ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Kolo, Philipp

Working Paper

Measuring a new aspect of ethnicity: The appropriate diversity index

IAI Discussion Papers, No. 221

Provided in Cooperation with:
Ibero-America Institute for Economic Research, University of Goettingen

Suggested Citation: Kolo, Philipp (2012) : Measuring a new aspect of ethnicity: The appropriate diversity index, IAI Discussion Papers, No. 221, Georg-August-Universität Göttingen, Ibero-America Institute for Economic Research (IAI), Göttingen

This Version is available at:
https://hdl.handle.net/10419/92997

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

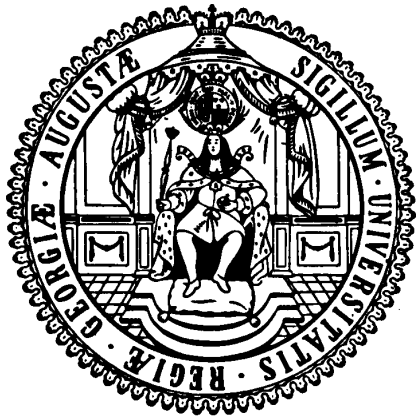
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Ibero-Amerika Institut für Wirtschaftsforschung
Instituto Ibero-Americano de Investigaciones Económicas
Ibero-America Institute for Economic Research
(IAI)

Georg-August-Universität Göttingen
(founded in 1737)



Diskussionsbeiträge · Documentos de Trabajo · Discussion Papers

Nr. 221

Measuring a New Aspect of Ethnicity –
The Appropriate Diversity Index

Philipp Kolo

July 2012

Platz der Göttinger Sieben 3 · 37073 Goettingen · Germany · Phone: +49-(0)551-398172 · Fax: +49-(0)551-398173
e-mail: uwia@gwdg.de · http://www.iai.wiwi.uni-goettingen.de

Measuring a New Aspect of Ethnicity – The Appropriate Diversity Index

Philipp Kolo*
Georg-August University Göttingen, Germany

March 2012

Abstract

Combined with the expansion of economic literature on the role of ethnicity, new indices were developed to do justice to its complexity. The current indices are generally based on pre-defined groups, disregarding the (dis)similarities between them. This is sufficient to calculate the most common indices of ethnic fractionalization and polarization. But they do not measure ethnic diversity as for any diversity index the introduction of distances between groups is essential.

This paper includes the distance between groups as a crucial aspect of a country's ethnic set-up. Language, ethno-racial and religious characteristics are combined in a consistent way for a composite (dis)similarity value. The resulting distance adjusted ethno-linguistic fractionalization index ($DELF$) is based on an extensive amount of data, containing more than 12,000 groups defined along all three characteristics and covering a wide range of countries. By applying the equivalent approach as that of the diversity measure for single countries, the $DELF$ offers an assessment of cultural differences between countries. As the new index measures a country's ethnic diversity, it is a good starting point to review some of the existing approaches linking ethnicity to economic outcomes.

Key words: Composite Index, Distance, Ethno-Linguistic Fractionalization (ELF), Measurement
JEL classification: C43, D63, O10, Z10

*Ph.D. candidate at the Chair of Development Economics, Georg-August University Goettingen, Germany. Contact: philipp.kolo@stud.uni-goettingen.de. I would like to thank Prof. Stephan Klasen for his continuous support as well as Axel Dreher, Joan Esteban, Olaf de Groot, Laura Mayoral and Walter Zucchini for their useful suggestions and helpful comments. I benefited a lot from the comments of participants to seminars at the Georg-August University Göttingen and the DIW (Berlin).

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.[1] 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 ethnolinguistic 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, 2004) or the role of regional segregation of ethnicity (Alesina and Zhuravskaya, 2011) have been studied more intensively.[2]

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.[3] Fearon (2003) summarizes the absence of a clear-cut definition of ethnic groups and states "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 absolutely fundamental concept in the measurement of dissimilarity that it must play an essential role in any meaningful theory of diversity or classification" (Weitzman, 1992, p. 365).[4] 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

[1] Ethnic 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 (Akdede, 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, 2005, 2008).

[2] For 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).

[3] For a similar line of critique, see Lind (2007).

[4] 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 (1992).

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., infrastructure 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).[5] The main aim of this article 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 article is structured as follows. Section 2 briefly summarizes the current discussion surrounding the conceptual und measurement problems. In section 3, the theoretical background of the new similarity parameters is outlined. Section 4 introduces the data sources used. Section 5 discusses the operationalization of the new distance adjusted ethno-linguistic fractionalization index ($DELF$), and compares it with existing measures. Section 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 7 summarizes the key findings, concludes and gives an outlook for further research.

2 Different aspects of ethnicity and its measurement

Alesina et al. (2003) describe ethnicity as a "rather vague and amorphous con-

[5] 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.

cept" (Alesina et al., 2003, p. 160) that makes any measurement hard to grasp.[6] To better operationalize ethnicity, this article 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 article.[7] Following these authors, ethnicity is defined in this article 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 article 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.[8] 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.[9] 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:

$$ELF = 1 - \sum_{i=1}^{K} p_i^2, \qquad i = 1, ...K \tag{1}$$

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

[6] Brown and Langer (2010) offer a broad summary of the recent discussion surrounding the definitions of ethnicity as well as its measurement problems.

[7] They include language, ethnic origin, skin pigmentation, race, culture or religion, and nationality as characteristics to describe ethnicity.

[8] 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).

[9] 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.

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).[10] 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:

$$POL = 1 - \sum_{i=1}^{K} \left(\frac{0.5 - p_i}{0.5} \right)^2 \cdot \; p_i, \qquad i = 1, ...K \tag{2}$$

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 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.

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.[11] 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 be-

[10] Their approach goes back to earlier work of Esteban and Ray (1994).

[11] This, however, is the main drawback of its operationalization, as reliable data on individuals are seldom available, especially in developing countries.

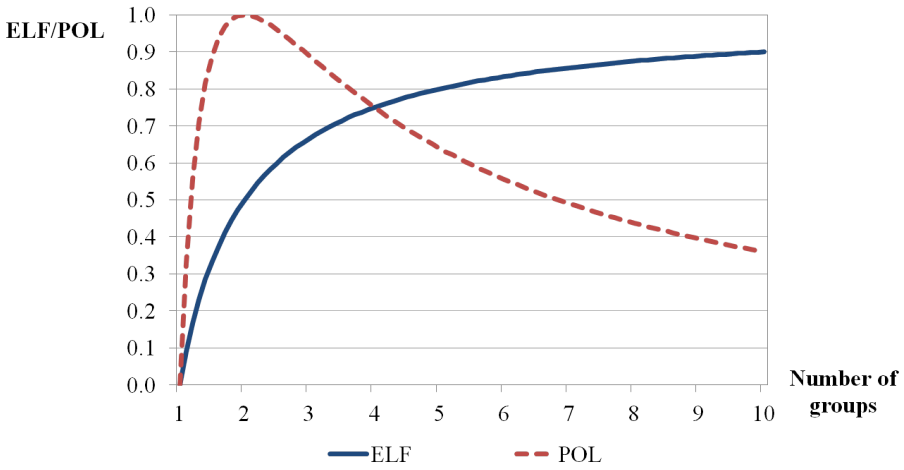


Figure 1: ELF and POL values depending on the number of equally sized groups

tween two individuals i and j for all $i, j \in \{1, ..., N\}$ is given through s_{ij}, with:

$$1 \geq \quad s_{ij} \quad \geq 0 \tag{3}$$

$$s_{ii} \quad = \quad 1 \tag{4}$$

$$s_{ij} \quad = \quad s_{ji} \tag{5}$$

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:

$$G(S_N) = 1 - \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} s_{ij} \tag{6}$$

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 (6)*, it follows that:

$$\begin{aligned} G(S_n) &= 1 - \frac{1}{N^2} \sum_{k=1}^{K} \sum_{l=1}^{K} m_k \; m_l \; \hat{s}_{kl} \\ &= 1 - \sum_{k=1}^{K} \sum_{l=1}^{K} \frac{m_k}{N} \; \frac{m_l}{N} \; \hat{s}_{kl} \\ &= 1 - \sum_{k=1}^{K} \sum_{l=1}^{K} p_k \; p_l \; \hat{s}_{kl} = DELF \end{aligned} \tag{7}$$

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 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 (1)*, where the ELF is specified.[12] 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.[13]

Finally, de Groot (2009) assessed the ethnic affinity between African nations.[14] 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 article. De Groot (2009), however, only offers data on ethnic affinity between countries and limits his assessment to Africa. This article consequently extends the work of all three studies.

[12] Note that due to the construction of *Equation (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 (2005, 2008) and Esteban and Ray (2011). In the latter, the index is labeled as the 'Greenberg-Gini' index.

[13] The superior theoretical explanatory power of such an index is also discussed in Ginsburgh and Weber (2011).

[14] 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 article follows the logic of the ELF, where higher values denote more fragmented countries.

3 Calculation of the distance values

For the calculation of the distance values, this article draws on Fearon (2003). His approach is adapted for three ethnicity characteristics: language, ethnoracial and religious identification. Taking a broader set of characteristics and similarity measures into account offers a more multifaceted picture.[15]

3.1 Language classification

Language is probably the most researched and operationalized characteristic.[16] 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 decision 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 article 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.[17] In total, 6,656 distinct languages are contained in the data analyzed. Two persons speaking one language are treated as completely similar ($s_{ij} = 1$).[18] The more charac-

[15] 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.

[16] Ginsburgh and Weber (2011, Ch. 3) offer a good overview of the different approaches to assess the distances between languages.

[17] 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`.

[18] 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

ters of the assigned code two languages share, the more similar they are. The structure is depicted in *Table 1*.

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

Table 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 $\bar{s}_{kl}^{L}$. 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 $\bar{s}_{kl}^{L}$ with $\bar{s}_{kl}^{L} = 1$.[19] In the case of the example $\bar{s}_{kl}^{L}$ takes a value of 0.35.

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 evolutionary 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 phylographic 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.[20]

This was certainly a pioneering piece of work but also demonstrates some

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.

[19] For a discussion on alternative similarity values, see *Appendix A.2*.

[20] 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.

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.[21] 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.[22] 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.[23] 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.[24]

In order to combine these views and caveats, this article 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.[25] Although those characteristics are closely linked in their development, their role for mutual understanding differs and is treated as cumulative in the subsequent analysis.[26]

Analogous to the pure language case, the different levels of ethno-racial classification are summarized in *Table 2*.[27] 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

[21] 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).

[22] 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.

[23] 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.

[24] 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.

[25] 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).

[26] This approach is also followed by de Groot (2009).

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

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.[28]

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

Table 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 ($\bar{s}^E_{kl}$) are calculated with the same decrease in slope of the similarity values being found as that of the language characteristic.[29]

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 $\bar{s}^E_{kl}$ of 0.88.

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.[30] Religious identification is in a certain way, an especially potent, but easily implemented instrument to expand ones political power through mobilizing one's followers.

[28] 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.

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

[30] 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).

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 commonalties. 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 article 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.[31] This approach should be adjusted as better data become available.

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.[32] 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 article, 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 var-

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

[32] 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.

ious 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.[33] 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.

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.[34] 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).[35] It contains data for over 12,000 groups in 210 countries, classified according to language, ethno-racial group and religion.[36] 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.[37] Due to the Christian background of the publishing

[33] The same might be true for religion and languages, or even dialects.

[34] Akdede (2010) gives a good overview of the data sources used in a broad set of influential articles and discusses their differences.

[35] 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.

[36] 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.

[37] See, for example, Annett (2001), Barro (1999), Barro and McCleary (2003), Collier and

Table 3: Descriptive statistics of ethnic groups by geographical area

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

[a]Western Europe and Australia, Canada, Greenland, Japan, New Zealand and United States.

[b]Includes the Caribbean.

[c]Includes the Pacific islands.

[d]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.

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.[38] The high granularity of data might still raise some questions 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.[39]

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* 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.[40] 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.[41] Excluding these three countries, Sub-Saharan Africa is again the region with the most fragmented countries.[42] This becomes even clearer when one compares the other figures in *Table 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.

Hoeffler (2002, 2004), Collier et al. (2004), Garcia-Montalvo and Reynal-Querol (2005, 2008, 2010), Loh and Harmon (2005), or Okediji (2005).

[38] 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 Hoeffler (2002, 2004) and Collier et al. (2004) used it for their index on religious fractionalization. However, Garcia-Montalvo and Reynal-Querol (2005) 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.

[39] For more details on these robustness checks, see *Appendix A.1.*

[40] The Asian region includes the Pacific countries and islands.

[41] 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.

[42] Excluding these three countries, the average number of groups per country in Asia would only amount to 56.

Also, the number of countries that have a majority group of 50% is significantly lower.

Source		Obs.	Mean	Std. Dev.	Min.	Max.
ELF	ANM	169	0.458	0.273	0.000	0.984
	Alesina	186	0.440	0.257	0.000	0.930
	Annett	144	0.479	0.275	0.010	0.950
	Fearon	153	0.471	0.270	0.002	0.953
	WCE	210	0.563	0.270	0.019	0.982

Table 4: Main statistical characteristics of ELF values for different sources

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.[43] *Table 4* confirms this by showing the summary statistics of the ELF values for the various sources described earlier.

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.[44] For the same reason, Okediji (2005) proposes including ethnic differentiation alongside racial and religious characteristics.[45] 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).[46]

[43] The theoretical attributes of the ELF and POL are nicely met by the *WCE* data. *Figure 9* of *Appendix A.3* shows the increasing ELF values in conjunction with a rising number of groups within a country.

[44] 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)

[45] 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.

[46] 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 char-

The most common approach when incorporating different characteristics into a combined index is to assign equal weights to all of its components.[47] Following this approach, the $DELF$ is calculated according to *Equation (7)* as:

$$DELF = 1 - \sum_{k=1}^{K} \sum_{l=1}^{K} p_k \; p_l \; \hat{s}_{kl} \tag{8}$$

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

$$\hat{s}_{kl} = \frac{1}{3} \left[\; \bar{s}_{kl}^{L} \; + \; \bar{s}_{kl}^{E} \; + \; \bar{s}_{kl}^{R} \; \right] \tag{9}$$

where $\bar{s}_{kl}^{L}$, $\bar{s}_{kl}^{E}$ and $\bar{s}_{kl}^{R}$ are the respective similarity values for the language, ethno-racial and religious classification.[48] The single characteristic $DELF$s are equally calculated using *Equation (9)*. Instead of the composite similarity measure ($\hat{s}_{kl}$) the characteristics specific similarity values ($\bar{s}_{kl}^{L}$, $\bar{s}_{kl}^{E}$, $\bar{s}_{kl}^{R}$) 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.[49] 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 5*.[50]

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

acteristic. If there is no strong reason for deviating from the equal weighting, Haq (2006) argues strongly for this principle.

[47] 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 Booysen (2002).

[48] The main focus of this article 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 (2)*. This would then transform to a distance adjusted POL index with: D-POL $= \sum_{k=1}^{K} \sum_{l=1}^{K} p_k^2 \cdot p_l \cdot \hat{s}_{kl}$ (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.

[49] Cahill (2005), McGillivray and Noorbakhsh (2004), Branisa et al. (2009) and Cho et al. (2011) subsequently used this decision rule.

[50] 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.

	$DELF$	$DELF_L$	$DELF_E$	$DELF_R$
$DELF$	1			
$DELF_L$	0.904	1		
$DELF_E$	0.714	0.537	1	
$DELF_R$	0.665	0.529	0.195	1

Table 5: Rank correlation for the composite $DELF$ and its components

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).[51] 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 6* shows the correlations seen in *Table 5*, split between equally sized quintiles.

		Quintiles				
	All obs.	**1**	**2**	**3**	**4**	**5**
	$DELF$	$DELF$	$DELF$	$DELF$	$DELF$	$DELF$
$DELF_L$	0.904*	0.282	0.483*	0.401*	0.556*	0.814*
$DELF_E$	0.714*	0.056	0.156	0.050	0.141	0.815*
$DELF_R$	0.665*	0.569*	0.142	0.004	0.276	0.372*

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

Table 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.[52] One approach widely discussed is the principal component

[51] 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.

[52] 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

analysis (PCA).[53] 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.[54]

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 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.[55]

		$DELF$	$DELF_{PCA}$	$DELF_{Geo}$	$DELF_{Pc}$	ANM	Alesina	Annett
DELF	$DELF$	1						
	$DELF_{PCA}$	0.999	1					
	$DELF_{Geo}$	0.963	0.963	1				
	$DELF_{Pc}$	0.994	0.994	0.959	1			
ELF	ANM	0.698	0.697	0.707	0.736	1		
	Alesina	0.628	0.630	0.632	0.662	0.800	1	
	Annett	0.630	0.630	0.651	0.671	0.874	0.883	1
	Fearon	0.607	0.606	0.626	0.621	0.748	0.817	0.795

Table 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 geometric mean.[56] 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

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 $DELF$.

[53] For a discussion and its application, mainly to the HDI, see Jolliffe (1973), Ram (1982), Ogwang (1994), Noorbakhsh (1998) or Ogwang and Abdou (2003).

[54] For the results of the PCA and further details, see *Appendix B*.

[55] 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 8*.

[56] An additional aggregation for the $DELF_{PCA}$ index is not necessary because, by construction, the distance vector of the first principal components contains the weights and aggregation implicitly.

compensation of one component through another is not desired.[57] 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.[58] 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 $DELF_{Geo}$ has a lower, yet still very high correlation to all the other $DELF$ 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 $DELF_{Pc}$. In line with Nardo et al. (2005), in this approach the weights are interpreted as trade-offs and not as importance coefficients.[59]

Finally, the $DELF$ 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 7*. All rank correlations between the most common ELF indices and the new $DELF$ fall below both redundancy thresholds.[60] 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.

The arithmetical average between the single characteristics is therefore the

[57] 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 $DELF$, which describes a state between two extremes without valuation.

[58] 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.

[59] 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.

[60] Note that the number of observations varies across the correlation values with the ELF indices due to their more limited observations.

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.[61]

6 Results

For each country, a similarity matrix is calculated, containing all $\hat{s}_{kl}$ for the weighting of mutual group similarities. *Tables 13* and *14* of *Appendix B* detail the general similarity matrix calculation. The group similarity calculations are comparable to the ones within a country and for the difference between countries.

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.
To make the differences between the ELF and $DELF$ values clear, *Figure 2* shows the influence of the various characteristics.[62] 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 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,[63] Spanish is the dominant language, although there are different ethno-racial and/or religious groups. The language similarities add

[61] For further details on all weighting schemes, see *Appendix B*. 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.

[62] Both indices are based on *WCE* data.

[63] Includes the Caribbean.

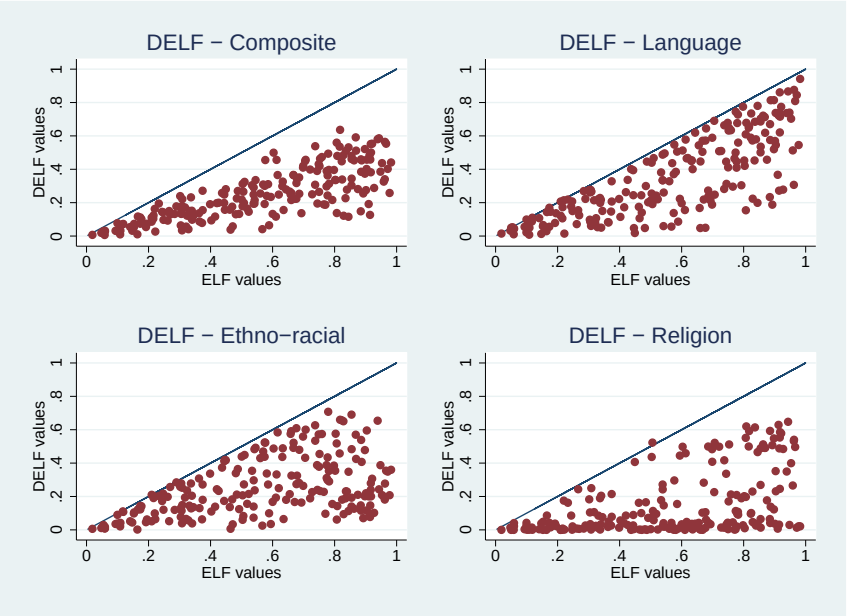


Figure 2: Combined and single characteristic $DELF$ values against ELF values.

to a higher affinity between the groups and, in turn, lower the $DELF$ values. *Table 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.

Index	Observations	Mean	Std. Dev.	Min.	Max.
ELF	210	0.563	0.270	0.019	0.982
$DELF$	210	0.252	0.157	0.006	0.636
$DELF_L$	210	0.353	0.243	0.008	0.942
$DELF_E$	210	0.255	0.176	0.002	0.708
$DELF_R$	210	0.148	0.188	0.000	0.648

Table 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 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.

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

Table 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 18* of *Appendix C*. 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.[64] 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.

6.2 Similarity measure between countries

To date, most authors have focused on the assessment of ethnicity within a country, as has this article. 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 7*, and is defined through:

$$DELF_{ij} = 1 - \sum_{k=1}^{K} \sum_{m=1}^{M} p_{ik}\ p_{jm}\ \hat{s}_{km} \tag{10}$$

where country i hosts groups $k = 1, ..., K$, and country j groups $m = 1, ..., M$, 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.[65] Due

[64] From *Figure 10* of *Appendix A.3*, one can see that the adjustments will tend to be more significant for higher values of ELF than for lower ones, where both indices are much closer. This is clearly visible for the higher ELF values at the top of *Table 18*.

[65] This significantly exceeds the 2,809 dyadic relations offered by de Groot (2009) for the 53

to the amount of country-pairs, only a discussion of averages and some tuples with the highest discrepancy is offered here.[66] Naturally, all $DELF$ values are much higher than those for individual countries. *Table 19* of *Appendix C* lists the mutually most similar and dissimilar countries at the single country level.[67] 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 not find many shared ethnic characteristics between these countries and all their values are close to one.

A regional aggregation also offers come interesting insights. For the calculation of the regional averages, the $DELF$ values between countries are adjusted for the different population sizes of the respective country pairs.[68] *Table 10* summarizes the regional and global averages.

	Regional $DELF$	Country pairs
Asia	0.719	1,600
Eastern Europe	0.479	841
Latin America	0.340	1,444
North Africa & Middle East	0.430	441
Sub-Saharan Africa	0.643	2,401
Western Countries	0.572	1,089
World	0.841	44,100

Table 10: $DELF$ average by main geographical regions

The global cultural diversity measured by the $DELF$ 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.[69] Latin America, in contrast, displays the least interregional

African countries.

[66]The complete data set can be received upon request.

[67]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 completly 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 B.6*. I would like to thank Walter Zucchini for this important comment.

[68]For the weighting, population data averages for 2005–2010 from the *World Development Indicators* World Bank (2011) were used. For more details on how regional averages are calculated and the differences in the calculation of $DELF$ values between countries, see *Appendix B.7*.

[69]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

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.

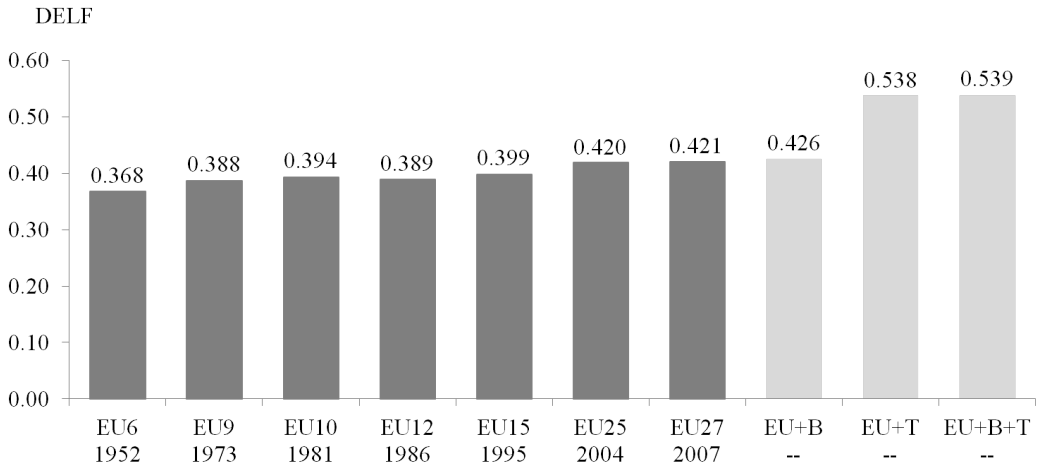


Figure 3: Average $DELF$ values of the EU per enlargement wave

Figure 3 shows the diversity level of the EU for each wave of enlargement.[70] 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 $DELF$ 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 $DELF$ 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.[71] Looking at potential future enlargements, the

more 'endogenously' than the method used in SSA by the colonial powers.

[70] For more details on the different waves of enlargement in the EU, and the respective diversity levels, see *Table 20* of *Appendix C*.

[71] One important caveat applies for this. As Kolo (2011) 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.

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 11*).[72] 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.[73] This comparison underlines that the genetic distance data is hardly a good proxy for the 'cultural' differences between countries.

	$DELF$	$DELF_L$	$DELF_E$	$DELF_R$
$DELF$	1			
	43890			
$DELF_L$	0.566	1		
	43890	*43890*		
$DELF_E$	0.489	0.636	1	
	43890	*43890*	*43890*	
$DELF_R$	0.899	0.363	0.193	1
	43890	*43890*	*43890*	*43890*
Genetic Distance	0.245	0.484	0.697	0.018
	30800	*30800*	*30800*	*30800*

Table 11: Rank correlations between $DELF$, its sub-indices and genetic distance data (observations in italics)

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 changes over time.

[72] 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 article.

[73] 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.

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.[74] 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 article. 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.

[74] 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.

References

Ahlerup, P. and Olsson, O. (2007). The Roots of Ethnic Diversity. Working Papers in Economics 281, Department of Economics, Göteborg University.

Akdede, S. H. (2010). Do more Ethnically and Religiously Diverse Countries have Lower Democratization? *Economics Letters* 106: 101–104.

Alesina, A., Baqir, R. and Easterly, W. (1999). Public Goods and Ethnic Divisions. *The Quarterly Journal of Economics* 114: 1243–1284.

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S. and Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth* 8: 155–194.

Alesina, A. and La Ferrara, E. (2000). Participation in Heterogeneous Communities. *The Quarterly Journal of Economics* 115: 847–904.

Alesina, A. and Zhuravskaya, E. (2011). Segregation and the Quality of Government in a Cross-Section of Countries. *The American Economic Review* 101: 1872–1911.

Annett, A. (2001). Social Fractionalization, Political Instability, and the Size of Government. *IMF Staff Papers* 48: 561–592.

Ashraf, Q. and Galor, O. (2011). The "Out of Africa" Hypothesis, Human Genetic Diversity, and Comparative Economic Development. NBER Working Papers 17216, National Bureau of Economic Research (NBER).

Barrett, D. B., Kurian, G. T. and Johnson, T. M. (2001). *World Christian Encyclopedia; A Comparative Survey of Churches and Religions in the Modern World*. Oxford: Oxford University Press, 2nd ed.

Barro, R. J. (1999). Determinants of Democracy. *Journal of Political Economy* 107: 158–183.

Barro, R. J. and McCleary, R. M. (2003). Religion and Economic Growth. NBER Working Papers 9682, National Bureau of Economic Research (NBER).

Bjørnskov, C. (2008). Social Trust and Fractionalization: A Possible Reinterpretation. *European Sociological Review* 24: 271–283.

Blattman, C. and Miguel, E. (2010). Civil War. *Journal of Economic Literature* 48: 3–57.

Booysen, F. (2002). An Overview and Evaluation of Composite Indices of Development. *Social Indicators Research* 59: 115–151.

Bossert, W., D'Ambrosio, C. and La Ferrara, E. (2011). A Generalized Index of Fractionalization. *Economica* 78: 723–750.

Bossert, W., Pattanaik, P. K. and Xu, Y. (2003). Similarity of Options and the Measurement of Diversity. *Journal of Theoretical Politics* 15: 405–421.

Branisa, B., Klasen, S. and Ziegler, M. (2009). The Construction of the Social Institutions and Gender Index (SIGI). Discussion Papers 184, Ibero America Institute for Economic Research (IAI), Georg-August-Universität Göttingen.

Brown, G. K. and Langer, A. (2010). Conceptualizing and Measuring Ethnicity. *Oxford Development Studies* 38: 411–436.

Bruk, S. I. (1964). *Atlas Narodov Mira*. Moskva: N. N. Miklucho-Maklaja, Institut Etnografii Imeni.

Cahill, M. B. (2005). Is the Human Development Index Redundant? *Social Indicators Research* 31: 1–5.

Caselli, F. and Coleman, W. J. (2008). On the Theory of Ethnic Conflict. CEDI Discussion Paper Series 08-08, Centre for Economic Development and Institutions, Brunel University West London.

Cavalli-Sforza, L. L. and Feldmann, M. W. (1981). *Cultural Transmission and Evolution; A Quantitative Approach*. Princeton, NJ: Princeton University Press.

Cavalli-Sforza, L. L., Menozzi, P. and Piazza, A. (1993). *The History and Geography of Human Genes*. Princeton, NJ: Princeton University Press.

Cavalli-Sforza, L. L., Piazza, A., Menozzi, P. and Mountain, J. (1988). Reconstruction of Human Evolution: Bringing Together Genetic, Archaeological, and Linguistic Data. *Proceedings of the National Academy of Sciences of the United States of America* 85: 6002–6006.

Chandra, K. and Wilkinson, S. (2008). Measuring the Effect of 'Ethnicity'. *Comparative Political Studies* 41: 515–563.

Cho, S.-Y., Dreher, A. and Neumayer, E. (2011). The Spread of Anti-Trafficking Policies - Evidence from a New Index. CESifo Working Paper Series 3376, CESifo Group Munich.

Chong, A., Guillen, J. and Rios, V. (2010). Language Nuances, Trust and Economic Growth. *Public Choice* 143: 191–208.

Chowdhury, S. and Squire, L. (2006). Setting Weights for Aggregate Indices: An Application to the Commitment to Development Index and Human Development Index. *The Journal of Development Studies* 42: 761–771.

CIA (2011). The World Factbook. `https://www.cia.gov/library/publications/the-world-factbook/`.

Clévenot, M. (1987). *L' Etat des Religions dans le Monde*. Paris: La Découverte.

Collier, P. (1998). The Political Economy of Ethnicity. Working Paper Series 98-8, Centre for the Study of African Economies, University of Oxford.

Collier, P. and Hoeffler, A. (1998). On Economic Causes of Civil War. *Oxford Economic Papers* 50: 563–573.

Collier, P. and Hoeffler, A. (2002). On the Incidence of Civil War in Africa. *The Journal of Conflict Resolution* 46: 13–28.

Collier, P. and Hoeffler, A. (2004). Greed and Grievance in Civil War. *Oxford Economic Papers* 56: 563–595.

Collier, P., Hoeffler, A. and Rohner, D. (2009). Beyond Greed and Grievance: Feasibility and Civil War. *Oxford Economic Papers* 61: 1–27.

Collier, P., Hoeffler, A. and Söderbom, M. (2004). On the Duration of Civil War. *Journal of Peace Research* 41: 253–273.

Dalby, D. and Williams, C. (1999). *The Linguasphere Register of the World's Languages and Speech Communities*. Hebron, Wales: Linguasphere Press.

Desmet, K., Le Breton, M., Ortín, I. Ortuño and Weber, S. (2011). The Stability and Breakup of Nations: A Quantitative Analysis. *Journal of Economic Growth* 16: 183–213.

Desmet, K., Ortín, I. Ortuño and Wacziarg, R. (2012). The Political Economy of Ethnolinguistic Cleavages. *Journal of Development Economics* 97: 322–332.

Desmet, K., Ortín, I. Ortuño and Weber, S. (2009). Linguistic Diversity and Redistribution. *Journal of the European Economic Association* 7: 1291–1318.

Easterly, W. and Levine, R. (1997). Africa's Growth Tragedy: Policies and Ethnic Divisions. *The Quarterly Journal of Economics* 112: 1203–1250.

Encyclopædia Britannica (ed.) (2007). *The New Encyclopædia Britannica*. Chicago [u.a.]: Encyclopædia Britannica Inc., 15th ed.

Esteban, J. and Mayoral, L. (2011). Ethnic and Religious Polarization and Social Conflict. Working paper 857.11, Unitat de Fonaments de l'Anàlisi Econòmica (UAB) and Institut d'Anàlisi Econòmica (CSIC).

Esteban, J., Mayoral, L. and Ray, D. (2010). Ethnicity and Conflict: An Empirical Study. Working Papers 840.10, Unitat de Fonaments de l'Anàlisi Econòmica (UAB) and Institut d'Anàlisi Econòmica (CSIC).

Esteban, J. and Ray, D. (1994). On the Measurement of Polarization. *Econometrica* 62: 819–851.

Esteban, J. and Ray, D. (2008). Polarization, Fractionalization and Conflict. *Journal of Peace Research* 45: 163–182.

Esteban, J. and Ray, D. (2011). Linking Conflict to Inequality and Polarization. *American Economic Review* 101: 1345–1374.

Falck, O., Heblich, S., Lameli, A. and Suedekum, J. (2010). Dialects, Cultural Identity, and Economic Exchange. IZA Discussion Papers 4743, Institute for the Study of Labor (IZA).

Fearon, J. D. (2003). Ethnic and Cultural Diversity by Country. *Journal of Economic Growth* 8: 195–222.

Fearon, J. D. and Laitin, D. D. (2003). Ethnicity, Insurgency, and Civil War. *The American Political Science Review* 97: 75–90.

Felbermayr, G. J. and Toubal, F. (2010). Cultural Proximity and Trade. *European Economic Review* 54: 279–293.

Filmer, D. and Pritchett, L. (2001). Estimating Wealth Effects Without Expenditure Data – Or Tears: An Application To Educational Enrollments In States Of India. *Demography* 38: 115–132.

Garcia-Montalvo, J. and Reynal-Querol, M. (2002). Why Ethnic Fractionalization? Polarization, Ethnic Conflict and Growth. Economics Working Papers 660, Department of Economics and Business, Universitat Pompeu Fabra.

Garcia-Montalvo, J. and Reynal-Querol, M. (2003). Religious Polarization and Economic Development. *Economics Letters* 80: 201–210.

Garcia-Montalvo, J. and Reynal-Querol, M. (2005). Ethnic Diversity and Economic Development. *Journal of Development Economics* 76: 293–323.

Garcia-Montalvo, J. and Reynal-Querol, M. (2008). Discrete Polarisation with an Application to the Determinants of Genocides. *The Economic Journal* 118: 1835–1865.

Garcia-Montalvo, J. and Reynal-Querol, M. (2010). Ethnic Polarization and the Duration of Civil Wars. *Economics of Governance* 11: 123–143.

Ginsburgh, V. (2005). Languages, Genes, and Cultures. *Journal of Cultural Economics* 29: 1–17.

Ginsburgh, V. and Weber, S. (2011). *How Many Languages Do We Need? The Economics of Linguistic Diversity*. Princeton, NJ: Princeton University Press.

Giuliano, P., Spilimbergo, A. and Tonon, G. (2006). Genetic, Cultural and Geographical Distances. IZA Discussion Papers 2229, Institute for the Study of Labor (IZA).

Greenberg, J. H. (1956). The Measurement of Linguistic Diversity. *Language* 32: 109–115.

Groot, O. J. de (2009). Measuring Ethno-Linguistic Affinity between Nations. Discussion Papers 921, DIW Berlin, German Institute for Economic Research.

Guiso, L., Sapienza, P. and Zingales, L. (2009). Cultural Biases in Economic Exchange? *The Quarterly Journal of Economics* 124: 1095–1131.

Haq, M. (2006). The Birth of the Human Development Index. In Fukuda-Parr, S. and Shiva Kumar, A. K. (eds), *Readings in Human Development*. Oxford: Oxford University Press, 127–137.

Hofstede, G. H. (2000). *Culture's Consequences: International Differences in Work-related Values*, Cross-cultural Research and Methodology Series *5*. Beverly Hills, Cal.: Sage Publications.

Johnson, T. M. (2010). World Christian Database. `http://www.worldchristiandatabase.org/wcd/`.

Jolliffe, I. T. (1973). Discarding Variables in a Principal Component Analysis. II: Real Data. *Applied Statistics* 22: 21–31.

Joshua Project (2007). Joshua Project: Bringing Definition to the Unfinished Task. `http://www.joshuaproject.net`.

Kolenikov, S. and Angeles, G. (2009). Socioeconomic Status Measurement With Discrete Proxy Variables: Is Principal Component Analysis A Reliable Answer? *Review of Income and Wealth* 55: 128–165.

Kolo, P. (2011). Questioning Ethnic Fragmentation's Exogeneity – Drivers of an Endogenous Formation. Discussion Papers 210, Ibero America Institute for Economic Research (IAI), Georg-August-Universität Göttingen.

La Porta, R., Silanes, F. Lopez-de, Shleifer, A. and Vishny, R. (1999). The Quality of Government. *Journal of Law, Economics, and Organization* 15: 222–279.

Lewis, M. P. (2009). *Ethnologue; Languages of the World*. Dallas, Tex.: Summer Institute of Linguistics (SIL), 16th ed.

Lind, J. T. (2007). Fractionalization and Inter-Group Differences. *Kyklos* 60: 123–139.

Loh, J. and Harmon, D. (2005). A Global Index of Biocultural Diversity. *Ecological Indicators* 5: 231–241.

Mauro, P. (1995). Corruption and Growth. *The Quarterly Journal of Economics* 110: 681–712.

McGillivray, M. and Noorbakhsh, F. (2004). Composite Indices of Human Well-being: Past, Present, and Future. UNU-WIDER Research Paper 63, World Institute for Development Economic Research (UNU-WIDER).

McGillivray, M. and White, H. (1993). Measuring Development? The UNDP's Human Development Index. *Journal of International Development* 5: 183–192.

Munda, G. and Nardo, M. (2005). Constructing Consistent Composite Indicators: The Issue of Weights. Technical Report EUR 21834 EN, European Commission.

Nardo, M., Saisana, M., Saltelli, S., Tarantola, A., Hoffman, A. and Giovannini, E. (2005). Handbook on Constructing Composite Indicators: Methodology and User Guide. Technical Report 2005/3, OECD.

Nehring, K. and Puppe, C. (2002). A Theory of Diversity. *Econometrica* 70: 1155–1198.

Nguefack-Tsague, G., Klasen, S. and Zucchini, W. (2011). On Weighting the Components of the Human Development Index: A Statistical Justification. *Journal of Human Development and Capabilities* 12: 183–202.

Noorbakhsh, F. (1998). The Human Development Index: Some Technical Issues and Alternative Indices. *Journal of International Development* 10: 589–605.

Ogwang, T. (1994). The Choice of Principle Variables for Computing the Human Development Index. *World Development* 22: 2011–2014.

Ogwang, T. and Abdou, A. (2003). The Choice of Principal Variables for Computing some Measures of Human Well-being. *Social Indicators Research* 64: 139–152.

Okediji, T. O. (2005). The Dynamics of Ethnic Fragmentation. *American Journal of Economics and Sociology* 64: 637–662.

Posner, D. N. (2004). Measuring Ethnic Fractionalization in Africa. *American Journal of Political Science* 48: 849–863.

Ram, R. (1982). Composite Indices of Physical Quality of Life, Basic Needs Fulfilment, and Income: A 'Principal Component' Representation. *Journal of Development Economics* 11: 227–247.

Ramachandran, S., Deshapande, O., Roseman, C., Rosenberg, N., Feldmann, M. W. and Cavalli-Sforza, L. L. (2005). Support from the Relationship of Genetic and Geographic Distance in Human Populations for a Serial Founder Effect Originating in Africa. *Proceedings of the National Academy of Sciences* 102: 15942–15947.

Spolaore, E. and Wacziarg, R. (2009). The Diffusion of Development. *The Quarterly Journal of Economics* 124: 469–529.

Taylor, C. L. and Hudson, M. C. (1972). *World Handbook of Political and Social Indicators*. New Haven: Yale University Press.

Weidmann, N. B., Rød, J. K. and Cederman, L.-E. (2010). Representing Ethnic Groups in Space: A New Dataset. *Journal of Peace Research* 47: 491–499.

Weitzman, M. L. (1992). On Diversity. *The Quarterly Journal of Economics* 107: 363–405.

World Bank (2011). World Development Indicators 2011. `http://data.worldbank.org/data-catalog/world-development-indicators`.

A Data robustness and alternative data

A.1 Data robustness checks

Although the discussion in this chapter already showed the general strength of the *WCE* data, some additional robustness check shall be applied. Two new data sets are created that add some noise to the original data. If all three datasets do not differ in a significant way, it should be reasonable to use the original data. In doing so, one accepts errors in the range of the noise added to the original data set. The noise data is created by altering the original group size p_i to the new size $\tilde{p_i}$ with a normal distributed random variable in a way that:

$$\tilde{p_i} = p_i \cdot (1+e) \qquad \text{,with} \qquad e \sim N(0;\sigma) \tag{11}$$

For σ two different values are assumed; σ_1 uses the standard deviation of the group distribution over all observations, and is thus equal for all countries. In contrast, σ_2 uses a country specific standard deviation. The scatter plot of *Figure 4* shows $DELF$ values for both alternative data sets against the original data.

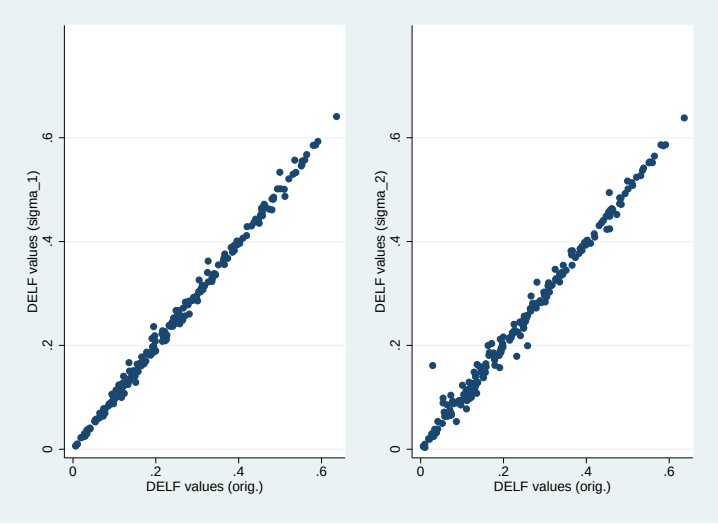


Figure 4: Original $DELF$ values against newly created random data sets

The Spearman rank correlation is over 0.99 for both data sets and confirms their high congruency. For the new data based on country specific variations, some small outliers are identifiable. These are rather homogeneous countries with a limited number of groups and a clear majority group. By construction, they have a much higher probability of being distant from the original data.

The granularity of the data, which is one of its major advantages, leads to a sizeable number of very small groups. The data quality, especially for these groups, might be debatable. Following Fearon (2003), a reduced data set is constructed excluding these very small groups.[75] Doing this reduces the number of groups from 12,432 down to 5,674. Excluding groups would either alter the group shares of all groups, because one would need to rescale them, or one can alternatively create new groups that differ from all existing groups. Subsequently the second approach is followed. Although the groups are small, they represent some part of the population that seems to be different from the rest. In some countries, that new group corresponds to a rather sizeable one. Thus, to not account for them at all would be incorrect. Combining them into one group lowers the potential individual data inaccuracies. Analogous to the figure above, *Figure 5* compares the $DELF$ values of the reduced data set against the full data.

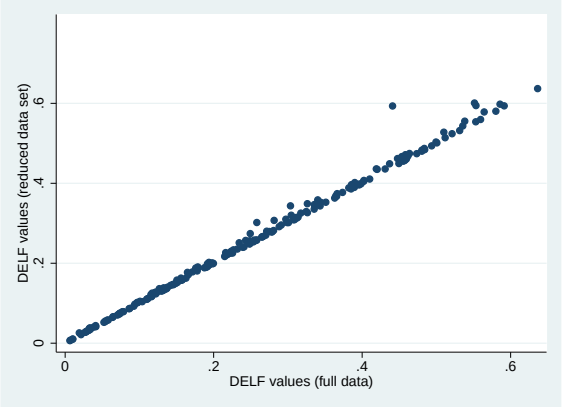


Figure 5: Original $DELF$ values against reduced data set

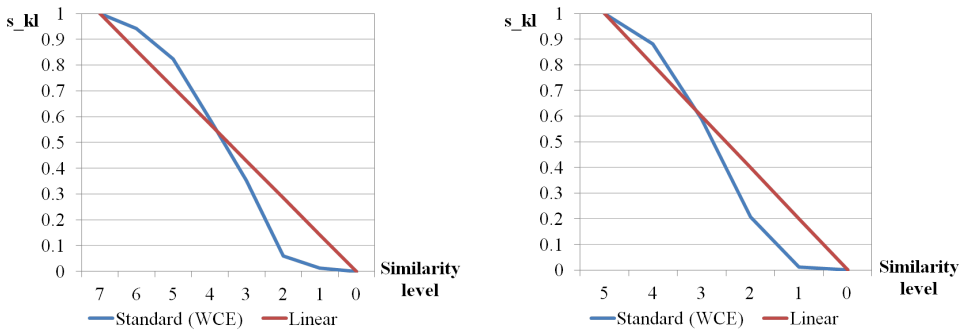
In this case, the most heterogeneous countries show an increased difference compared to the base data. Papua New Guinea is the most apparent outlier. Because Papua New Guinea has a huge number of small groups that are now combined into one group that differs completely from all other groups, it appears more diverse than when accounting for the mutual similarities of all the small groups. However, the similarity between both data sets is still very high.

[75] In contrast to Fearon (2003), who limits his ELF calculation to groups greater than 1%, here a lower threshold of 0.1% is used.

A.2 Alternative similarity values

The assignment of the similarity values according to the language classification is rather clear. Here, one can easily leverage the lexical congruency between two languages and transfer these similarity levels to the assigned $\bar{s}_{kl}$ values. When the $\bar{s}_{kl}$ were differently assigned to correspond directly with the similarity levels and the values of 1, 0.85, 0.80, 0.70, 0.50, 0.30 and 0.05 for $\bar{s}_{kl}$ were used, the overall results show only marginal changes. However, for single countries, some slightly larger adjustments in their rank order accrue.

For the ethno-racial classification, however, the congruency is more ordinal in nature. In the essay, the assigned $\bar{s}_{kl}$ follow the same decreasing slope as that of the language classification. Nevertheless, one could also argue in favor of a linear assignment of the $\bar{s}_{kl}$ values to mirror the ordinal similarity levels. For both classifications, both similarity slopes are pictured in *Figure 6*.



(a) Language classification

(b) Ethno-racial classification

Figure 6: Used similarity values $\bar{s}_{kl}$ vs. linear similarity levels

From the differences in the slopes, one can easily see that for both classifications, less distant groups are assigned higher $\bar{s}_{kl}$ values under the *WCE* method than under a linear assignment. For more distant groups, the opposite is the case. Countries with groups that speak more distant languages would exhibit lower *DELF* values in the *WCE* case than under a linear $\bar{s}_{kl}$ allocation. *Figure 7* contrasts the *DELF* values used in the essay with the corresponding values calculated with a linear scale for the language and the ethno-racial classification.

The impact differs between both characteristics. Whereas the Spearman rank correlation between both scales is again over 0.99 for the ethno-racial values, it is slightly less, at 0.94, for the language classification. The countries with the highest downward adjustments are Papua New Guinea, Solomon Islands, Senegal, Vanuatu, Northern Mariana Islands, Niger, Uganda, Nigeria, Switzerland and Sierra Leone. The country with a significant upward adjustment is

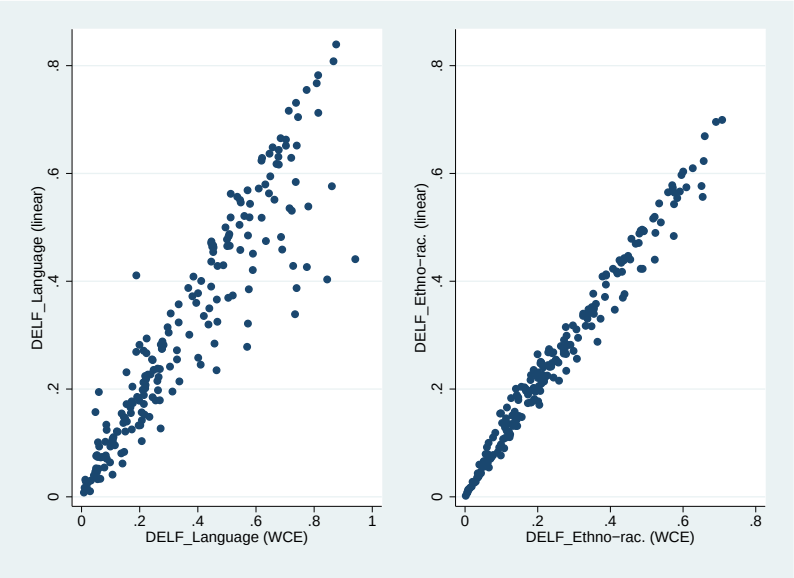


Figure 7: *DELF* values based on *WCE* similarity values against linear scale *DELF* values per characteristic

Trinidad and Tobago. Due to the high correlation values which remain, the results should not be significantly impacted.

Extending the discussion above, Fearon (2003) defines his measure of cultural diversity with:

$$r_{kl} = \left(\frac{n}{m}\right)^{\alpha} \tag{12}$$

where m are the highest number of classifications two groups may share and n the number they actual share. This naturally leads to linear similarity values. The parameter $\alpha \in [0, ..., 1]$ then influences the course of the similarity value function to give it a concave shape.[76] For the application here, this would translate into:

$$f(\hat{s}_{kl}) = (\hat{s}_{kl})^{\alpha} \tag{13}$$

The idea behind assigning such a function is that early divergence between two groups might signify more differences than small differences at a later stage. In other words, with a rising α, more severe differences are proportionally less important and small differences increase in importance. Desmet et al. (2012) assume that more severe splits (i.e., completely different languages) are more

[76] This is at least the range within which Fearon (2003) limits α. However, much larger values could still apply and for $\alpha = \infty$ any continuous distance measure fades and the indices merge with their dichotomous forms.

relevant for more drastic conflicts of interest (e.g., incidence of civil wars). More nuanced differences (i.e., different dialects), in contrast, affect the transaction costs of coordination for any economic activity and are relevant, for example, in explaining differences in economic growth. As a consequence, the choice of α might depend on the problem under scrutiny. The final selection of a value for α, however, remains completely arbitrary. Fearon (2003) uses a value of $\alpha = 0.5$, whereas Desmet et al. (2009) and Esteban et al. (2010) use a value of $\alpha = 0.05$.[77] *Figure 8* shows the courses of the applied similarity values for different concavity assumptions.

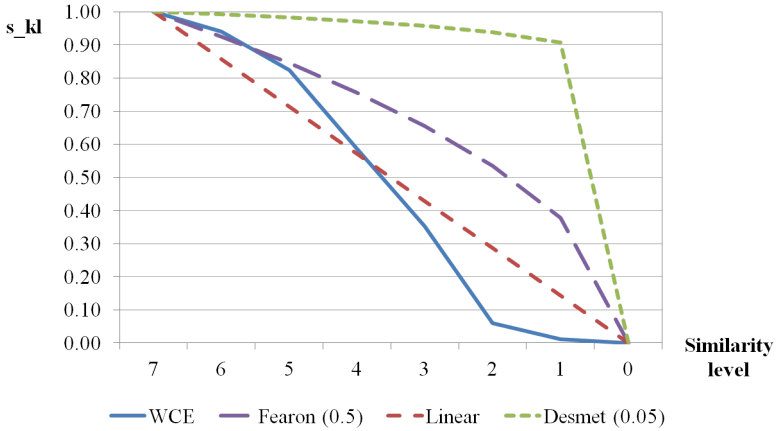


Figure 8: Similarity functions depending on different concavity assumptions

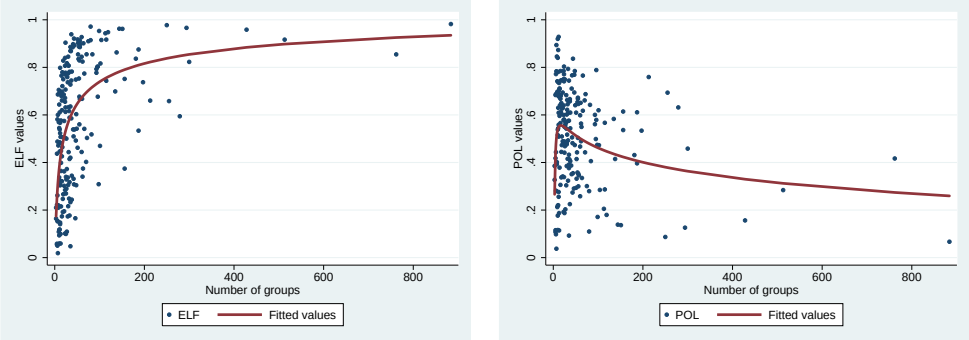
For the three highest similarity levels, the course for a linear similarity function with $\alpha = 0.5$ (Fearon, 2003) and assigned values of the *WCE* are quite comparable, yet somewhat distinct from the linear values. Thereafter, the *WCE* drops faster. With the assumption that $\alpha = 0.05$, the similarity between two groups stays very high for quite a while, dropping steeply afterwards. The latter thus assigns rather extreme (dis)similarity values, whereas the other functions are more continuous. As the *WCE* similarity classification has an inbuilt non-linearity of similarity measures, assigning values of α is less important here than it is for Fearon (2003), Desmet et al. (2009) and Esteban et al. (2010). In addition, as the similarity values assigned by Barrett et al. (2001) in the *WCE* seem to be more grounded in the real difficulties between two individuals to communicate, this essay refrains from assigning an arbitrary value to α.[78]

[77] Indeed, Desmet et al. (2009) vary the values of α and conclude that these low levels show the best performance in their analysis of redistribution.

[78] Nevertheless, $DELF$ and $D-POL$ values with the commonly used values of 0.05 and 0.5 for α, may be obtained from the author.

A.3 Characteristics of different ethnicity measures depending on the number of groups

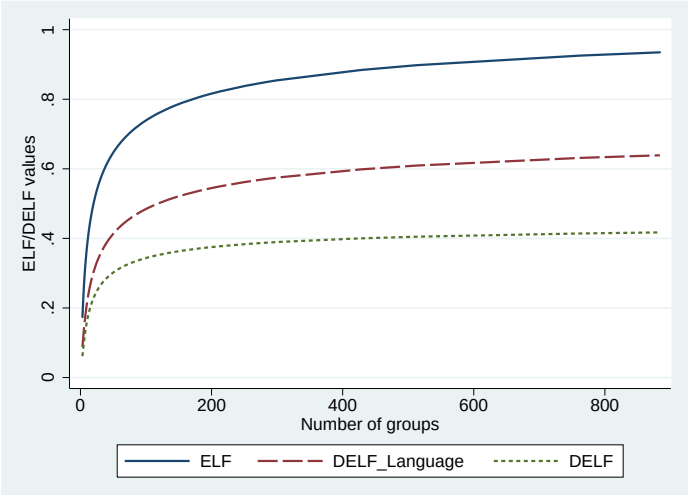
Figure 9: ELF and POL values against number of groups for 210 countries based on *WCE* data



(a) ELF values

(b) Polarization (POL) values

Figure 10: Fitted ELF and *DELF* values against number of groups for all 210 countries



B Details on similarity calculations, weighting and its implication for the interpretation of results

B.1 Similarity matrix calculations

Groups, the integral component of all ELF, POL and $DELF$ calculations, can generally be defined for each single characteristic or by all three at the same time.[79] For all 210 countries, he WCE data consists of 11,657 groups defined by language, 4,625 groups defined by culture, and 883 groups defined by religion. If the groups are defined by all three characteristics at the same time, more groups can emerge as characteristics might be combined. The definition of a group is the most granular possible, i.e., along all three characteristics, and results in 12,432 groups in the data set used. This means that any two groups differ slightly in at least one of the characteristics.

The following example shall illustrate the calculation of the similarity values per characteristic and the combination to arrive at the composite $DELF$ values. The exemplary country consists of three groups. Thus, for the example, it follows that the number of K groups within this country is given by $K = \{A;B;C\}$. There exist two languages, L1 and L2, two ethno-racial groups, E1 and E2, and only one religion, R1. Combining these characteristics results in three groups with the specifications below:

Group	Language	Ethno-racial	Religion
A	L1	E1	R1
B	L2	E1	R1
C	L2	E2	R1

Table 12: Specifications of characteristics per group

For each characteristic, language, ethno-racial, and religion, similarity values ($\bar{s}^L_{kl}$, $\bar{s}^E_{kl}$, and $\bar{s}^R_{kl}$), with $k,l \in K = \{A;B;C\}$ between two groups can be assigned. Based on these specifications, one can calculate a $DELF$ value for each of the characteristics:

$$DELF_L = 1 - \sum_{k \in K} \sum_{l \in K} p_k p_l \bar{s}^L_{kl} \tag{14}$$

$$DELF_E = 1 - \sum_{k \in K} \sum_{l \in K} p_k p_l \bar{s}^E_{kl} \tag{15}$$

[79] Naturally, one could also combine any two of the characteristics if such a combination was recommended for the research problem at hand.

$$DELF_R = 1 - \sum_{k \in K} \sum_{l \in K} p_k p_l \bar{s}^R_{kl} \tag{16}$$

with $k,l \in \{A;B;C\}$ and p_k and p_l the relative group sizes. To arrive at the similarity values, one can set up a similarity matrix for each characteristic. For the above example, these matrices are shown in *Table 13*.

(a)

	A	B	C
A	$\bar{s}^L_{AA}$	$\bar{s}^L_{AB}$	$\bar{s}^L_{AC}$
B	$\bar{s}^L_{BA}$	$\bar{s}^L_{BB}$	$\bar{s}^L_{BC}$
C	$\bar{s}^L_{CA}$	$\bar{s}^L_{CB}$	$\bar{s}^L_{CC}$

(b)

	A	B	C
A	$\bar{s}^E_{AA}$	$\bar{s}^E_{AB}$	$\bar{s}^E_{AC}$
B	$\bar{s}^E_{BA}$	$\bar{s}^E_{BB}$	$\bar{s}^E_{BC}$
C	$\bar{s}^E_{CA}$	$\bar{s}^E_{CB}$	$\bar{s}^E_{CC}$

(c)

	A	B	C
A	$\bar{s}^R_{AA}$	$\bar{s}^R_{AB}$	$\bar{s}^R_{AC}$
B	$\bar{s}^R_{BA}$	$\bar{s}^R_{BB}$	$\bar{s}^R_{BC}$
C	$\bar{s}^R_{CA}$	$\bar{s}^R_{CB}$	$\bar{s}^R_{CC}$

Table 13: Exemplary similarity matrices for the three groups (a) with mutual language $\bar{s}^L_{kl}$ values, (b) with mutual ethno-racial $\bar{s}^E_{kl}$ values and (c) with mutual religion $\bar{s}^R_{kl}$ values

The assumptions that $\bar{s}_{kk} = 1$ and $\bar{s}_{kl} = \bar{s}_{lk}$ for all $k,l \in \{A;B;C\}$ hold, and for all groups that belong to one language or ethno-racial group, a respective similarity value of one is assigned. In the case of the religious classification, all belong to one religion, i.e., one group. Based on the characteristic definitions in *Table 12*, it follows that $\bar{s}^E_{AC} = \bar{s}^E_{BC} = \bar{s}^E_{CA} = \bar{s}^E_{CB}$. The distance is labeled in the following simplified $\bar{s}^E$. This analogously holds for the language similarity values. The matrices of *Table 13* can be further defined with:

(a)

	A	B	C
A	1	$\bar{s}^L$	$\bar{s}^L$
B	$\bar{s}^L$	1	1
C	$\bar{s}^L$	1	1

(b)

	A	B	C
A	1	1	$\bar{s}^E$
B	1	1	$\bar{s}^E$
C	$\bar{s}^E$	$\bar{s}^E$	1

(c)

	A	B	C
A	1	1	1
B	1	1	1
C	1	1	1

Table 14: Similarity matrices for the three groups, taking into account the specifications of their (a) language, (b) ethno-racial and (c) religious characteristic

With the relative group sizes p_A, p_B and p_C, one obtains an exemplary $DELF_E$ index for the ethno-racial characteristic:

$$\begin{aligned} DELF_E &= 1 - \sum_{k \in K} \sum_{k \in K} p_k p_l \bar{s}^E_{kl} = \\ &= 1 - (p_A \cdot p_A \cdot 1 + p_A \cdot p_B \cdot 1 + p_A \cdot p_C \cdot \bar{s}^E + \\ &\quad + p_B \cdot p_A \cdot 1 + p_B \cdot p_B \cdot 1 + p_B \cdot p_C \cdot \bar{s}^E + \\ &\quad + p_C \cdot p_A \cdot \bar{s}^E + p_C \cdot p_B \cdot \bar{s}^E + p_C \cdot p_C \cdot 1) = \end{aligned}$$

$$= \quad 1 - \Big((p_A + p_B)^2 \cdot 1 + 2 \cdot (p_A + p_B) \cdot p_C \cdot \bar{s}^E + p_C^2 \cdot 1 \Big)$$

One can clearly see that for the single characteristics $DELF$, the respective most granular split per characteristic is decisive. The group definition at a more detailed level does not add additional information. In the above example, this would lead to a reduced 2×2 matrix of the one found in *Table 13(b)* with one group $(A + B)$, and the remaining group C with the respective relative group sizes $(p_A + p_B)$ and p_C.[80] However, for the composite $DELF$, combining all three characteristics into a composite similarity measure $\hat{s}_{kl}$ is key. The general matrix for the composite $DELF$ calculation is then given in *Table 15*.

	A	B	C
A	$\hat{s}_{AA}$	$\hat{s}_{AB}$	$\hat{s}_{AC}$
B	$\hat{s}_{BA}$	$\hat{s}_{BB}$	$\hat{s}_{BC}$
C	$\hat{s}_{CA}$	$\hat{s}_{CB}$	$\hat{s}_{CC}$

Table 15: Similarity matrix for composite $DELF$ calculation

The calculation of the $\hat{s}_{kl}$ depends on the mode of weighting and combining the three characteristics. The averaging of the characteristics has important implications for the interpretation of the resulting $DELF$ values.[81] Extending the discussions in section 5, especially their mathematical attributes, is discussed in the following. In contrast to the exemplary case used here to demonstrate the similarity calculation, the following discussions apply to the general case.

B.2 Arithmetic mean

In the case of an arithmetic mean, as discussed in section 5, the composite $DELF$ value is calculated as:

$$DELF = 1 - \sum_{k=1}^{K} \sum_{l=1}^{K} p_k p_l \hat{s}_{kl} \tag{17}$$

[80] This is equivalent to the discussion in section 2. Only perfectly similar individuals are grouped together and groups are meant to emerge 'endogenously'. Here, two identical groups merge into one group.

[81] All approaches portrayed here share a common, implicit assumption. They all assume that a combination follows the same pattern, independent of the specific combination of the single characteristics, and that the combination is equivalent in all countries. This assumption is further discussed in *Appendix B.5*.

with

$$\hat{s}_{kl} = \frac{1}{3} \left[\bar{s}^L_{kl} + \bar{s}^E_{kl} + \bar{s}^R_{kl} \right] \tag{18}$$

where $\bar{s}^L_{kl}$, $\bar{s}^E_{kl}$ and $\bar{s}^R_{kl}$ for all $k,l \in K$ are again the respective similarity values for the language, ethno-racial and religious classification. In the general case, K is again the total number of groups in the given country. For the specifications of the above example, the matrix in *Table 15* transforms, with $k,l \in K = \{A;B;C\}$, to *Table 16*.

	A	B	C
A	$\frac{1}{3}(1+1+1)$	$\frac{1}{3}(1+\bar{s}^L+1)$	$\frac{1}{3}(\bar{s}^E+\bar{s}^L+1)$
B	$\frac{1}{3}(1+\bar{s}^L+1)$	$\frac{1}{3}(1+1+1)$	$\frac{1}{3}(\bar{s}^E+1+1)$
C	$\frac{1}{3}(\bar{s}^E+\bar{s}^L+1)$	$\frac{1}{3}(\bar{s}^E+1+1)$	$\frac{1}{3}(1+1+1)$

Table 16: Similarity matrix for the exemplary $DELF$ calculation

For the arithmetic mean, there exists an identity between the calculation of the composite similarity value $\hat{s}_{kl}$, as in *Equation (18)*, to arrive at the composite $DELF$ values and the arithmetic mean of the single $DELF$ indices. With *Equations (14)–(16)*, for the general case, one obtains :

$$\begin{aligned}
DELF &= \frac{1}{3}(DELF_L + DELF_E + DELF_R) = \\
&= \frac{1}{3}\left[\left(1 - \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \bar{s}^L_{kl}\right) + \left(1 - \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \bar{s}^E_{kl}\right) + \left(1 - \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \bar{s}^R_{kl}\right)\right] = \\
&= \frac{1}{3}\left[3 - \left(\sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \bar{s}^L_{kl} + \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \bar{s}^E_{kl} + \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \bar{s}^R_{kl})\right)\right] = \\
&= 1 - \frac{1}{3}\left[\sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \bar{s}^L_{kl} + \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \bar{s}^E_{kl} + \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \bar{s}^R_{kl}\right] = \\
&= 1 - \frac{1}{3}\left[\sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \left(\bar{s}^L_{kl} + \bar{s}^E_{kl} + \bar{s}^R_{kl}\right)\right] = \\
&= 1 - \left[\sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \frac{1}{3}\left(\bar{s}^L_{kl} + \bar{s}^E_{kl} + \bar{s}^R_{kl}\right)\right] = \\
&= 1 - \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \hat{s}_{kl} = \\
&= DELF
\end{aligned}$$

Thus, in the case of the arithmetic mean, there is no difference between the

$DELF$ calculation following *Equations (17)* and *(18)*, and an arithmetic mean over the single $DELF$ values. The arithmetic mean is therefore the most practical way of combining the single indices. Besides the arguments discussed in section 5, this is one of the main reasons why this approach is used.

B.3 Geometric mean and partly compensating methods

In the case of the geometric mean, there is no complementarity between the three characteristics. If two groups differ completely in one characteristic, which is quite often the case for religion, they are also classified to be completely different overall. For the geometric mean, the $\hat{s}_{kl}$ calculation follows:

$$\hat{s}_{kl}^{Geo} = \left[\ \bar{s}_{kl}^{L} \ \cdot \ \bar{s}_{kl}^{E} \ \cdot \ \bar{s}_{kl}^{R} \ \right]^{\frac{1}{3}} \tag{19}$$

Although the calculation of $\hat{s}_{kl}^{Geo}$ is not much more difficult than the standard $\hat{s}_{kl}$, it implies a further limitation. In contrast to the arithmetic mean, where one finds equality in the calculation of the $\hat{s}_{kl}$ and averaging the single $DELF$ values, this is not possible for the geometric mean. Relying on *Equation (19)*, one obtains:

$$\begin{aligned} DELF_{Geo} &= 1 - \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \hat{s}_{kl}^{Geo} \\ &= 1 - \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \ \left(\ \bar{s}_{kl}^{L} \ \bar{s}_{kl}^{E} \ \bar{s}_{kl}^{R} \ \right)^{\frac{1}{3}} \\ &= 1 - \left[\sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l (\bar{s}_{kl}^{L})^{\frac{1}{3}} \cdot \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l (\bar{s}_{kl}^{E})^{\frac{1}{3}} \cdot \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l (\bar{s}_{kl}^{R})^{\frac{1}{3}} \right] \end{aligned} \tag{20}$$

In contrast, calculating the geometric average of the single indices under the consideration of *Equations (14)–(16)* leads to:

$$\begin{aligned} DELF_{Geo2} &= (DELF_L \cdot DELF_E \cdot DELF_R)^{\frac{1}{3}} \\ &= (DELF_L)^{\frac{1}{3}} \cdot (DELF_E)^{\frac{1}{3}} \cdot (DELF_R)^{\frac{1}{3}} \\ &= \left(1 - \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \bar{s}_{kl}^{L} \right)^{\frac{1}{3}} \left(1 - \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \bar{s}_{kl}^{E} \right)^{\frac{1}{3}} \left(1 - \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \bar{s}_{kl}^{R} \right)^{\frac{1}{3}} \end{aligned} \tag{21}$$

That *Equations (20)* and *(21)* are not equivalent is straightforward to see.

Between the geometric mean, which does not mirror the complementarity of the characteristics at all, and the arithmetic mean, which does reflect this,

Branisa et al. (2009) suggest a third alternative. They square the components before the calculation of the arithmetic mean. This leads to an adjusted $\hat{s}_{kl}$ with:

$$\hat{s}_{kl}^{Pc} = \frac{1}{3} \left[\ (\bar{s}_{kl}^{L})^2 \ + \ (\bar{s}_{kl}^{E})^2 \ + \ (\bar{s}_{kl}^{R})^2 \ \right] \tag{22}$$

Analogously one obtains:

$$\begin{aligned}
DELF_{Pc} &= 1 - \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \hat{s}_{kl}^{Pc} = \\
&= 1 - \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \ \left(\frac{1}{3} \left[\ (\bar{s}_{kl}^{L})^2 \ + \ (\bar{s}_{kl}^{E})^2 \ + \ (\bar{s}_{kl}^{R})^2 \ \right] \right) = \\
&= 1 - \frac{1}{3} \left[\sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \ (\bar{s}_{kl}^{L})^2 \ + \ \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \ (\bar{s}_{kl}^{E})^2 \ + \ \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \ (\bar{s}_{kl}^{R})^2 \right] = \\
&= \frac{1}{3} \left[3 - \left(\sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \ (\bar{s}_{kl}^{L})^2 \ + \ \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \ (\bar{s}_{kl}^{E})^2 \ + \ \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l \ (\bar{s}_{kl}^{R})^2 \right) \right] = \\
&= \frac{1}{3} \left[\underbrace{(1 - \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l (\bar{s}_{kl}^{L})^2)}_{\neq (DELF_L)^2} + \underbrace{(1 - \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l (\bar{s}_{kl}^{E})^2)}_{\neq (DELF_E)^2} + \underbrace{(1 - \sum_{k=1}^{K}\sum_{l=1}^{K} p_k p_l (\bar{s}_{kl}^{R})^2)}_{\neq (DELF_R)^2} \right]
\end{aligned} \tag{23}$$

As is the case with the geometric mean, one first needs to calculate the respective composite $\hat{s}_{kl}$ values on the most granular group setting and then follow *Equation (17)* to arrive at the composite $DELF$ values. *Figure 11* shows a matrix scatter plot of the different weighting schemes. Their high correlation is again confirmed by the scatter outline:

B.4 Principal component analysis

Principal component analysis (PCA) is becoming a more and more utilized approach to assess weights, not on theoretical grounds, but based on the data itself. Whenever one deals with continuous data, the PCA approach is indeed a promising one. Bossert et al. (2011) also use this approach to calculate the composite GELF values for different diversity characteristics in the US. However, they also used predominantly continuous data like income, for example. For categorical data, the PCA is much more difficult to apply (Kolenikov and Angeles, 2009). For a PCA calculation, the data need to be in a number format and not in categories. A possible solution for this is to turn the categories in

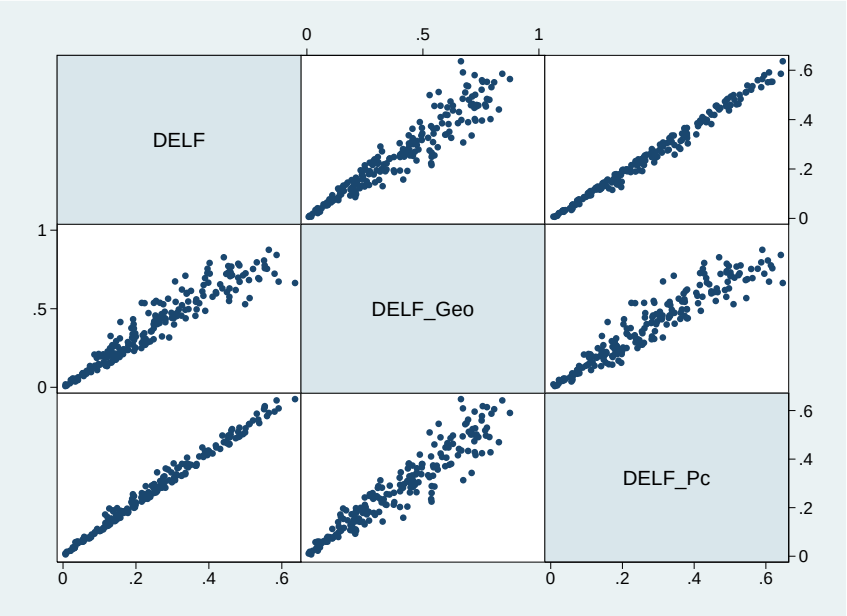


Figure 11: Scatter plots of the differently weighted $DELF$ values

dummy variables and use them for the PCA calculation.[82] To apply this procedure, one would need to define fixed categories of groups, which would work against the credo of this essay to refrain from such an approach. Additionally, the granularity of the data would, in any case, yield a significant number of groups and thus subsequent dummy variables.[83]

To bypass these problems, a more straightforward approach is used. In contrast to the previous weighting methods between the characteristics, the single $DELF$ values for each individual characteristic are used as the starting point for the PCA. Thus, the principal components are calculated as linear combinations of the three single $DELF$ values per country. They are combined in a way that explains the largest part of their variation. The first principal component explains most of the variance (62%), followed by the second (27%), and third (11%) principal component. The assigned loading factors can then be used to weight the sub-indices. The results of the PCA based on the three components are displayed in *Table 17*.

The loading factors found for the components of 0.66 for $DELF_L$, 0.54 for $DELF_E$ and 0.52 for $DELF_R$, confirm the equal weighting rather strongly.[84]

[82] This procedure was raised by Filmer and Pritchett (2001). If the categories can be transferred into an ordinal scale, then there exist procedures that improve the results (Kolenikov and Angeles, 2009). This, however, is not the case for the detailed group information on which the $DELF$ is build.

[83] For example, Bossert et al. (2011) only used five racial, and four unemployment categories in their GELF calculation.

[84] Nguefack-Tsague et al. (2011) show that PCA leads to a rather equal weighting scheme

Components/Factors	Comp. 1	Comp. 2	Comp. 3
$DELF_L$	0.658	-0.018	-0.753
$DELF_E$	0.541	-0.684	0.490
$DELF_R$	0.523	0.730	0.441
Eigenvalue	1.860	0.798	0.342
Proportion of explained variance	0.620	0.266	0.114
Cumulative explained variance	0.620	0.886	1.000

Table 17: Results of the principal component analysis and factor loadings for the components of $DELF$ sub-indices

Nevertheless, two slightly different ways of using the loading factors can be applied in order to utilize the detailed information of the PCA. For both indices, only the first principal component is used as it explains most of the variance (Ogwang and Abdou, 2003).[85] The approaches differ in the way they apply the loading factors. The first uses the calculated principal components of each observation and follows the approach of Noorbakhsh (1998). It is calculated as:

$$DELF_{PCA} = 1 - \left(\frac{d_i}{\bar{d} + 2s_d} \right) \tag{24}$$

with $\bar{d}$ and s_d representing the mean and the standard deviation of all d_i. d_i is the distance vector of country i from the most diverse country and is calculated as:

$$d_i = |z_i - z_{max}|$$

where z_i are the calculated principal components for each country i.

A simpler alternative multiplies the components by the PCA loading factors, and divides them by their sum (Ogwang and Abdou, 2003). As the first approach is the more accurate one, and the results do not differ significantly, it is used here.

B.5 Implications of similarity value construction and possible future extensions

One problem that all outlined methods share is the loss of information. The *WCE* data stick out because of their granularity and the advantage that all groups are defined along the three characteristics. In constructing the composite

if its components more or less demonstrate comparable correlation values. Only when these values deviate significantly does PCA not deliver results near to an equal weighting.

[85] Additionally, the negative loading factors of the second principal component complicate the interpretation

(average) index, one loses two pieces of information in the case of the $DELF$.

Firstly, information pertaining to the spread of groups and their mutual similarities is lost. This is a problem for any mean construction. Average values might emerge from very different base data setups. The mutual similarities might scatter only slightly around the mean, or be quite far apart. In the case of the $DELF$, one averages not only the group sizes, but also the similarity values. Covering the spread of similarity values is an important piece of information, but is hard to include in the $DELF$ index.[86]

To include this information, the most straightforward statistical measure would be to leverage the variance of the similarity values. A more elaborate method is found in Nguefack-Tsague et al. (2011), who, regarding the HDI, assess whether development is equal across all sub-indices, or if one or the other index deviates strongly from the overall mean of the composite index. For this, Nguefack-Tsague et al. (2011) suggest, calculating a balance of development index (BODI).[87] When adjusted for the $DELF$, it follows:

$$BODI = 1 - 1.5 \cdot \left((DELF_L - DELF)^2 + (DELF_E - DELF)^2 + (DELF_R - DELF)^2\right) \tag{25}$$

A BODI of one indicates that all components and the composite index are rather equal, whereas a BODI of zero characterizes countries where the sub-indices differ as much as possible from the composite index. *Figure 12* displays the $DELF$ values versus their respective BODI values.
The most significant imbalance is for Papua New Guinea, which has a low BODI value due to the deviation in its language and religious diversity from the composite mean. Equally imbalanced are some other small islands, where some differences in the setup of one characteristic have a large impact. The other most imbalanced countries are Bolivia and Belize (due to religion), Senegal and Mali (due to language), and Togo (due to the ethno-racial classification). On the other side of the coin, there are some countries that show remarkably equal values across all the single characteristics despite a high $DELF$ overall. These are Nepal, Kazakhstan, Mauritius, and Suriname. The BODI thus analyzes how differently the diversity of countries is spread, depending on the single characteristics.

The more serious problem is the second piece of information lost. By using

[86] A comparable thought was behind the introduction of the POL measure. Compared to the ELF, it covers other information about group size spread away from a equally sized duopoly.
[87] The acronym is adopted as it may very well stand for a 'balance of diversity index'.

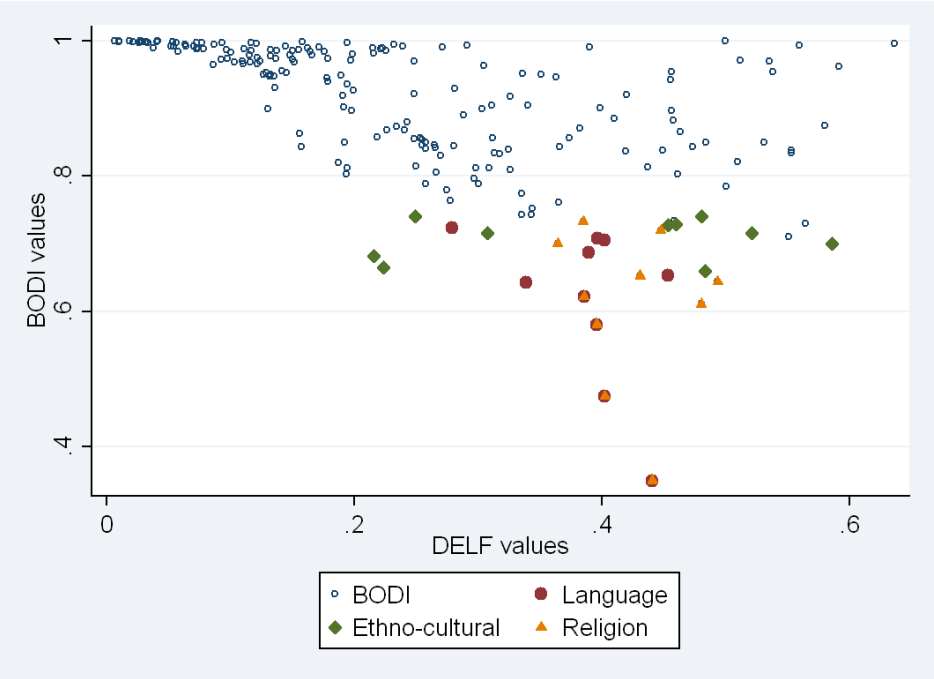


Figure 12: Scatter plots of BODI and *DELF* values. For countries with highest deviations, responsible characteristics marked

any of the above methods, one does not utilize the complete granularity of the data. This is easiest seen in the case for the arithmetic mean. There is mathematically no difference between using the average per characteristic, and the calculation of the composite similarity values at the most granular level. This equally applies for all other methods. To use this level of detail, one would like to assign similarity values not only per characteristic, but also to take the specific combination of the characteristics into account. Thus, one would need to assign specific complementarity factors between the characteristics to answer the question, whether a Christian, German speaking, Austrian is more distant for a Muslim, English speaking, Brit than for a Muslim, Urdu speaking, Brit. Based on these combinations, their mutual similarities might be less similar than only defined by the difference in their languages. It is obvious that these differences might also vary between cultural areas. Differences in religions might affect (dis)similarities between groups more in the Middle East than in Europe, whereas language differences are more important in the latter.[88] The mutual distance between an Christian, English speaking, American and a Muslim, Punjabi speaking Pakistani might be more profound in Pakistan than in

[88]These group distances might even be problem specific. Some combinations might be more prone to conflicts (e.g., religion), whereas other combinations might be more important in the field of trade (e.g., language).

the US. This is certainly a very important and interesting field of research. For the time being, however, the data to assess these differences are not available. Thus, for now it is assumed that the way of combining different characteristics is independent of the specific combination of single characteristics, and that it is comparable in all countries. Assessing the role of specific characteristic combinations in different cultural areas, and subsequently taking them into account, is a crucial step in improving the $DELF$ in the future.

B.6 Details of similarity interpretation between countries

Considering the $DELF$ between countries, it is obvious to explore how a theoretical country would look like to maximize (or minimize) this similarity measure with respect to a given country. Following its definition in chapter 6 the $DELF$ measures the expected dissimilarity between two individuals randomly drawn from each country. Thus, the similarity between two countries, as measured by the $DELF$, results not from the comparable structure of their respective people but from the consideration how similar two individuals are when they randomly meet.[89]

How would one expect a country j, whose group constellation (profile q) would make it most similar to country i, given its group profile p? Simplifying *Equation (10)* using p instead of p_i and q instead of p_j, p and q are row vectors of length K and M representing the respective group sizes/structures. Their elements range between zero and one and add up to a total of one. S is the $K \times M$ symmetric distance matrix with its elements equally ranging between zero and one. The $DELF$ between two countries is then given through:

$$DELF_{ij} = 1 - pSq' \tag{26}$$

As outlined earlier the key building bloc for the $DELF$ is the similarity vector S. If all its elements are zero (no group exist in both countries) the resulting $DELF$ is equal to one, attributed with two countries that are completely different. This is in line with what one would expect for two countries whose groups do not share any characteristic.[90] For the case when the groups in both countries share some characteristics (and more elements of S are non-zero) their

[89] This directly follows from the general construction of the GELF (Bossert et al., 2011) and taking the individual as the starting point of all considerations. However, the interpretation is slightly counter-intuitive as one would spontaneously regard two countries i and j as being 'similar' if their group profiles are similar, i.e., if $p_{ik} \approx p_{jm}$ and the corresponding $\hat{s}_{km} \approx 1$ for all $k = 1, 2, ..., K$ and $m = 1, 2, ..., M$.

[90] Note that the respective group constellations p and q for both countries are irrelevant in this case.

group profiles p and q are relevant. If the group sizes p_{ik} and p_{jm} with a corresponding similarity value of $\hat{s}_{km} > 0$ are small enough both countries are still approximately completely different with a $DELF$ tending to one.

On the contrary, a $DELF$ value of zero between two countries is attained if both countries consist of only one, completely similar group in both countries. For any country i with more than one group ($K, M > 1$), which is the case in all countries covered by the *WCE* data, the extreme value of zero is not attained. The more elements of the similarity matrix S are non-zero the lower will be the resulting $DELF$ value. Thus, lower values of $DELF$ correctly indicate country pairs where the expected dissimilarity between two individuals is lower. However, given two countries have the identical groups ($\hat{s}_{km} \approx 1$), not the identical group constellations minimizes the $DELF$ value. *Equation (26)* is minimized with respect to the group constellation of the second country q when

$$pSq' = \sum_{k=1}^{K} a_k q_k \tag{27}$$

is maximized, where a_m is the m-th element of the vector pS. Now

$$\sum_{k=1}^{K} a_k q_k \leq a_n \sum_{k=1}^{K} q_k = a_n \tag{28}$$

where a_n is the largest entry of the vector $(a_1, a_2, ..., a_K)$, and this maximum is attained by setting $q_n = 1$ and $q_m = 0$ for all $m \neq n$. A country j would be most similar to country i is one in which the entire population of country j belongs to a single group, namely the group n, where n is the subscript of the largest entry of the vector $(a_1, a_2, ..., a_K)$.

Despite the maximization result, the general interpretation of lower levels of $DELF$ reflecting countries that share more groups with similar characteristics is still valid. As most countries have a high number of groups the result of the theoretical maximization process leading to a single group maximizing the similarity level between both is less relevant than the similarity values between those groups. However, one has to consider that the way the $DELF$ measures 'similarity' between two countries slightly deviates from ones general expectation of two 'similar' countries.

B.7 Details of population weighting for regional means

The $DELF$ values between countries represent the expected dissimilarity between two individuals randomly drawn, each from a different country. Thus,

one individual is randomly drawn from country A and the other from country B, and their mutual diversity is then assessed. For this assessment different population sizes of the two countries do not matter, as only the relative group sizes determine the probabilities to be matched. This concept is thus only applicable for tuples.

As soon as an expected level of diversity between more than two countries is concerned, for example, in the case of regional averages, a different calculation applies and population size matters. The two individuals are no longer drawn randomly from each country, instead two individuals are randomly drawn out of the region. To be drawn from one country or the other depends on the relative sizes of their population in relation to the region's overall population. The expected (average) diversity between any two individuals drawn is then easily given by the $DELF$ value between those two countries. Mathematically, the formula for the regional average of region r is given through:

$$\begin{aligned} DELF_r &= \sum_{i=1}^{R}\sum_{j=1}^{R} \frac{n_i}{N_r} \cdot \frac{n_j}{N_r} \cdot DELF_{ij} \\ &= \frac{1}{N_r^2} \cdot \sum_{i=1}^{R}\sum_{j=1}^{R} n_i \cdot n_j \cdot DELF_{ij} \end{aligned} \tag{29}$$

where region r consists of countries $i, j \in \{1, ...R\}$. Their between country $DELF$s are given by $DELF_{ij}$ for all $i, j \in \{1, ..., R\}$. n_i and n_j are the respective populations of country i and j and $N_r = \sum_{i=1}^{R} n_i \in \{1, ..., R\}$ the region's total population size.

In contrast to the $DELF$ formula in *Equation (8)*, the sum does not need to be subtracted from one. In *Equation (8)*, $\hat{s}_{kl}$ is a measure of similarity, whereas the $DELF$ in *Equation (29)* is already a measure of dissimilarity or diversity.

For dynamic regions it does, however, have an important implication when new countries join or members secede. When an additional country joins a specific region (e.g., the EU) it brings two different types of diversity into this region. First, it enters the new region with its internal (rather homogeneous) diversity. Secondly, it has its external (rather heterogeneous) diversity towards all members of the region. Depending on population size differences and the two types of diversity values, the additional country can either increase or decrease the diversity in the region.

C Detailed DELF data per country

Table 18: ELF and DELF values and ranks for 210 countries

Country	ELF	Rank	DELF	Rank	Delta	$DELF_L$	$DELF_E$	$DELF_R$
Papua New Guinea	0.982	1	0.441	36	-35	0.942	0.360	0.021
Congo, Dem. Rep.	0.977	2	0.258	91	-89	0.545	0.208	0.021
Solomon Islands	0.971	3	0.402	42	-39	0.845	0.349	0.013
Cameroon	0.966	4	0.553	7	-3	0.809	0.354	0.497
Chad	0.963	5	0.564	5	0	0.876	0.277	0.540
Tanzania	0.962	6	0.340	60	-54	0.307	0.181	0.533
India	0.958	7	0.326	66	-59	0.513	0.200	0.266
Central African Republic	0.953	8	0.437	37	-29	0.703	0.208	0.399
Vanuatu	0.948	9	0.386	49	-40	0.740	0.388	0.030
Cote d'Ivoire	0.943	10	0.586	3	7	0.867	0.243	0.648
United Arab Emirates	0.939	11	0.580	4	7	0.737	0.654	0.350
Mozambique	0.927	12	0.288	80	-68	0.278	0.102	0.485
Liberia	0.921	13	0.553	8	5	0.774	0.307	0.578
Singapore	0.917	14	0.501	16	-2	0.715	0.201	0.586
Nigeria	0.917	16	0.551	9	7	0.861	0.240	0.553
Kenya	0.917	15	0.382	51	-36	0.621	0.279	0.246
Ghana	0.915	17	0.458	27	-10	0.740	0.147	0.488
Zambia	0.914	18	0.127	158	-140	0.272	0.077	0.031
Togo	0.913	19	0.484	20	-1	0.723	0.099	0.629
Congo, Rep.	0.910	20	0.192	125	-105	0.367	0.201	0.007
Timor-Leste	0.904	21	0.458	28	-7	0.546	0.596	0.231
Israel	0.903	22	0.402	43	-21	0.738	0.116	0.352
Uganda	0.901	23	0.275	85	-62	0.570	0.219	0.036
Benin	0.885	29	0.460	26	3	0.671	0.115	0.593
South Africa	0.898	24	0.374	52	-28	0.520	0.478	0.123
Guinea-Bissau	0.898	25	0.521	13	12	0.814	0.201	0.548
Madagascar	0.892	26	0.255	94	-68	0.188	0.070	0.507
Mali	0.887	27	0.453	33	-6	0.814	0.407	0.139
Namibia	0.886	28	0.385	50	-22	0.575	0.539	0.041
Zimbabwe	0.884	30	0.148	144	-114	0.233	0.147	0.065
Ethiopia	0.863	34	0.453	32	2	0.721	0.127	0.512
Philippines	0.875	31	0.281	81	-50	0.457	0.210	0.177
Bhutan	0.869	32	0.512	14	18	0.619	0.425	0.491
Fiji	0.868	33	0.591	2	31	0.713	0.570	0.491
Indonesia	0.855	37	0.303	75	-38	0.501	0.140	0.269
Iran, Islamic Rep.	0.855	35	0.344	58	-23	0.536	0.483	0.014
Burkina Faso	0.855	36	0.462	25	11	0.703	0.193	0.489
New Caledonia	0.855	38	0.480	21	17	0.686	0.691	0.065
Sierra Leone	0.845	39	0.531	12	27	0.780	0.348	0.466
Angola	0.845	40	0.116	166	-126	0.199	0.113	0.035
Micronesia, Fed. Sts.	0.840	41	0.278	84	-43	0.580	0.229	0.026
Malaysia	0.836	42	0.510	15	27	0.685	0.231	0.614
Gabon	0.835	43	0.227	107	-64	0.453	0.189	0.039
Italy	0.829	44	0.122	161	-117	0.224	0.094	0.047
Qatar	0.828	45	0.484	19	26	0.572	0.651	0.230
Senegal	0.824	46	0.339	61	-15	0.734	0.181	0.101
United States	0.823	47	0.448	35	12	0.589	0.657	0.097

Continued on next page

Table 18 – continued from previous page

Country	ELF	Rank	DELF	Rank	Delta	$DELF_L$	$DELF_E$	$DELF_R$
Suriname	0.818	48	0.636	1	47	0.657	0.660	0.592
Lao PDR	0.816	49	0.536	11	38	0.649	0.458	0.500
Niger	0.782	58	0.396	45	13	0.728	0.353	0.108
Brunei Darussalam	0.809	50	0.480	22	28	0.679	0.143	0.620
Malawi	0.807	51	0.138	148	-97	0.154	0.062	0.197
Mauritius	0.807	52	0.560	6	46	0.609	0.518	0.551
Peru	0.803	53	0.336	63	-10	0.421	0.576	0.010
France	0.802	54	0.336	62	-8	0.453	0.355	0.202
N. Mariana Islands	0.798	55	0.396	46	9	0.775	0.385	0.028
Thailand	0.793	56	0.216	113	-57	0.304	0.155	0.189
Belgium	0.782	57	0.314	69	-12	0.560	0.290	0.091
Belize	0.779	59	0.494	18	41	0.677	0.708	0.096
Kuwait	0.777	60	0.363	56	4	0.446	0.434	0.209
Pakistan	0.777	61	0.243	102	-41	0.410	0.299	0.021
Gambia, The	0.774	62	0.390	48	14	0.745	0.311	0.113
Afghanistan	0.774	63	0.297	78	-15	0.500	0.388	0.003
Morocco	0.770	64	0.187	128	-64	0.464	0.097	0.002
Monaco	0.765	65	0.190	127	-62	0.296	0.228	0.045
Oman	0.759	66	0.474	23	43	0.634	0.574	0.212
Guinea	0.753	67	0.464	24	43	0.647	0.233	0.512
Canada	0.751	68	0.419	40	28	0.632	0.455	0.171
Mauritania	0.750	69	0.265	90	-21	0.412	0.378	0.004
Bolivia	0.749	70	0.431	38	32	0.678	0.572	0.043
Spain	0.745	71	0.195	120	-49	0.313	0.240	0.032
Nepal	0.744	72	0.390	47	25	0.446	0.388	0.336
Sudan	0.738	73	0.538	10	63	0.664	0.534	0.417
Ecuador	0.737	74	0.307	73	1	0.282	0.627	0.013
Latvia	0.728	75	0.250	97	-22	0.510	0.226	0.014
Eritrea	0.721	76	0.398	44	32	0.508	0.189	0.498
Guyana	0.707	77	0.457	29	48	0.248	0.600	0.522
Nauru	0.705	78	0.449	34	44	0.690	0.432	0.226
Myanmar	0.699	79	0.420	39	40	0.589	0.264	0.408
Trinidad and Tobago	0.698	80	0.410	41	39	0.188	0.559	0.483
Andorra	0.693	81	0.137	149	-68	0.213	0.164	0.034
Cayman Islands	0.686	82	0.253	96	-14	0.237	0.480	0.043
Bosnia and Herzegovina	0.686	83	0.351	57	26	0.273	0.281	0.499
Guam	0.679	84	0.343	59	25	0.645	0.325	0.061
Switzerland	0.677	85	0.317	68	17	0.572	0.274	0.106
Colombia	0.677	86	0.224	109	-23	0.050	0.609	0.012
Montenegro	0.671	87	0.223	110	-23	0.219	0.167	0.283
Guatemala	0.668	88	0.364	55	33	0.571	0.522	0.000
New Zealand	0.667	89	0.366	53	36	0.505	0.491	0.103
French Polynesia	0.661	90	0.258	93	-3	0.447	0.325	0.001
Brazil	0.660	91	0.216	114	-23	0.048	0.591	0.008
Mexico	0.658	92	0.249	98	-6	0.168	0.575	0.005
Equatorial Guinea	0.655	93	0.266	88	5	0.543	0.214	0.042
Djibouti	0.644	94	0.279	83	11	0.619	0.180	0.037
Algeria	0.635	95	0.156	139	-44	0.401	0.065	0.003
Iraq	0.633	96	0.326	65	31	0.454	0.489	0.036
Estonia	0.631	97	0.299	77	20	0.449	0.437	0.010
Luxembourg	0.620	98	0.248	101	-3	0.468	0.250	0.028

Continued on next page

Table 18 – continued from previous page

Country	ELF	Rank	DELF	Rank	Delta	$DELF_L$	$DELF_E$	$DELF_R$
Panama	0.616	99	0.366	54	45	0.465	0.584	0.048
Macedonia, FYR	0.613	100	0.456	30	70	0.578	0.332	0.459
Grenada	0.611	101	0.116	165	-64	0.156	0.193	0.000
Kazakhstan	0.603	102	0.499	17	85	0.513	0.487	0.498
St. Lucia	0.600	103	0.133	154	-51	0.197	0.168	0.033
China	0.594	104	0.234	105	-1	0.223	0.035	0.445
Egypt, Arab Rep.	0.589	105	0.065	185	-80	0.086	0.099	0.008
Georgia	0.586	106	0.311	71	35	0.506	0.272	0.155
Greenland	0.581	107	0.241	103	4	0.385	0.338	0.000
Bahrain	0.576	108	0.455	31	77	0.548	0.522	0.296
Nicaragua	0.575	109	0.301	76	33	0.371	0.524	0.008
Bermuda	0.574	110	0.192	124	-14	0.138	0.438	0.001
Virgin Islands (U.S.)	0.570	111	0.309	72	39	0.437	0.470	0.020
Comoros	0.567	112	0.041	192	-80	0.057	0.025	0.042
Mongolia	0.506	125	0.266	89	36	0.191	0.083	0.523
Turkey	0.560	113	0.255	95	18	0.328	0.430	0.006
Mayotte	0.545	114	0.335	64	50	0.495	0.492	0.019
Netherlands	0.542	115	0.215	115	0	0.261	0.237	0.147
Venezuela, RB	0.542	116	0.194	122	-6	0.059	0.484	0.040
Kyrgyz Republic	0.539	117	0.291	79	38	0.334	0.297	0.242
Albania	0.539	118	0.248	100	18	0.334	0.140	0.272
Ireland	0.539	119	0.194	123	-4	0.488	0.073	0.020
Australia	0.534	120	0.305	74	46	0.381	0.354	0.178
Sri Lanka	0.503	126	0.312	70	56	0.440	0.060	0.437
Bahamas, The	0.523	121	0.146	145	-24	0.220	0.215	0.002
Germany	0.518	122	0.165	135	-13	0.242	0.156	0.096
Tajikistan	0.510	123	0.325	67	56	0.467	0.449	0.058
St. Vincent & the Gr.	0.508	124	0.199	117	7	0.210	0.272	0.113
Sweden	0.503	127	0.179	130	-3	0.255	0.207	0.074
Chile	0.500	128	0.219	112	16	0.213	0.439	0.004
Norway	0.492	129	0.133	152	-23	0.202	0.124	0.072
Cape Verde	0.488	130	0.270	87	43	0.446	0.364	0.000
Liechtenstein	0.485	131	0.225	108	23	0.300	0.211	0.165
Dominican Republic	0.481	132	0.130	156	-24	0.048	0.340	0.003
Tuvalu	0.471	133	0.058	187	-54	0.141	0.033	0.000
United Kingdom	0.470	134	0.176	132	2	0.244	0.183	0.101
Bangladesh	0.341	153	0.098	172	-19	0.050	0.039	0.204
Botswana	0.462	136	0.158	137	-1	0.175	0.137	0.162
Tunisia	0.464	135	0.038	194	-59	0.107	0.006	0.002
Cuba	0.449	137	0.281	82	55	0.018	0.417	0.407
Puerto Rico	0.446	138	0.157	138	0	0.048	0.419	0.005
Argentina	0.444	139	0.249	99	40	0.245	0.412	0.089
Moldova	0.444	140	0.198	118	22	0.395	0.173	0.027
Palau	0.437	141	0.258	92	49	0.401	0.373	0.000
Netherlands Antilles	0.426	142	0.200	116	26	0.337	0.233	0.029
Saudi Arabia	0.420	143	0.197	119	24	0.263	0.243	0.086
Libya	0.415	144	0.117	164	-20	0.172	0.139	0.039
Ukraine	0.403	145	0.094	174	-29	0.115	0.110	0.057
Aruba	0.399	146	0.191	126	20	0.222	0.337	0.013
Uzbekistan	0.375	147	0.155	140	7	0.207	0.180	0.078
Russian Federation	0.374	148	0.271	86	62	0.328	0.272	0.215

Continued on next page

Table 18 – continued from previous page

Country	ELF	Rank	DELF	Rank	Delta	$DELF_L$	$DELF_E$	$DELF_R$
Somalia	0.372	149	0.079	178	-29	0.147	0.063	0.026
Jamaica	0.364	150	0.087	176	-26	0.081	0.130	0.050
Costa Rica	0.363	151	0.136	150	1	0.083	0.308	0.018
Bulgaria	0.337	156	0.232	106	50	0.228	0.278	0.190
Turkmenistan	0.344	152	0.121	162	-10	0.151	0.136	0.076
Syrian Arab Republic	0.340	154	0.152	141	13	0.217	0.204	0.033
Dominica	0.337	155	0.110	169	-14	0.199	0.129	0.002
Austria	0.332	157	0.151	142	15	0.221	0.145	0.085
Belarus	0.329	158	0.041	193	-35	0.053	0.057	0.013
Barbados	0.324	159	0.122	160	-1	0.107	0.236	0.024
Jordan	0.321	160	0.057	188	-28	0.082	0.066	0.023
Serbia	0.318	161	0.171	133	28	0.214	0.194	0.106
Vietnam	0.309	162	0.221	111	51	0.265	0.149	0.250
Paraguay	0.308	163	0.179	129	34	0.269	0.252	0.016
Lesotho	0.308	164	0.034	195	-31	0.061	0.039	0.002
American Samoa	0.307	165	0.135	151	14	0.277	0.115	0.014
Uruguay	0.305	166	0.133	153	13	0.085	0.279	0.034
Greece	0.304	167	0.166	134	33	0.261	0.132	0.104
Swaziland	0.304	168	0.064	186	-18	0.098	0.078	0.016
Lebanon	0.302	169	0.239	104	65	0.276	0.259	0.183
Hungary	0.290	170	0.178	131	39	0.223	0.285	0.026
Lithuania	0.284	171	0.132	155	16	0.269	0.120	0.008
Honduras	0.270	172	0.129	157	15	0.124	0.257	0.006
West Bank and Gaza	0.266	173	0.150	143	30	0.155	0.052	0.243
Antigua and Barbuda	0.262	174	0.093	175	-1	0.072	0.198	0.008
Croatia	0.248	175	0.097	173	2	0.150	0.121	0.021
Slovak Republic	0.247	176	0.142	147	29	0.207	0.217	0.001
Azerbaijan	0.244	177	0.145	146	31	0.177	0.173	0.086
Cambodia	0.233	178	0.195	121	57	0.219	0.203	0.163
Isle of Man	0.222	179	0.027	204	-25	0.015	0.064	0.002
Kosovo	0.220	180	0.163	136	44	0.214	0.099	0.175
Romania	0.216	181	0.124	159	22	0.173	0.191	0.008
El Salvador	0.215	182	0.104	170	12	0.106	0.204	0.001
Marshall Islands	0.210	183	0.111	168	15	0.122	0.210	0.000
Samoa	0.210	184	0.086	177	7	0.207	0.051	0.000
Yemen, Rep.	0.195	185	0.074	180	5	0.137	0.063	0.023
Slovenia	0.192	186	0.054	190	-4	0.079	0.046	0.037
Finland	0.177	187	0.101	171	16	0.146	0.142	0.015
Cyprus	0.173	188	0.112	167	21	0.170	0.123	0.042
Portugal	0.173	189	0.074	181	8	0.056	0.144	0.023
Denmark	0.165	190	0.117	163	27	0.144	0.122	0.086
San Marino	0.164	191	0.010	207	-16	0.029	0.002	0.000
St. Kitts and Nevis	0.153	192	0.073	182	10	0.066	0.105	0.049
Sao Tome and Principe	0.153	193	0.052	191	2	0.058	0.098	0.000
Rwanda	0.147	194	0.032	198	-4	0.013	0.044	0.039
Iceland	0.141	195	0.054	189	6	0.107	0.052	0.004
Malta	0.119	196	0.073	183	13	0.110	0.108	0.001
Seychelles	0.117	197	0.070	184	13	0.087	0.110	0.014
Czech Republic	0.109	198	0.033	197	1	0.050	0.042	0.006
Haiti	0.108	199	0.010	208	-9	0.008	0.021	0.001

Continued on next page

Table 18 – continued from previous page

Country	ELF	Rank	DELF	Rank	Delta	$DELF_L$	$DELF_E$	$DELF_R$
Poland	0.102	200	0.033	196	4	0.065	0.035	0.001
Armenia	0.100	201	0.077	179	22	0.099	0.090	0.042
Burundi	0.099	202	0.028	202	0	0.022	0.038	0.025
Tonga	0.094	203	0.031	200	3	0.055	0.035	0.004
Korea, Rep.	0.059	204	0.032	199	5	0.045	0.009	0.041
Maldives	0.059	205	0.028	203	2	0.043	0.018	0.022
Faeroe Islands	0.058	206	0.006	210	-4	0.010	0.009	0.000
Channel Islands	0.055	207	0.029	201	6	0.053	0.029	0.005
Kiribati	0.050	208	0.021	205	3	0.050	0.014	0.000
Japan	0.048	209	0.019	206	3	0.032	0.011	0.014
Korea, Dem. Rep.	0.019	210	0.007	209	1	0.015	0.006	0.000

Table 19: Country-pairs with highest mutual (dis)similarities[91]

	Region	Country A	Region	Country B	DELF	$DELF_L$	$DELF_E$	$DELF_R$
Most similar countries	SSA	Burundi	SSA	Rwanda	0.047	0.068	0.041	0.032
	MENA	Jordan	MENA	Egypt	0.072	0.118	0.083	0.015
	MENA	Jordan	MENA	Yemen.	0.081	0.155	0.065	0.023
	MENA	Egypt	MENA	Yemen	0.083	0.151	0.082	0.015
	LA	Antigua	LA	St. Kitts	0.085	0.070	0.155	0.029
	Western	Iceland	Western	Faeroe I.	0.086	0.115	0.141	0.002
	MENA	Jordan	MENA	Tunisia	0.089	0.217	0.037	0.012
	MENA	Egypt	MENA	Tunisia	0.091	0.214	0.055	0.005
	MENA	Egypt	MENA	Libya	0.093	0.136	0.120	0.024
	MENA	Yemen	MENA	Tunisia	0.098	0.247	0.035	0.012
Most dissimilar countries	Asia	Kiribati	MENA	Algeria	1.000	1.000	1.000	1.000
	Asia	Korea, Rep.	SSA	Niger	1.000	1.000	1.000	1.000
	Asia	Lao PDR	SSA	Eritrea	1.000	1.000	1.000	1.000
	Asia	Bhutan	SSA	Gabon	1.000	1.000	1.000	1.000
	Asia	Bhutan	SSA	Congo, Rep.	1.000	1.000	1.000	1.000
	SSA	Djibouti	Asia	Lao PDR	1.000	1.000	1.000	1.000
	Asia	Lao PDR	MENA	Tunisia	1.000	1.000	1.000	1.000
	Asia	Lao PDR	SSA	Mauritania	1.000	1.000	1.000	1.000
	Asia	Lao PDR	MENA	West Bank	1.000	1.000	1.000	1.000
	Asia	Lao PDR	MENA	Morocco	1.000	1.000	1.000	1.000

Table 20: Details of EU enlargement waves and respective *DELF* averages

	Enlargement waves							**Potential future enlargement**		
EU group	**EU6**	**EU9**	**EU10**	**EU12**	**EU15**	**EU25**	**EU27**	**EU+B**	**EU+T**	**EU+B+T**
Year	1952	1973	1981	1986	1995	2004	2007	–	–	–
Countries	Belgium	Denmark	Greece	Portugal	Austria	Cyprus	Bulgaria	Albania	Turkey	Albania
	France	Ireland		Spain	Finland	Czech Rep.	Romania	Croatia		Croatia
	Germany	Britain			Sweden	Estonia		Iceland		Iceland
	Italy					Hungary		Macedonia		Macedonia
	Luxembourg					Latvia		Montenegro		Montenegro
	Netherlands					Lithuania		Serbia		Serbia
						Malta				Turkey
						Poland				
						Slovak Rep.				
						Slovenia				
DELF	0.3685	0.3875	0.3940	0.3893	0.3987	0.4196	0.4206	0.4263	0.5383	0.5393
Delta		*0.019*	*0.006*	*-0.005*	*0.009*	*0.021*	*0.001*	*0.006*	*0.118*	*0.119*