>30</td>
<td>37</td>
<td>104</td>
<td>35</td>
<td>72</td>
</tr>
<tr>
<td>Max. number of groups</td>
<td>884</td>
<td>300</td>
<td>71</td>
<td>255</td>
<td>884</td>
<td>156</td>
<td>513</td>
</tr>
<tr>
<td>Min. number of groups</td>
<td>3</td>
<td>3</td>
<td>14</td>
<td>6</td>
<td>3</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>Average pop. share of largest group</td>
<td>57%</td>
<td>68%</td>
<td>60%</td>
<td>64%</td>
<td>52%</td>
<td>75%</td>
<td>39%</td>
</tr>
<tr>
<td>Number of countries with a group ≥ 50%</td>
<td>123</td>
<td>25</td>
<td>14</td>
<td>27</td>
<td>19</td>
<td>26</td>
<td>12</td>
</tr>
<tr>
<td>Fraction of all countries</td>
<td>59%</td>
<td>76%</td>
<td>67%</td>
<td>71%</td>
<td>48%</td>
<td>90%</td>
<td>24%</td>
</tr>
</tbody>
</table>

<sup>a</sup>Western Europe and Australia, Canada, Greenland, Japan, New Zealand and United States.

<sup>b</sup>Includes the Caribbean.

<sup>c</sup>Includes the Pacific islands.

<sup>d</sup>In total data for 239 countries and constituencies are provided. Data on small islands and legally unclear constituencies were excluded: Anguilla, Bougainville, British Indian Ocean, British Virgin Islands, Christmas Island, Cocos (Keeling) Islands, Cook Islands, Falkland Islands, French Guiana, Gibraltar, Guadeloupe, Holy See, Martinique, Montserrat, Niue, Norfolk Island, Northern Cyprus, Pitcairn Islands, Reunion, Western Sahara, Saint Helena, Saint Pierre & Miquelon, Somaliland, Spanish North Africa, Svalbard & Jan Mayen, Taiwan, Tokelau Islands, Turks & Caicos Islands, Wallis & Futuna Islands.
about its accuracy. To test the robustness of the base data, two additional data sets with some noise based on a normal randomization are created. Additionally, the consistency of the data was tested if very small groups in the data were excluded. For both robustness checks, no significant deviation from the results employing the base data set occur.\textsuperscript{164}

Below, the most granular group data is used to offer the best possibility of endogenous group formation. Although data at the individual level is not available, this very granular data is very close to the desired approach outlined earlier. Table 3.3 gives an overview of the data, which is structured according to Alesina et al. (2003) and Fearon (2003).

The WCE data clearly show much more groups. Alesina et al. (2003) have, on average, less than six groups per country. While 59 groups are counted in the present data set, on average. Besides the higher number of groups in general, the pattern of fractionalization across the regions is quite similar, with one exception. In contrast to the previous sources, this data show that most groups are located in Asia.\textsuperscript{165} This is nearly exclusively driven by three countries that contribute half of all groups in this region: Papua New Guinea with 884 groups, Indonesia 762 and India 428.\textsuperscript{166} Excluding these three countries, Sub-Saharan Africa is again the region with the most fragmented countries.\textsuperscript{167} This becomes even clearer when one compares the other figures in Table 3.3. The average population share of the largest group is only 39\% of the population’s total in Sub-Saharan Africa, whereas it is at least 50\% in all other regions. Also, the number of countries that have a majority group of 50\% is significantly lower.

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|c|c|c|}
\hline
Source & Obs. & Mean & Std. Dev. & Min. & Max. \\
\hline
ANM & 169 & 0.458 & 0.273 & 0.000 & 0.984 \\
Alesina & 186 & 0.440 & 0.257 & 0.000 & 0.930 \\
ELF & 144 & 0.479 & 0.275 & 0.010 & 0.950 \\
Annett & 153 & 0.471 & 0.270 & 0.002 & 0.953 \\
Fearon & 210 & 0.563 & 0.270 & 0.019 & 0.982 \\
\hline
\end{tabular}
\caption{Main statistical characteristics of ELF values for different sources}
\end{table}

The higher amount of small groups also has an effect on the ELF values based on the WCE data, reflected in a noticeably higher mean value. A higher number of groups will increase the ELF index by design.\textsuperscript{168} Table 3.4 confirms this by showing the summary statistics of the ELF values for the various sources described earlier.

\textsuperscript{164}For more details on these robustness checks, see Appendix C.1.1.
\textsuperscript{165}The Asian region includes the Pacific countries and islands.
\textsuperscript{166}Although this number seems to be high, it is very much in line with other very detailed sources. Lewis (2009) lists 860 languages for Papua New Guinea, over 10\% of the world’s total in his data set.
\textsuperscript{167}Excluding these three countries, the average number of groups per country in Asia would only amount to 56.
\textsuperscript{168}The theoretical attributes of the ELF and POL are nicely met by the WCE data. Figure C.6 of Appendix C.1.3 shows the increasing ELF values in conjunction with a rising number of groups within a country.
3.5 DELF operationalization

For the construction of the new composite distance adjusted ethno-linguistic fractionalization index (DELF), two major, partly interconnected, questions arise. The first is, whether the single components are redundant when compared to each other. The second is the assignment of weights and the way of combining the single characteristics.

Based on theoretical considerations, no single characteristic out of the three is deemed to be superior or more sound than the others, with all of them seeming to be of equal relevance.\(^{169}\) For the same reason, Okediji (2005) proposes including ethnic differentiation alongside racial and religious characteristics.\(^{170}\) Finally, one can argue that the distance between the groups increases, if more differences are in place, which would be in line with the cumulative statement of de Groot (2009).\(^{171}\)

The most common approach when incorporating different characteristics into a combined index is to assign equal weights to all of its components.\(^{172}\) Following this approach, the DELF is calculated according to Equation (3.7) as:

\[
DELF = 1 - \sum_{k=1}^{K} \sum_{l=1}^{K} p_k \cdot p_l \cdot \hat{s}_{kl}
\]  (3.8)

where the combined \(\hat{s}_{kl}\) is the equally weighted average of the similarity values of each ethnicity characteristic.

\[
\hat{s}_{kl} = \frac{1}{3} \left[ \bar{z}_{kl}^L + \bar{z}_{kl}^E + \bar{z}_{kl}^R \right]
\]  (3.9)

where \(\bar{z}_{kl}^L\), \(\bar{z}_{kl}^E\) and \(\bar{z}_{kl}^R\) are the respective similarity values for the language, ethno-racial and religious classification.\(^{173}\) The single characteristic DELFs are equally calculated

\(^{169}\)See, for example, Chandra and Wilkinson (2008) and Barrett et al. (2001). Hofstede (2000) concludes similarly that “the world population has diversified in three ways: in genes, in languages, and in cultures” (Hofstede, 2000, p. 3).

\(^{170}\)Okediji (2005) constructs his social diversity index based on the complementary nature of the three characteristics and also uses WCE data. However, he does not take into account the mutual (dis)similarities between the groups.

\(^{171}\)One could argue that by design, the language and ethno-racial classification is not without overlaps. This is why one should weight their sum less. On the other hand, the religious classification is less accurate and would, in contrast, argue for a lower weighting of this characteristic. If there is no strong reason for deviating from the equal weighting, Haq (2006) argues strongly for this principle.

\(^{172}\)The most well-known index calculated utilizing this approach is the UNDP’s Human Development Index (HDI). More recent examples are the SIGI index on gender equality (Branisa et al., 2009) or the 3P index on trafficking policies (Cho et al., 2011). For an analysis of different operationalization strategies for a broad set of composite development indicators, see Boysen (2002).

\(^{173}\)The main focus of this essay is to assess the diversity of a country, which is well reflected by the DELF. However, from the discussion above, one can easily apply the similarity values \(\hat{s}_{kl}\) to an adapted version of the polarization index found in Equation (3.2). This would then transform to a distance adjusted POL index with: D-POL = \(\sum_{k=1}^{K} \sum_{l=1}^{K} p_k \cdot p_l \cdot \hat{s}_{kl}\) (Esteban and Ray, 1994). For further theoretical discussions on this kind of index, see Esteban and Ray (2008) and Esteban and Ray (2011). For rare examples of an empirical application of this index, see Desmet et al. (2009), Esteban et al. (2010), Esteban and Ray (2011) and Esteban and Mayoral (2011). The data for the D-POL index based on the WCE data can be obtained from the author upon request.
using Equation (3.9). Instead of the composite similarity measure ($\hat{s}_{kl}$) the characteristics specific similarity values ($\bar{s}_{L}^{L}, \bar{s}_{E}^{E}, \bar{s}_{R}^{R}$) are used. To decide on the redundancy of the composite index and its components, McGillivray and White (1993) propose two thresholds of correlation values between the components: 0.90 and 0.70.\footnote{Cahill (2005), McGillivray and Noorbakhsh (2004), Branisa et al. (2009) and Cho et al. (2011) subsequently used this decision rule.} The Spearman’s rank correlations of the DELF values based on the components (labeled with a respective subscript for (L)anguage, (E)thno-culture and (R)eligion) and the composite DELF index are shown in Table 3.5.\footnote{Because all conditions are fulfilled, Pearson’s correlation coefficients can also be used. The results are comparable throughout, but slightly lower. As, in the following, the focus is mainly on ranking comparison, Spearman’s rank correlations are consequently used.} 

<table>
<thead>
<tr>
<th>DELF</th>
<th>DELF$_{L}$</th>
<th>DELF$_{E}$</th>
<th>DELF$_{R}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DELF</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DELF$_{L}$</td>
<td>0.904</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>DELF$_{E}$</td>
<td>0.714</td>
<td>0.537</td>
<td>1</td>
</tr>
<tr>
<td>DELF$_{R}$</td>
<td>0.665</td>
<td>0.529</td>
<td>0.195</td>
</tr>
</tbody>
</table>

\textbf{Table 3.5:} Rank correlation for the composite DELF and its components

The correlations between the single components are no higher than 0.54, falling clearly below both thresholds. Thus, any form of double counting by using collinear indicators can be neglected. As the composite index is partly matched to its components, the resulting correlations are naturally higher. By correlating the components with reduced forms of the DELF (by excluding the respective component), most correlations again fall below both thresholds (McGillivray and White, 1993; Ogwang and Abdou, 2003).\footnote{The correlation between DELF$_{L}$ and the reduced DELF by excluding DELF$_{L}$ shows a value of 0.69. The respective values for excluding DELF$_{E}$ and DELF$_{R}$ are 0.48 and 0.43, all falling below both thresholds.} In addition to the overall correlations, Noorbakhsh (1998) proposes to split the total observations into different groups. A high correlation overall might hide differences within groups, e.g., split into quintiles. Table 3.6 shows the correlations seen in Table 3.5, split between equally sized quintiles.

<table>
<thead>
<tr>
<th>Quintiles</th>
<th>All obs.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>DELF</td>
<td>DELF$_{L}$</td>
<td>0.904*</td>
<td>0.282</td>
<td>0.483*</td>
<td>0.401*</td>
<td>0.556*</td>
</tr>
<tr>
<td>DELF$_{E}$</td>
<td>0.714*</td>
<td>0.056</td>
<td>0.156</td>
<td>0.050</td>
<td>0.141</td>
<td>0.815*</td>
</tr>
<tr>
<td>DELF$_{R}$</td>
<td>0.665*</td>
<td>0.569*</td>
<td>0.142</td>
<td>0.004</td>
<td>0.276</td>
<td>0.372*</td>
</tr>
</tbody>
</table>

* indicate rank correlations that are significant at the 5% level

\textbf{Table 3.6:} Rank correlation for equally sized quintiles (according to their DELF values)

Indeed this shows that the higher correlations between the components and the composite DELF vanish completely, or are at least far below both thresholds, except for the
fifth quintile. In light of the above discussion, it is reasonable to assume that all components are individually relevant, they indeed measure different characteristics, and the combination of all three is a valid way to cover the complexities of ethnic diversity.

To come up with the composite DELF, an equal weighting scheme has been applied to date. Following an extensive critique on the rather simplistic equal weighting of composite indices (Cahill, 2005; McGillivray and White, 1993), the call for a more elaborate weighting scheme, or at least a better foundation, is understandable.177 One approach widely discussed is the principal component analysis (PCA).178 Principal components are calculated as linear combinations of the original variables (the single characteristic DELF values in this case) in a way of explaining the largest part of its variation. The first principal component explains most of the variance, followed by the second and third principal component. In doing so, principal component analysis transforms correlated variables into uncorrelated ones and all principal components are orthogonal. The assigned loading factors can then be used to weight the sub-indices.179

The very high correlation of 0.999 between the DELF and the index based on PCA calculations (DELF_{PCA}) is seen in the upper part of Table 3.7. This suggests that one can resign from using the more complex weighting schemes and it underlines that none of the components dominates the other components in a problematic way.180

<table>
<thead>
<tr>
<th></th>
<th>DELF</th>
<th>DELF_{PCA}</th>
<th>DELF_{Geo}</th>
<th>DELF_{Pc}</th>
<th>ANM</th>
<th>Alesina</th>
<th>Annett</th>
</tr>
</thead>
<tbody>
<tr>
<td>DELF</td>
<td>1</td>
<td>0.999</td>
<td>0.963</td>
<td>0.994</td>
<td>0.698</td>
<td>0.628</td>
<td>0.630</td>
</tr>
<tr>
<td>DELF_{PCA}</td>
<td>0.999</td>
<td>1</td>
<td>0.963</td>
<td>0.994</td>
<td>0.697</td>
<td>0.630</td>
<td>0.630</td>
</tr>
<tr>
<td>DELF_{Geo}</td>
<td>0.963</td>
<td>0.963</td>
<td>1</td>
<td>0.959</td>
<td>0.707</td>
<td>0.632</td>
<td>0.651</td>
</tr>
<tr>
<td>DELF_{Pc}</td>
<td>0.994</td>
<td>0.994</td>
<td>0.959</td>
<td>1</td>
<td>0.736</td>
<td>0.662</td>
<td>0.671</td>
</tr>
<tr>
<td>ANM</td>
<td>0.698</td>
<td>0.697</td>
<td>0.707</td>
<td>0.736</td>
<td>1</td>
<td>0.800</td>
<td>0.874</td>
</tr>
<tr>
<td>Alesina</td>
<td>0.628</td>
<td>0.630</td>
<td>0.632</td>
<td>0.662</td>
<td>1</td>
<td>0.800</td>
<td>0.883</td>
</tr>
<tr>
<td>Annett</td>
<td>0.630</td>
<td>0.630</td>
<td>0.651</td>
<td>0.671</td>
<td>1</td>
<td>0.874</td>
<td>0.817</td>
</tr>
<tr>
<td>Fearon</td>
<td>0.607</td>
<td>0.606</td>
<td>0.626</td>
<td>0.621</td>
<td>0.748</td>
<td>0.817</td>
<td>0.795</td>
</tr>
</tbody>
</table>

Table 3.7: Rank correlation matrix for differently weighted DELF values and the most common ELF indices

Having discussed the possible redundancy of the components and ways to assign their weights, there are two ways to aggregate the components; using the arithmetic, or the

177 Chowdhury and Squire (2006) show that the vast majority of scholars still opt for the equally weighted average regarding aggregated development indices, despite ongoing discussions. For the HDI, Nguefack-Tsague et al. (2011) also provide a statistical reinforcement of the equal weighting scheme. An additional problem often raised is the implicit weighting due to different scales of the sub-indices (McGillivray and Noorbakhsh, 2004; Noorbakhsh, 1998). Through construction of the sub-indices, this problem does not apply to the DELF.


179 For the results of the PCA and further details, see Appendix C.2.

180 Additionally, the variances of the sub-indices are rather similar. So, none of the sub-indices would significantly bias the equally weighted index. For details on key statistical attributes of the single sub-indices, see Table 3.8.
geometric mean.\(^{181}\) Using a geometric mean does ‘penalize’ high dissimilarity in one of the components, however. This is often used in composite indices on various inequality measures, e.g., poverty, where the direct compensation of one component through another is not desired.\(^{182}\) Two individuals from the same ethno-racial and language backgrounds, who adhere to different religions, would be completely different in the case of a geometric mean because the religious component would be zero.\(^{183}\) That a certain similarity still prevails between both individuals/groups is obvious. Thus, for the application here, a form of compensation between components seems reasonable. In connection with the discussion above, the interpretation of the cumulative nature of the characteristics is more perspicuous and, additionally, argues in favor of an arithmetic mean. Due to these very different attributes, it is not surprising that the \(D ELF_{Geo}\) has a lower, yet still very high correlation to all the other \(D ELF\) values.

As an alternative, the introduction of a certain non-linearity of compensation between characteristics might be reasonable. This is, for example, promoted by Branisa et al. (2009). To allow for a certain compensation, one squares the components before the calculation of the arithmetic mean. This leads to an adjusted value of \(D ELF_{Pc}\). In line with Nardo et al. (2005), in this approach the weights are interpreted as trade-offs and not as importance coefficients.\(^{184}\)

Finally, the \(D ELF\) index should contain different information than other indices that try to measure ethnic fragmentation or diversity. Thus, the redundancy considerations regarding the components can be applied as a comparison to existing ELF indices. The results are found in the lower part of Table 3.7. All rank correlations between the most common ELF indices and the new \(D ELF\) fall below both redundancy thresholds.\(^ {185}\) Although already alluded to the theoretical discussion, where it was apparent that both indices measure different things (fragmentation versus diversity), the statistical results provide additional confirmation.

\(^{181}\) An additional aggregation for the \(D ELF_{PCA}\) index is not necessary because, by construction, the distance vector of the first principal components contains the weights and aggregation implicitly.

\(^{182}\) The HDI just recently switched from an arithmetic mean to a geometric one. To advance a country’s development it now needs to advance much more equally across the sub-indices than before, where one could compensate for one index with another. A geometric mean for an index would also imply a clear assignment of both a bad and good state for the values of zero and one. This is possible for poverty and development indices but not for the \(D ELF\), which describes a state between two extremes without valuation.

\(^{183}\) Collier and Hoeffler (2002), Collier and Hoeffler (2004) and Collier et al. (2004) use a multiplicative combination of the ethnic and religious fractionalization measure to assess ‘social fractionalization’. To avoid the dominance of one characteristic, where two groups are completely different, they add the index which is the greater to the product of both indices.

\(^{184}\) Thus, an individual can reduce the distance between another individual that does not adhere to the same religion by learning his language. For further theoretical discussions on weighting and differences between compensatory and non-compensatory approaches, see Munda and Nardo (2005). Branisa et al. (2009) offer a functional operationalization.

\(^{185}\) Note that the number of observations varies across the correlation values with the ELF indices due to their more limited observations.
The arithmetical average between the single characteristics is therefore the easiest way to operationalize the composite DELF index. Furthermore, it has the compensatory attributes between the characteristics that reflects their complementarity. This is not given by using the geometric mean, for example. By using the part compensation method and principal components, comparably adequate results are found to those of the simple arithmetic mean. As their correlation is rather high, the method used here follows the principle of keeping it as simple as possible.

### 3.6 Results

For each country, a similarity matrix is calculated, containing all \( s_{kl} \) for the weighting of mutual group similarities. *Tables C.2 and C.3 of Appendix C.2* detail the general similarity matrix calculation. The group similarity calculations are comparable to the ones within a country and for the difference between countries.

#### 3.6.1 Diversity measure within countries

The size of the respective \( K \times K \) matrices for each country is defined by the number of groups found in it, ranging from 3 to 884.

![Graphs showing combined and single characteristic DELF values against ELF values.](image)

**Figure 3.2:** Combined and single characteristic DELF values against ELF values.

---

186For further details on all weighting schemes, see *Appendix C.2*. A detailed discussion of the superiority of the equal weighting scheme is found in McGillivray and Noorbakhsh (2004), who conclude that more elaborate weighting schemes “produce values which are generally indistinguishable from values of the equally weights index” (McGillivray and Noorbakhsh, 2004, p. 15). Comparably, de Groot (2009) uses the same approach in his ethno-linguistic affinity index.
To make the differences between the ELF and DELF values clear, Figure 3.2 shows the influence of the various characteristics.\textsuperscript{187} By adjusting for the language differences only, reduces the values by less than when all three characteristics are considered. The most influential changes emerge if religion is taken into account, since in many countries a majority religion is present, which acts as a unifying characteristic. The combined DELF, weighting all three characteristics, yields more consistent values, which is confirmed in Table 3.8. The standard deviation of the composite DELF is considerably lower than those of the decomposed indices.

Religious and language homogeneity, in particular, are spread differently across regions. This is why the adjustments also vary significantly between regions. In Latin-America,\textsuperscript{188} Spanish is the dominant language, although there are different ethno-racial and/or religious groups. The language similarities add to a higher affinity between the groups and, in turn, lower the DELF values. Table 3.9 summarizes the mean values for different ELF and DELF specifications across regions. Additionally, it compares the average ranks of the countries in the respective groups. A rank of one is assigned to the most heterogeneous countries, i.e., the countries with the highest ELF or DELF values. Comparing both ranks gives a good indication of how large the adjustments in the DELF calculation are compared to the standard ELF values.

<table>
<thead>
<tr>
<th>Index</th>
<th>Observations</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELF</td>
<td>210</td>
<td>0.563</td>
<td>0.270</td>
<td>0.019</td>
<td>0.982</td>
</tr>
<tr>
<td>DELF</td>
<td>210</td>
<td>0.252</td>
<td>0.157</td>
<td>0.006</td>
<td>0.636</td>
</tr>
<tr>
<td>DELFL</td>
<td>210</td>
<td>0.353</td>
<td>0.243</td>
<td>0.008</td>
<td>0.942</td>
</tr>
<tr>
<td>DELFE</td>
<td>210</td>
<td>0.255</td>
<td>0.176</td>
<td>0.002</td>
<td>0.708</td>
</tr>
<tr>
<td>DELFR</td>
<td>210</td>
<td>0.148</td>
<td>0.188</td>
<td>0.000</td>
<td>0.648</td>
</tr>
</tbody>
</table>

Table 3.8: Main statistical characteristics of DELF values, decomposed for all ethnicity characteristics

Sub-Saharan Africa (SSA) demonstrates a much higher value when measured by the ELF compared to the DELF, resulting in a negative rank delta. As seen earlier, this region includes countries with the highest number of groups, mirrored by high ELF values. However, if one takes the similarity between the groups into account, the ranks decrease. Eastern Europe, in contrast, shows much more diversity when considering the DELF value rather than the ELF value.

More interesting is the decomposition of the DELF into its single characteristics. For the language characteristic, Latin America hosts the most homogeneous countries, whereas Sub-Saharan Africa again shows the most heterogeneous ones. Taking into account only the ethno-racial aspect, Latin America shows the highest diversity. This might come from the interbreeding of the native Indian population with the high number of descendants from the Western colonial powers and the resulting Mestizo progeny. The region with

\textsuperscript{187}Both indices are based on WCE data.

\textsuperscript{188}Includes the Caribbean.
the most homogeneous countries in this regard is Eastern Europe, a region where outside powers have interfered less. The religious characteristic again demonstrates the expected distribution. Sub-Saharan Africa has the most religiously heterogeneous countries and Western and Latin American countries, with high numbers of Christians, host the most homogeneous ones. Not surprisingly, the Middle East and Northern African (MENA) countries also show values indicating rather homogeneous religious characteristics, which is not surprising considering the high proportion of Muslims in these areas. Most countries that have a majority religion, i.e., more than 60% of the population either adhere to Christianity (133 countries) or to Islam (43 countries), exhibit rather low religious DELF values. For all other countries, where there is either no majority religion or it is made up of another denomination, show significantly higher religious DELF values. Also, their average overall DELF rank is substantially higher than when only taking the number of groups in the ELF value into account.

<table>
<thead>
<tr>
<th>Region</th>
<th>Obs.</th>
<th>ELF</th>
<th>DELF</th>
<th>DELFL</th>
<th>DELFE</th>
<th>DELFR</th>
<th>Rank ELF</th>
<th>Rank DELF</th>
<th>Delta Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asia</td>
<td>40</td>
<td>0.608</td>
<td>0.290</td>
<td>0.435</td>
<td>0.240</td>
<td>0.194</td>
<td>93.3</td>
<td>90.8</td>
<td>2.5</td>
</tr>
<tr>
<td>E. Europe</td>
<td>29</td>
<td>0.389</td>
<td>0.197</td>
<td>0.261</td>
<td>0.204</td>
<td>0.126</td>
<td>145.9</td>
<td>125.0</td>
<td>20.8</td>
</tr>
<tr>
<td>L. America</td>
<td>38</td>
<td>0.509</td>
<td>0.227</td>
<td>0.220</td>
<td>0.386</td>
<td>0.075</td>
<td>121.3</td>
<td>114.5</td>
<td>6.8</td>
</tr>
<tr>
<td>MENA</td>
<td>21</td>
<td>0.558</td>
<td>0.249</td>
<td>0.358</td>
<td>0.275</td>
<td>0.114</td>
<td>108.1</td>
<td>107.0</td>
<td>1.2</td>
</tr>
<tr>
<td>SSA</td>
<td>49</td>
<td>0.741</td>
<td>0.319</td>
<td>0.490</td>
<td>0.219</td>
<td>0.248</td>
<td>62.6</td>
<td>81.2</td>
<td>-18.6</td>
</tr>
<tr>
<td>W. Count.</td>
<td>33</td>
<td>0.465</td>
<td>0.184</td>
<td>0.279</td>
<td>0.206</td>
<td>0.066</td>
<td>128.7</td>
<td>130.9</td>
<td>-2.2</td>
</tr>
<tr>
<td>World</td>
<td>210</td>
<td>0.563</td>
<td>0.252</td>
<td>0.353</td>
<td>0.255</td>
<td>0.148</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Muslim</td>
<td>43</td>
<td>0.571</td>
<td>0.262</td>
<td>0.389</td>
<td>0.271</td>
<td>0.127</td>
<td>105.6</td>
<td>100.7</td>
<td>4.9</td>
</tr>
<tr>
<td>Christian</td>
<td>133</td>
<td>0.519</td>
<td>0.208</td>
<td>0.299</td>
<td>0.251</td>
<td>0.076</td>
<td>115.7</td>
<td>121.2</td>
<td>-5.7</td>
</tr>
<tr>
<td>Other</td>
<td>34</td>
<td>0.729</td>
<td>0.407</td>
<td>0.519</td>
<td>0.249</td>
<td>0.454</td>
<td>65.6</td>
<td>50.1</td>
<td>15.5</td>
</tr>
</tbody>
</table>

Table 3.9: Mean ELF and DELF values and ranks for all regions and countries with main majority religions

The single country perspective shows even more considerable adjustments. The ELF and DELF values of each country are listed in Table C.7 of Appendix C.3. The countries are ordered according to their ELF values in descending order, from the most heterogeneous country to the most homogeneous country. The third column depicts their corresponding DELF values and DELF ranks. The difference between the ELF and DELF ranks is shown in column four. The next column outlines the DELF values, decomposed for each characteristic, which helps to better illustrate the adjustments. An adjustment of over 40 places is seen by half of the 10 most diverse countries. Looking at the lower end, one sees only marginal adjustments, as expected. The 15 most homogeneous countries are, with three exceptions, the same for both indices. For the other countries, however, significant adjustments are found. For example, Zambia, the Republic of Congo, Zimbabwe, Angola...
and Italy, which are treated as much more homogeneous by the \( DELF \) compared to the \( ELF \), show difference in ranking of more than 100 places are. Nevertheless, one also finds adjustments in the opposite direction, i.e., countries that have a higher diversity rank based on \( DELF \) values. The countries with the most significant adjustments in this regard – all more than sixty places – are Kazakhstan, Bahrain, Macedonia, Lebanon, Sudan and the Russian Federation. These upward changes are mainly driven by relatively high language diversity.

### 3.6.2 Similarity measure between countries

To date, most authors have focused on the assessment of ethnicity within a country, as has this essay. This has also been the case in analyzing a country’s growth or conflict incidence. De Groot (2009) expands upon this and proposes his index of ethno-linguistic affinity (ELA) to measure the similarities between two neighboring countries. He shows that conflict spillovers are more likely between contiguous countries sharing stronger ethnic similarities. The extended calculation for the \( DELF \) between countries is nearly identical to Equation 3.7, and is defined through:

\[
DELF_{ij} = 1 - \sum_{k=1}^{K} \sum_{m=1}^{M} p_{ik} p_{jm} \hat{s}_{km} \tag{3.10}
\]

where country \( i \) hosts groups \( k = 1,\ldots,K \), and country \( j \) groups \( m = 1,\ldots,M \), respectively. The distance between the two groups \( k \) and \( m \) is given through \( \hat{s}_{km} \). The result is the expected dissimilarity between two individuals randomly drawn from each country.

The 210 countries analyzed here give a matrix containing over 150 million similarity values and nearly 44,000 dyadic relations between countries.\(^{190}\) Due to the amount of country-pairs, only a discussion of averages and some tuples with the highest discrepancy is offered here.\(^{191}\) Naturally, all \( DELF \) values are much higher than those for individual countries. Table C.8 of Appendix C.3 lists the mutually most similar and dissimilar countries at the single country level.\(^{192}\) Many of the mutually most similar countries come from the MENA region. The religious homogeneity of this region plays an important role in their overall similarity level. It is not surprising that the most dissimilar pairs are matches between Asian and African countries. Except for some minority migrant groups, one does

---

\(^{190}\) This significantly exceeds the 2,809 dyadic relations offered by de Groot (2009) for the 53 African countries.

\(^{191}\) The complete data set can be received upon request.

\(^{192}\) In general the interpretation of the \( DELF \) value between countries ranging between zero and one is comparable to the case of \( DELF \) values within countries. Two countries that consist of groups that share not a single characteristic show a mutual \( DELF \) value of one, being completely different. Lower values of \( DELF \) correctly indicate countries that share more characteristics and thus are more ‘similar’. However, the theoretical country setup maximizing the similarity between two countries (minimizing the \( DELF \) value) deviates in its limit from the generally understood meaning of the word ‘similar’. This is discussed in more detail in Appendix C.2.6. I would like to thank Walter Zucchini for this important comment.
not find many shared ethnic characteristics between these countries and all their values are close to one.

A regional aggregation also offers some interesting insights. For the calculation of the regional averages, the \textit{DELF} values between countries are adjusted for the different population sizes of the respective country pairs.\footnote{For the weighting, population data averages for 2005–2010 from the \textit{World Development Indicators} World Bank (2011) were used. For more details on how regional averages are calculated and the differences in the calculation of \textit{DELF} values between countries, see Appendix C.2.7.} Table 3.10 summarizes the regional and global averages.

<table>
<thead>
<tr>
<th>Regional \textit{DELF}</th>
<th>Country pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asia</td>
<td>0.719 1,600</td>
</tr>
<tr>
<td>Eastern Europe</td>
<td>0.479 841</td>
</tr>
<tr>
<td>Latin America</td>
<td>0.340 1,444</td>
</tr>
<tr>
<td>North Africa &amp; Middle East</td>
<td>0.430 441</td>
</tr>
<tr>
<td>Sub-Saharan Africa</td>
<td>0.643 2,401</td>
</tr>
<tr>
<td>Western Countries</td>
<td>0.572 1,089</td>
</tr>
<tr>
<td>World</td>
<td>0.841 44,100</td>
</tr>
</tbody>
</table>

\textbf{Table 3.10:} \textit{DELF} average by main geographical regions

The global cultural diversity measured by the \textit{DELF} displays an average of 0.84. Asia exhibits the highest diversity level compared to all other regions. Thus, from a regional perspective, Asia seems to be the most diverse region, and not SSA.\footnote{Note that from the single country perspective, SSA still has the countries with the highest internal heterogeneity. This is an indication that the drawing up of borders in Asia proceeded more ‘endogenously’ than the method used in SSA by the colonial powers.} Latin America, in contrast, displays the least interregional diversity.

The regional level of diversity plays an important role in the European Union (EU). The success of European integration is often questioned by the high level of cultural diversity. This was debated before the last enlargement in particular, when the EU grew from 15 to 25 and shortly after to 27 member states. It will eventually lead to even more controversial debates regarding future enlargement plans. With the above approach, the developments in the level of diversity through language, ethno-racial, and religious characteristics, can easily be traced.

\textit{Figure 3.3} shows the diversity level of the EU for each wave of enlargement.\footnote{For more details on the different waves of enlargement in the EU, and the respective diversity levels, see Table C.9 of Appendix C.3.} The predecessor of today’s EU was initiated in 1952, including Belgium, France, Germany, Italy, Luxembourg and the Netherlands. This ‘core Europe’, which it is often referred to, displayed a regional \textit{DELF} value of 0.37. The next two enlargement waves added nearly 25% to the total population. However, these countries were not overly different from the existing group and were internally rather homogeneous. Hence, the \textit{DELF} only slightly increased. The addition of Portugal and Spain in 1986, two populous and very homogeneous countries, slightly decreased the overall level of European diversity, whereas
the huge enlargement of 10 countries in 2004, and of two more in 2007, again increased the DELF level significantly.\textsuperscript{196} Looking at potential future enlargements, the admission of mainly Balkan states, as well as Iceland (EU+B), would not change the status quo greatly. The highest increase in diversity within the EU would result from admitting Turkey (EU+T). The increased cultural diversity Turkey would bring to the EU can’t be judged as good or bad, per se – however, it offers an easy target for exploitation of these differences and political agitation. This was already the case during earlier waves of enlargement which only displayed marginal increases in the EU’s diversity. The increase Turkey would bring, as stated, would be far greater, thus the potential for exploitation and political agitation could be far greater.

Finally, the DELF values between countries are compared with the most widely used measure of cultural distance between countries, its genetic distance. By matching these with the detailed data on genetic diversity compiled by Spolaore and Wacziarg (2009), yields only a very limited correlation (Table 3.11).\textsuperscript{197} The rank correlation of genetic distance and the composite DELF is only 0.25, and thus fail to meet both of the redundancy thresholds discussed above.\textsuperscript{198} This comparison underlines that the genetic distance data is hardly a good proxy for the ‘cultural’ differences between countries.

\textsuperscript{196}One important caveat applies for this. As essay 2 outlined, cultural heterogeneity levels are subject to change. As the underlying data for the DELF calculation is dated for the years 2005–2010, using it for time frames of over 50 years ago will lead to distorted values. Thus, the DELF values for the EU enlargement for the earlier years can only be taken as an indication. The changing DELF values are only attributable to compositional changes of the European Union and not to expectable changes over time.

\textsuperscript{197}Spolaore and Wacziarg (2009) construct two measures of genetic relatedness between countries. One is based only on the genetic distances between the plurality ethnic groups of each country. The second is a measure of weighted genetic distance of all groups. The latter construction is more comparable to the one employed in this essay.

\textsuperscript{198}As expected from the characteristic definition, the highest correlation of the genetic data is with the ethno-racial DELF values at 0.7. Both are correlated but still seem to measure different things.
Chapter 3. Measuring Ethnic Diversity

<table>
<thead>
<tr>
<th></th>
<th>(DELF)</th>
<th>(DELF_L)</th>
<th>(DELF_E)</th>
<th>(DELF_R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(DELF)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43890</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(DELF_L)</td>
<td>0.566</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>43890</td>
<td>43890</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(DELF_E)</td>
<td>0.489</td>
<td>0.636</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>43890</td>
<td>43890</td>
<td>43890</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(DELF_R)</td>
<td>0.899</td>
<td>0.363</td>
<td>0.193</td>
<td>1</td>
</tr>
<tr>
<td>43890</td>
<td>43890</td>
<td>43890</td>
<td>43890</td>
<td></td>
</tr>
<tr>
<td>Genetic Distance</td>
<td>0.245</td>
<td>0.484</td>
<td>0.697</td>
<td>0.018</td>
</tr>
<tr>
<td>30800</td>
<td>30800</td>
<td>30800</td>
<td>30800</td>
<td>30800</td>
</tr>
</tbody>
</table>

Table 3.11: Rank correlations between \(DELF\), its sub-indices and genetic distance data (observations in italics)

3.7 Conclusion

Taking the mutual (dis)similarities between ethnic groups into account, the new \(DELF\) index covers a new and very important aspect of ethnicity: its diversity. This additional aspect was ignored by the most commonly used measures of ethnicity. The \(DELF\) index for 210 countries shows considerable differences between countries and regions. The differences suggest that it indeed measures different aspects of ethnicity, which might have a contrasting effect on the socio-economic problems under investigation.

Many current papers analyzing the role of ethnicity based on the ELF index can profit from taking the mutual (dis)similarities between individual groups into account. In countries, where ethnic groups show higher differences, it might be even more difficult to agree on public goods (e.g., infrastructure or social security systems), as has already been shown by Alesina et al. (1999). Caselli and Coleman (2008) discuss the importance of barriers between groups to prevent assimilation between them on the incidence of wars. This is exactly what Collier and Hoeffler (1998, 2004), Collier et al. (2009) and Fearon and Laitin (2003) try to find in their analyses. i.e., whether ethnic fragmentation increases the incidence of wars. Their results do not find a robust influence of ELF on conflict incidence. It might still be the case that there is a strong influence of ethnic diversity on conflicts, but the applied ELF index does not measure the appropriate aspect of ethnicity in order to prove this. Additionally, the possibility to analyze the single characteristic \(DELF\) for very specific questions offers new room for application. Akdede (2010), for example, shows the different implications of ethnic and religious fractionalization on democratic institutions.

Research that leveraged genetic distances to assess the dissimilarity between countries should equally profit from employing the \(DELF\) between countries. It offers a much more comprehensive data set of ‘cultural’ affinity between nations. As de Groot (2009) concludes, it is not necessarily the geographical distance, often used in spatial economics, which is being applied to assess the influences one country might have on others. Nor does genetic distance really offer a satisfying alternative. The \(DELF\) values between countries
offer an excellent and valid extension of the analysis into spillover effects between countries. De Groot (2009) shows the role cultural affinity between neighboring countries plays in the spillover of conflicts.

Trust is associated closely with more homogeneous and similar country setups. Genetic distance only covers trust in a very limited way. Trust is seldom hidden in the genetic code, evolving out of the interaction between individuals whose cultural backgrounds play an important role.199 Leveraging genetic distance is even more problematic in Spolaore and Wacziarg’s (2009) analysis on the spillover effect of innovations and development between countries. Imitation and adaptation costs of innovations rely significantly more on the ‘cultural’ barriers (different language, ethno-racial background and beliefs) than on the biological ones (genes).

Nevertheless, there are some caveats that one cannot overlook. As the data source used is somewhat unique in its combination of all characteristics, only limited robustness checks with other sources on the combination of the characteristics are possible. Secondly, the weighting of the three sub-indices is debatable, as is the case for most composite index calculations. Here, the most general approach is used. For specific questions, different emphasis might be given to specific characteristics. The clear discussion and overview of the single sub-indices should encourage every researcher to do so. Finally, there might be country or region-specific characteristics influencing cultural diversity not covered in the (globally comparable) three characteristics treated in this essay. The caste system in India would be one example. Thus, for a country or region-specific analysis, the diversity data offered might have restricted relevance. Nevertheless, the approach discussed here can still be applied.

In the above cases the DELF index should be more appropriate than the ELF index as it incorporates the fundamental concept of diversity. The extension to measure cultural dissimilarities between nations offers a good alternative to the applied genetic distance data. The broad foundation and the detailed new data set should be a call to critically review the usage of the ELF index and the genetic distance data. Additionally, it provides a starting point for new research on the specific role of the diversity of countries.

199 For an indication of how a common language increases trust and common identification in a case study for the US, see Chong et al. (2010). Falck et al. (2010) show that German cross-regional migration and economic exchange can be attributed to dialect similarities from the 19th century that remain today.