Economics School of Louvain - ESL

Economics School of Namur - ESN

Local Ethnic Contact and Civil Conflicts

Author : MASSON Laurent

Thesis Director : GOMES Joseph

Thesis Reader : CASSAN Guilhem

Academic Year 2021-2022

Master in Economics – 120 credits – Focus: Professional

Université catholique de Louvain - Economics School of Louvain - Place Montesquieu 3, 1348 Louvain-La-Neuve

Université de Namur – Economics School of Namur – Rempart de la Vierge 8, 5000 Namur

Abstract

This paper examines empirically the impact of ethnolinguistic divisions on civil conflicts through the prism of local interactions. Adding the local-global complementarity from Desmet et al. (2020) to Esteban et al. (2012) framework on ethnic divisions, I measure how local interactions play a role in the intensity of conflict. A spatial resolution of 5 km by 5 km is used to capture the local component and I look attentively at the effects of ethnic polarization, fractionalization, and the local-global complementarity on conflict. I provide empirical evidence that an increase in local interactions with other groups increases the overall antagonism felt towards those groups in the society at large.

1. Introduction

This work examines the impact of local interactions between different ethnolinguistic groups on civil conflicts.

The fact that conflicts have always been part of our societies and can be considered inherent to human being is common sense. Nevertheless, the cause of those conflicts is not always easy to detect. In our modern societies, economic inequalities between poor and rich or political fights over a new policy are often good reasons to cause civil conflicts but those are not exhaustive. Therefore, number of works examining the diverse potential causes of conflicts already exist in the literature.

In this paper, I analyze whether ethnolinguistic differences matter in the explanation of civil conflict following the work of Esteban et al. (2012) (hereafter EEA). To assess those differences, several variables are used. The first one is the fractionalization index which has been used in many empirical studies in the literature. The second one is the ethnic polarization from Montalvo and Reynal-Querol (2005) and the third one is a Greenberg-Gini index of ethnic differences. Those three variables have different importance depending on whether the conflict's payoffs are private or public and whether the within-group cohesion level is high or low.

Moreover, the specificity of my work is the addition to those variables of a geographic component because I believe that the spatial distribution of diversity matters in explaining conflicts. Recent papers in the literature are focused on ethnic polarization and this work is intended to give another story looking at a variable that has never been taken into account in the computation of the intensity of civil conflict. I use the data of ethnolinguistic diversity created by Desmet et al. (2020) (hereafter DEA) which divided the world into a grid composed of 5 km by 5 km cells to evaluate local interactions between ethnic groups. Local ethnic contact is captured by a local-global complementarity variable at different levels of linguistic aggregation. The higher the level, the wider the aggregation. Even though the dataset from DEA was built to estimate the effect of local interactions on the provision of public goods, I will use it in this work to test whether including this local-global complementarity variable makes a difference on conflicts compared to the results obtained by EEA.

Differences on conflict could indeed be expected following two major theories in the literature: contact theory and conflict theory. The former states that different ethnolinguistic groups, by interacting together, would more easily accept others and identify to them. Local interactions would then mitigate conflict because people would feel less antagonism. On the other hand, conflict theory shows that the society is in permanent conflict due to incompatible objectives and class or ethnic group differences. Everyone is looking for power, but resources are limited. An increase in interactions between ethnic groups would then foster conflict.

Using the PRIO/UCDP conflict data, in the baseline analysis, I estimate whether local interactions with other groups affecting the overall antagonism towards those other ethnolinguistic groups in the society at large impact the intensity of conflict. In this paper, only civil conflicts are examined which are conflicts where one of the parties involved is representing a country's current government. I find that even though the local-global complementarity is mostly not statistically significant, an increase in local interactions is consistent with an increase in the intensity of conflict. Those results are consistent with conflict theory: interacting with other groups at the local level increases the overall antagonism towards those groups in the society at large leading to more conflicts.

Next, the importance of the local-global complementarity is tested against regressions using alternative measures from EEA and others. First, different dependent variables are tested as proxies for the intensity of conflict with among them, higher thresholds from the PRIO/UCDP dataset. Second, alternative groupings are taken using data from the Ethnologue (Lewis, 2009). Third, in DEA's dataset, they also computed the local-global complementarity using 10 km by 10 km cells, suggesting that 5 km by 5 km cells might be too small to evaluate local interactions. I examine the differences on the intensity of conflict between those two local-global complementarity measures. Fourth, using the language tree from the Ethnologue, I recompute the results but for the levels 2 and 15 of linguistic aggregation. Fifth, due to the potential presence of endogeneity, I will use the instrument for the local-global complementarity as developed by DEA. This instrument has been created because they were evaluating the effect of spatial mixing on the provision of public goods while there was a possibility that public goods provision impacts the spatial mixing.

Previous works e.g., Esteban and Ray (2011), focused only on the ethnic fractionalization, polarization, and inequalities to explain conflicts. My work contributes to the literature by showing an empirical study of the effect of local-global complementarity on conflict. Since, to the best of my knowledge, it has not been done before, my work provides a baseline for further studies analyzing the effect of local interactions on conflict.

This work is organized as follows. Section 2 describes the conceptual framework while the data is presented in Section 3. Section 4 provides the results from the baseline analysis and Section 5 extends the analysis with the use of alternative measures. Section 6 concludes.

2. Conceptual framework

The background I am using for this work comes from Desmet et al. (2020) and Esteban et al. (2012). EEA tried to find out the way ethnic divisions mattered in explaining conflicts and the private or public prize yielded. To define "ethnic divisions" they first used the well-known fractionalization index. This index, that is measuring the probability that two randomly picked individuals belong to different ethnolinguistic groups, has been frequently used as an independent variable in the literature; see, e.g., Collier and Hoeffler (2004), Fearon and Laitin (2003 a), and Miguel, Satyanath, and Sergenti (2004). Nevertheless, in previous empirical works, a connection between this index and civil conflicts has not always been found and, as DEA suggests, those papers were surely omitting a variable, but I will discuss that below.

Then EEA used the index of ethnic polarization from Montalvo and Reynal-Querol (2005) who were the first to propose an index of polarization capturing how far the distribution of ethnic groups is from the (1/2,0, 0, 0, …0, 1/2) bipolar distribution which is the highest level of ethnic polarization. When the minority ethnic group is large enough and undivided, then the index is close to its maximum. Collier and Hoeffler (1998) emphasize that coordination costs would be extremely low in the case of a minority ethnic group being of the same size as the group identified as the governmental group. Therefore, we can assume that this situation would easily lead to civil conflict.

In the literature, Duclos et al. (2004) and Esteban and Ray (2011) added the notion of intergroup distance in their definition of ethnic polarization. This distance parameter is used to assess the payoffs a group will derive from the victory of another group in the conflict. Since in most conflicts, winners yield a public prize that can be dominance in the political, religious, or cultural sector, each other group will be affected depending on their distance from the winner.

EEA used the data and the theory from Esteban and Ray (2011) which report that conflicts over public and private goods should be differentiated and that every conflict can be explained by three factors: Fractionalization, Polarization, and a Greenberg-Gini index of ethnic difference. In function of the nature of the conflict, those indices will have different importance in explaining conflict intensity. When the group cohesion is high, ethnic fractionalization and polarization will impact conflict if the prize is respectively private or public. A private prize can be defined as the capture of oil exploitations or mining resources. On the other hand, the Greenberg-Gini index will be important if the group cohesion is low. To define privateness, a dummy for the production of oil and diamond is used and publicness is captured by several controls of autocracy.

EEA described a measure of ethnic polarization (POL) based on Duclos et al. (2004) and Esteban and Ray (2011) and, the Greenberg-Gini index (GINI) with m groups engaged in conflict, s_i being the population share of group i and d_{ij} the distance between group i and group j :

$$POL = \sum_{i=1}^{m} \sum_{j=1}^{m} s_i^2 s_j d_{ij}$$

$$GINI = \sum_{i=1}^{m} \sum_{j=1}^{m} s_i s_j d_{ij}$$

Where the only difference between the measure of polarization and the Greenberg-Gini index is the squaring in the own-group share which does not seem to be of importance but in fact matters a lot. In GINI, the interpretation of the own-group share is a trivial counting of the group size while for POL it reflects the effect of identification for an individual to its own group. DEA gives us an interpretation that I will discuss below. Next, I define the ethnolinguistic fractionalization index (FRAC) (called global ethnolinguistic fractionalization in DEA).

$$FRAC = \sum_i \sum_{j \neq i} s_i s_j$$

DEA analyze how linguistic diversity can impact the provision of the public goods. They first show that the level of antagonism influences the public goods, then they brought to the discussion that analyses of ethnolinguistic differences would always be biased if local interactions were not considered. They divided the world into 5 km by 5 km cells and defined average antagonism in society by two terms. The first one is the ethnolinguistic fractionalization and the second one is the local-global ethnic complementarity (LGC) which assesses how interactions with other ethnic groups inside a cell affect the overall antagonism towards those same groups in the society at large. This variable is the probability that when three individuals are randomly picked: individuals 1 and 2 from the same cell and individual 3 from the society at large, 2 and 3 would be from the same group and 1 from a different one.

The local-global ethnic complementarity (LGC) in a cell k is

$$LGC = \beta \sum\nolimits_k s_k \sum\nolimits_i \sum\nolimits_{j \neq i} s_{ki} s_{kj} s_j$$

Where β can have a positive or negative sign depending on local interactions increasing or decreasing the average antagonism in the society at large and s_{ki} being the share of the population of ethnolinguistic group i in cell k.

DEA made an interesting comparison of LGC with ethnic polarization. In a society that would be perfectly geographically mixed, $s_{ki} = s_i$ for all cells and all groups, the local-global complementarity would be equal to $\sum_i \sum_{j \neq i} s_i s_j^2$. Now, when assuming that the intergroup distance between all groups in the society equals to 1, we would face an equation similar to the polarization ($POL = \sum_i \sum_{j \neq i} s_i^2 s_j$). In the polarization equation, the squared own-group share suggests that the more an individual identifies to his own group, the greater the antagonism towards other groups. On the opposite, a possible explanation for the squared other-group share in the local-global complementarity would be that the local interactions with other groups would lead to an identification to those other groups and thus would potentially mitigate the overall antagonism felt towards those groups.

I will use DEA's data, specifically the local-global complementarity variable, and insert it in the dataset from EEA to see in which direction interactions at a local level affecting overall antagonism would impact civil conflicts.

3. Data

The dataset I use is a merger between the dataset from Esteban et al. (2012) and the data on local interactions from Desmet et al. (2020). It covers 138 countries for a period going from 1960 to 2008 divided into 10 subperiods of 5 years.[1] The baseline is taken from EEA where I have added the local-global complementarity variable.

3.1. Conflict

The intensity of conflict is measured based on the number of battle deaths from the UCDP/PRIO dataset.[2] Since available information cannot be trusted, a convention has been made to measure conflict by a binary variable. There are several levels in the dataset attesting for the severity of the conflict. In my baseline, I use the lowest level, PRIO25, which takes the value of 1 when conflicts reach 25 battle death per year and 0 otherwise. There are two other levels above: PRIOCW, which attests when conflicts reach 1000 battle deaths since onset and PRIO1000, which takes the value of 1 when conflicts reach 1000 deaths per year and that can be considered as "civil war".

3.2. Groups and Distances

I use the data from EEA which have made an analysis based on an update of the data from Fearon (2003 b). In the dataset, more than 800 "ethnic and ethnoreligious" groups in 160 countries are identified. That classification is the one adopted by EEA because considered to be among all attempts to describe ethnic divisions one of the most careful ones.

To define intergroup distances, the preferences over public goods would be needed but since this data does not exist, an alternative had to be found. In the literature, Greenberg (1956); Laitin (2000); Fearon (2003 a); Desmet et al. (2009); and Desmet et al. (2012), the linguistic distance has been used as a proxy for the distance between two groups in preferences over public goods. Now, to identify the linguistic distance between two groups, languages are classified on a language tree. There are 15 different levels of classification reported by the Ethnologue, the 15th being the most precise considering two close dialects like Walloon and Picard as two different languages while the first level is aggregating worldwide spoken languages. Depending on the level chosen, the distance between two groups will be reflected by the distance between their languages on the tree. In the baseline analysis, level 5 is chosen because intermediate levels of linguistic aggregation were considered the most relevant for the provision of public goods by Desmet et al. (2012). I assumed it would be the same for the intensity of conflict. Other analyses will be discussed later with other levels of branching.

3.3. Spatial distribution of languages

DEA split the world into cells to evaluate the number of individuals speaking each language at a local level. It was considered that 5 km by 5 km cells were reasonably well reflecting this local component and this resolution is used in the baseline analysis. Different data sources were combined to develop their grid of cells: the World Language Mapping System from the Ethnologue (Lewis, 2014)

[1] The last period, 2005-2008, contains only 4 years.
[2] This is a joint dataset of the Uppsala University and the Peace Research Institute, Oslo. See Gleditsch et al. (2002) for a presentation of the dataset.

and Landscan[3]. With those two sources, DEA had the data about the number of people per cell, the number of speakers of each language per country, and whether in a precise cell, a language is spoken. Nevertheless, what they did not have was the number of speakers of each language in each cell. Several papers had constructed datasets of linguistic groups at a local level, but DEA used an iterative proportional fitting algorithm that allocates individuals to cells by considering the cells' populations and the country numbers of speakers for each language. With this algorithm, the problem of allocating the right share of speakers in cells where several languages are spoken is solved.

3.4. Descriptive statistics

In Table (1) and Table (2), a summary of the statistics of the main variables is displayed. The variable of interest, the local-global complementarity, the ethnic fractionalization and polarization, and the Greenberg-Gini index.

Table (2) is more interesting because shows the correlation coefficients between all those four variables. The coefficients are relatively low suggesting that the variables are independent and non-correlated with each other. Hence, they are relevant to use together for this analysis.

Table 1: Summary Statistics

Variables	Mean	Std.Dev.	Min.	Max.
LGC5	0.047	0.055	0	0.25
POL	0.046	0.054	0	0.246
FRAC	0.417	0.245	0	0.842
GINI	0.044	0.129	0	1.775

Table 2: Correlations

Variables	LGC5	POL	FRAC	GINI
LGC5	1.000			
POL	0.287	1.000		
FRAC	0.344	0.205	1.000	
GINI	0.211	0.401	0.142	1.000

3.5. Additional variables

In this section, I present the set of controls used in the baseline and the extensions. Those controls were the ones already used by Montalvo and Reynal-Querol (2005) and by EEA; Log population (POP); log per capita GDP (GDPPC); a dummy for the production of oil/diamond (OILDIAMOND); the percentage of mountainous terrain in the country (MOUNT); the noncontiguity of country's regions (NCONT); the presence of democracy (DEMOC). Moreover, some additional governance controls are also added: the lack of executive constraints (EXCONS); the autocracy level (AUTOCR); the extent of suppression of civil liberties (CIVLIB), and political rights (POLRIGHT). Those additional variables are taken from Polity IV and Freedom House. The values of the coefficients are the ones from the first year of each five-year period.

[3] See the data from the Oak Ridge National Laboratory (1997)

4. Baseline analysis

The dependent variable PRIO25A measures the intensity of conflict based on the number of deaths related to a conflict. There are several levels of intensity in the PRIO dataset and PRIO25A is the lowest of them indicating when a conflict is reaching 25 battle deaths in the five-year period. In each column of Table (3) we are going to find the ethnic polarization, the ethnolinguistic fractionalization, a Greenberg-Gini index of ethnic difference, the population control variables, and the variable of interest: the local-global complementarity variable. In column 2, I added the lagged PRIO25A, and in column 3 the GDP per capita. Column 4 adds a dummy for the presence of oil or diamond production in the country. Column 5 adds geographical variables: the presence of mountains and non-contiguous regions. Column 6 adds the dummy for democracy. And finally, more specific political controls are added in column 7. Besides the local-global complementarity, the baseline is the same as in EEA.

The coefficient of POL is positive and highly significant in all columns. This means that the publicness of the prize and the group cohesion are important and as they increase, the intensity of conflict increases. Since in most conflicts, winners yield a public prize, the fact that the coefficient of POL is positive is not surprising. The coefficient of FRAC is also positive but significant in only 5 columns out of 7. This suggests that the privateness of the prize is also important in explaining conflict but to a lesser degree than the publicness reflected by the ethnolinguistic polarization. The coefficient of GINI is negative and statistically significant throughout the regressions.

The variable of interest, the local-global complementarity is positive and even though it is not statistically significant in the two first columns, it becomes significant in column 3 when the per capita GDP is introduced. In columns 4 and 7, the coefficient is not significant with the conventional levels but is right above the 10% of significance. Using the theory from DEA that I summarized in the conceptual framework, a possible explanation can be provided for the sign of the local-global complementarity. Even though it is not statistically significant in all the regressions, the fact that it has a positive value is interesting. It means, following the theory, that the value of β is positive suggesting that interactions with other groups at the local level impacts positively the overall antagonism felt towards those groups in the society at large leading to an increase in the intensity of conflict.

In the literature, papers like Duclos et al. (2004) and Esteban and Ray (2011) were putting forward the use of ethnic polarization to evaluate conflict. In this work, I analyze the breadth of the impact of local-global complementarity on conflict. Using a marginal effect analysis, I examine which of those two variables has the higher effect on the intensity of conflict. The results obtained show that polarization has a higher and more robust effect than local-global complementarity.

Then, the population control is highly significant and positive but does not have a huge impact on conflict. The lagged conflict coefficient, on the other hand, while highly significant is largely impacting conflict positively. It might be obvious but passed conflicts seems to lead to future conflicts. On the other hand, the per capita GDP influences conflict negatively and is also highly significant suggesting that a richer country will endure less conflict.

The production of oil and diamonds seems to not affect the incidence of conflict as the control for democracy and other political measures of autocracy. However, results show that the presence of mountainous terrains and non-contiguous regions will impact conflict positively.

Table 3: Baseline Specification

	(1) PRIO25A	(2) PRIO25A	(3) PRIO25A	(4) PRIO25A	(5) PRIO25A	(6) PRIO25A	(7) PRIO25A
POP	0.209***	0.117**	0.144***	0.142***	0.0906*	0.0908*	0.0917*
	(0.0687)	(0.0470)	(0.0525)	(0.0517)	(0.0498)	(0.0499)	(0.0516)
POL	4.126**	3.099***	3.483***	3.522***	3.699***	3.733***	3.404***
	(1.677)	(1.121)	(1.077)	(1.129)	(1.242)	(1.240)	(1.307)
FRAC	1.439***	0.974***	0.469	0.452	0.553*	0.568**	0.572*
	(0.373)	(0.258)	(0.305)	(0.305)	(0.283)	(0.284)	(0.298)
GINI	-3.312*	-2.925***	-2.436*	-2.466*	-2.671*	-2.744*	-2.882*
	(1.846)	(1.089)	(1.382)	(1.378)	(1.479)	(1.468)	(1.531)
LGC5	1.388	1.084	2.093*	2.077	2.363*	2.430*	2.102
	(1.785)	(1.166)	(1.270)	(1.288)	(1.335)	(1.340)	(1.332)
LAG		1.696***	1.631***	1.630***	1.581***	1.580***	1.617***
		(0.116)	(0.122)	(0.122)	(0.122)	(0.122)	(0.129)
GDPPC			-0.242***	-0.245***	-0.292***	-0.277***	-0.213***
			(0.0631)	(0.0648)	(0.0665)	(0.0709)	(0.0757)
OILDIAMOND				0.0285	0.0257	0.0129	-0.0685
				(0.109)	(0.113)	(0.116)	(0.122)
MOUNT					0.00385*	0.00377*	0.00397
					(0.00220)	(0.00221)	(0.00242)
NCONT					0.523**	0.536**	0.560***
					(0.212)	(0.210)	(0.207)
DEMOC						-0.0652	-0.0108
						(0.139)	(0.185)
EXCONS							-0.0532
							(0.214)
AUTOCR							0.0612
							(0.152)
POLRIGHT							0.170
							(0.183)
CIVLIB							0.0628
							(0.210)
_cons	-4.830***	-3.655***	-2.020**	-1.963**	-0.959	-1.052	-1.671*
	(1.125)	(0.765)	(0.870)	(0.860)	(0.849)	(0.863)	(0.988)
N	1248	1113	1089	1089	1089	1089	981

Standard errors in parentheses
* p¡0.1, ** p¡0.05, *** p¡0.01

5. Extensions

To challenge the robustness of the baseline analysis, some variations are examined in the following section. I analyze alternative measures of conflict (5.1.); alternative ethnic groups classification (5.2.); larger cell sizes (5.3.); higher and lower linguistic aggregation levels (5.4.); and an instrument for the LGC (5.5.). Furthermore, alternative results without ethnic polarization can be found in the appendix.

5.1. Alternative measures of Conflict

In the baseline, the dependent variable measuring conflict is the binary variable PRIO25A, but other variables can also be used to measure conflict.

In Table (4), column 1 is a repetition of column 6 from Table (3) used for comparison with the alternative measures of conflict. In column 2, the dependent variable is PRIOcwA which is a binary variable taking the value of 1 if the overall conflict has reached 1000 deaths since onset. Column 3 is using the highest-level variable from the PRIO dataset: PRIO1000A. This variable demands 1,000 deaths per year and is considered as “civil war”. Column 4 employs PRIOINT, a nonbinary variable based on the PRIO dataset that could take the value 0 when the 25-deaths per year threshold is not attained, 1 when it is reached, and 2 when conflicts yield 1000 deaths per year. In column 5, I use the continuous index of social conflict (ISC) computed by the Cross-National Time-Series Data Archive (CNTS) and used by Banks (2008) as an alternative measure of conflict intensity. A weighted average over eight different manifestations of internal conflict is used to form the index ISC, adopted by Rummel (1963). The lagged variables of conflict are adjusted for every regression in addition to the same controls used in the baseline.

The polarization coefficient is highly significant and positive; FRAC is positive but not significant in most of the columns as for the local-global complementarity; GINI has its expected negative coefficient but is not significant in most of the columns either. As it is interesting to have different dependent variables as proxies for conflict, I must recognize that some are not that useful for this analysis because they do not capture every conflict. The PRIO25A registers all conflicts that have more than 25 battle deaths in a year and is thus better fitting in this analysis that has for objective to assess the impact of local-global complementarity on every conflict. On the other hand, the other variables like PRIOcwA or PRIO1000A can be useful too in other analyses more focused on bigger conflicts for example but here I cannot afford to ignore conflicts that do not make hundreds of deaths.

5.2. Alternative Groupings

EEA trusted the Fearon classification of ethnic groups (Fearon 2003 a) and I use it in the baseline but in some cases like Papua where no linguistic group reaches one percent of the population, it appeared that the classification of the ethnic groups needed the researcher intervention to define which ethnic groups are the relevant ones. Hence, the information from the Ethnologue about linguistic groups is taken as an alternative way of measuring the number of groups. When Fearon identified more than 800 “ethnic and ethnoreligious” groups, the Ethnologue reports 6,912 known living languages and the size of each of them in each country. Even though nowadays, conflicts are not happening between the groups reported by the Ethnologue anymore, the linguistic distance is a good proxy for the difference over the preferences for public good. Table (5) is the exact replication of Table (4) using the Ethnologue groups.

Polarization is still highly significant and positive but the GINI, FRAC, and local-global coefficients are not significant anymore. It is not surprising that coefficients are no longer significant because the groups following the Ethnologue data lack the intergroup distance that we could find in the ethnic groups formed by Fearon (2003 a). Hence, group formation is less robust.

Table 4: Alternative Measures of Conflict, Fearon Groupings

	(1) PRIO25A	(2) PRIOcwA	(3) PRIO1000A	(4) PRIOmult2A	(5) domestic9_meanA
POL	3.733***	3.199**	5.086***	3.500***	26.01***
	(1.240)	(1.353)	(1.677)	(1.099)	(8.876)
FRAC	0.568**	0.443	0.533	0.603**	2.012
	(0.284)	(0.328)	(0.358)	(0.269)	(1.909)
GINI	-2.744*	-2.798	-5.959*	-2.791*	-1.780
	(1.468)	(2.610)	(3.444)	(1.445)	(2.490)
LGC5	2.430*	3.144**	0.904	1.785	4.339
	(1.340)	(1.575)	(1.662)	(1.187)	(7.441)
GDPPC	-0.277***	-0.231***	-0.331***	-0.241***	-1.701***
	(0.0709)	(0.0859)	(0.0859)	(0.0627)	(0.500)
POP	0.0908*	0.123**	0.0776	0.0761*	1.158***
	(0.0499)	(0.0509)	(0.0567)	(0.0435)	(0.244)
OILDIAMOND	0.0129	0.0359	0.0187	-0.0102	-0.498
	(0.116)	(0.151)	(0.164)	(0.106)	(0.802)
MOUNT	0.00377*	0.00578**	0.00300	0.00302	0.0382**
	(0.00221)	(0.00248)	(0.00284)	(0.00199)	(0.0159)
NCONT	0.536**	0.401*	0.462**	0.382**	4.439***
	(0.210)	(0.223)	(0.229)	(0.180)	(1.614)
DEMOC	-0.0652	-0.0726	-0.206	-0.0761	0.104
	(0.139)	(0.162)	(0.175)	(0.124)	(0.921)
LAG	1.580***	2.121***	1.569***	1.113***	0.495***
	(0.122)	(0.155)	(0.135)	(0.0787)	(0.0371)
_cons	-1.052	-2.341**	-0.864		-3.586
	(0.863)	(0.945)	(1.171)		(4.544)
cut1				0.977	
				(0.786)	
cut2				2.131***	
				(0.818)	
N	1089	1089	1089	1089	1075

Standard errors in parentheses
* p¡0.1, ** p¡0.05, *** p¡0.01

5.3. Alternative cell sizes (10km x 10km)

In the baseline, regressions were made using the local-global complementarity in 5 km by 5 km cells because it was determined by DEA to be the reasonable size to evaluate the “local” component of an individual’s life and the other individuals he was interacting with. Next, I am going to analyze the results when using the LGC in 10 km by 10 km cells. Table (6) reports those results.

Fundamentally, the results are not that different from the baseline analysis. The coefficients have similar values and significance which consolidates the idea that 5 km by 5 km cells are enough to perfectly reflect the local-global complementarity. Enlarging the size of the cells would not bring any value.

Table 5: Alternative Measures of Conflict, Ethnologue Groupings

	(1) PRIO25A	(2) PRIOcwA	(3) PRIO1000A	(4) PRIOmult2A	(5) domestic9_meanA
POL	3.650**	3.235**	4.347*	3.494**	25.79**
	(1.469)	(1.629)	(2.286)	(1.396)	(10.91)
FRAC	0.126	0.153	0.104	0.103	-0.864
	(0.253)	(0.297)	(0.314)	(0.236)	(1.632)
GINI	-1.326	-1.020	-3.368	-1.330	-0.222
	(0.883)	(1.137)	(2.791)	(0.812)	(2.326)
LGC5	2.459	2.782	1.664	2.003	3.608
	(1.635)	(1.829)	(1.936)	(1.380)	(8.763)
GDPPC	-0.317***	-0.254***	-0.354***	-0.287***	-2.054***
	(0.0676)	(0.0792)	(0.0810)	(0.0598)	(0.462)
POP	0.102*	0.142**	0.101	0.0880*	1.249***
	(0.0550)	(0.0561)	(0.0639)	(0.0482)	(0.293)
OILDIAMOND	0.0895	0.0998	0.0877	0.0665	0.0376
	(0.116)	(0.150)	(0.156)	(0.106)	(0.907)
MOUNT	0.00459**	0.00647**	0.00344	0.00388*	0.0433**
	(0.00234)	(0.00266)	(0.00301)	(0.00209)	(0.0171)
NCONT	0.462**	0.320	0.339	0.311*	4.228***
	(0.200)	(0.204)	(0.211)	(0.171)	(1.570)
DEMOC	-0.0224	-0.0225	-0.155	-0.0347	0.0820
	(0.139)	(0.159)	(0.171)	(0.127)	(0.953)
LAG	1.591***	2.148***	1.600***	1.125***	0.500***
	(0.126)	(0.151)	(0.134)	(0.0802)	(0.0350)
_cons	-0.783	-2.400**	-0.906		-1.223
	(0.897)	(0.997)	(1.149)		(4.806)
cut1				0.656	
				(0.802)	
cut2				1.797**	
				(0.829)	
N	1086	1086	1086	1086	1072

Standard errors in parentheses
* p¡0.1, ** p¡0.05, *** p¡0.01

5.4. Alternative levels of linguistic aggregation

In the baseline analysis languages are aggregated at level 5 out of 15 because an intermediate level of linguistic aggregation is the most relevant for the provision of public goods as elaborated by Desmet et al (2012). I assumed that it was the case for the analysis of the intensity of conflict. Nevertheless, I compute the same regressions as in the baseline but this time for levels 2 and 15 of linguistic aggregation, respectively Table (7) and Table (8), to see if the results would change. Level 2 is a very broad aggregation considering French, Italian, and the other Latin languages as part of the same linguistic group while level 15 is the most precise level of linguistic aggregation considering similar dialects like Walloon and Picard as different languages.

Table 6: Local-global Complementarity in 10km x 10km cells

	(1) PRIO25A	(2) PRIO25A	(3) PRIO25A	(4) PRIO25A	(5) PRIO25A	(6) PRIO25A	(7) PRIO25A
POP	0.209***	0.117**	0.145***	0.143***	0.0922*	0.0924*	0.0923*
	(0.0688)	(0.0471)	(0.0525)	(0.0518)	(0.0501)	(0.0502)	(0.0517)
POL	4.156**	3.116***	3.497***	3.532***	3.674***	3.706***	3.382***
	(1.676)	(1.118)	(1.074)	(1.125)	(1.241)	(1.240)	(1.306)
FRAC	1.450***	0.982***	0.466	0.450	0.548*	0.563**	0.575*
	(0.371)	(0.256)	(0.304)	(0.304)	(0.282)	(0.282)	(0.299)
GINI	-3.292*	-2.910***	-2.446*	-2.473*	-2.672*	-2.741*	-2.848*
	(1.838)	(1.086)	(1.380)	(1.377)	(1.480)	(1.467)	(1.522)
10kmLGC5	1.246	1.000	2.143*	2.127*	2.521*	2.578*	2.083
	(1.778)	(1.166)	(1.268)	(1.291)	(1.342)	(1.344)	(1.346)
LAG		1.696***	1.631***	1.630***	1.579***	1.579***	1.617***
		(0.116)	(0.122)	(0.123)	(0.122)	(0.122)	(0.129)
GDPPC			-0.246***	-0.248***	-0.296***	-0.282***	-0.217***
			(0.0631)	(0.0647)	(0.0660)	(0.0702)	(0.0754)
OILDIAMOND				0.0259	0.0237	0.0113	-0.0672
				(0.110)	(0.113)	(0.116)	(0.122)
MOUNT					0.00412*	0.00404*	0.00415*
					(0.00218)	(0.00220)	(0.00242)
NCONT					0.526**	0.539**	0.560***
					(0.213)	(0.212)	(0.208)
DEMOC						-0.0628	-0.00974
						(0.138)	(0.184)
EXCONS							-0.0484
							(0.213)
AUTOCR							0.0578
							(0.152)
POLRIGHT							0.161
							(0.184)
CIVLIB							0.0625
							(0.208)
_cons	-4.824***	-3.654***	-2.015**	-1.963**	-0.961	-1.050	-1.648*
	(1.126)	(0.767)	(0.871)	(0.861)	(0.852)	(0.865)	(0.986)
N	1248	1113	1089	1089	1089	1089	981

Standard errors in parentheses
* p¡0.1, ** p¡0.05, *** p¡0.01

5.4.1. Level 2 of Linguistic Aggregation

Polarization and fractionalization are highly significant and positive; GINI is significant and negative; the local-global complementarity coefficient at level 2 of linguistic aggregation is not significant. The LGC coefficient not being significant is consistent with the theory from Desmet et al. (2012). Next, I analyze if that theory is also true for the most precise level of aggregation.

Table 7: Linguistic Aggregation level 2

	(1) PRIO25A	(2) PRIO25A	(3) PRIO25A	(4) PRIO25A	(5) PRIO25A	(6) PRIO25A	(7) PRIO25A
POP	0.206***	0.116**	0.142***	0.139***	0.0889*	0.0895*	0.0911*
	(0.0693)	(0.0477)	(0.0531)	(0.0523)	(0.0502)	(0.0503)	(0.0525)
POL	4.231**	3.150***	3.524***	3.571***	3.766***	3.790***	3.446***
	(1.709)	(1.144)	(1.122)	(1.177)	(1.278)	(1.279)	(1.335)
FRAC	1.485***	1.003***	0.523*	0.499*	0.614**	0.626**	0.621**
	(0.369)	(0.257)	(0.300)	(0.302)	(0.283)	(0.283)	(0.292)
GINI	-3.233*	-2.889***	-2.421*	-2.461*	-2.636*	-2.720*	-2.898*
	(1.839)	(1.063)	(1.345)	(1.343)	(1.449)	(1.444)	(1.551)
LGC2	0.561	0.585	1.385	1.379	1.439	1.539	1.441
	(1.679)	(1.044)	(1.196)	(1.204)	(1.327)	(1.338)	(1.358)
LAG		1.698***	1.638***	1.637***	1.591***	1.591***	1.625***
		(0.117)	(0.122)	(0.123)	(0.125)	(0.125)	(0.130)
GDPPC			-0.234***	-0.238***	-0.279***	-0.265***	-0.198**
			(0.0629)	(0.0650)	(0.0690)	(0.0737)	(0.0801)
OILDIAMOND				0.0367	0.0350	0.0225	-0.0705
				(0.110)	(0.115)	(0.117)	(0.123)
MOUNT					0.00364	0.00356	0.00377
					(0.00223)	(0.00224)	(0.00241)
NCONT					0.506**	0.518**	0.545***
					(0.213)	(0.211)	(0.208)
DEMOC						-0.0650	-0.00800
						(0.139)	(0.185)
EXCONS							-0.0698
							(0.221)
AUTOCR							0.0719
							(0.153)
POLRIGHT							0.185
							(0.190)
CIVLIB							0.0757
							(0.213)
_cons	-4.768***	-3.621***	-2.042**	-1.968**	-1.011	-1.106	-1.774*
	(1.134)	(0.775)	(0.865)	(0.856)	(0.862)	(0.882)	(1.033)
N	1248	1113	1089	1089	1089	1089	981

Standard errors in parentheses
* p¡0.1, ** p¡0.05, *** p¡0.01

5.4.2. Level 15 of Linguistic Aggregation

The results of the analysis with the linguistic aggregation of LGC are in Table (8). Polarization and fractionalization are still highly significant and positive as before; GINI is not significant in most of the columns; the local-global complementarity coefficient at level 15 is not significant.

What is interesting here is that, although it is not significant, the LGC at level 15 of linguistic aggregation is negatively impacting conflict unlike the level 5 in the baseline and level 2 just above. This would suggest that local interactions with another group mitigate the overall antagonism felt towards this group in the society at large reducing the intensity of conflict. But since the coefficient is not statistically significant, I cannot give a real valid interpretation of it. Again, the fact that the LGC coefficient is not significant is consistent with the theory and consolidates the idea that an intermediate level of aggregation is the most relevant for the analysis of conflict intensity.

Table 8: Linguistic Aggregation level 15

	(1) PRIO25A	(2) PRIO25A	(3) PRIO25A	(4) PRIO25A	(5) PRIO25A	(6) PRIO25A	(7) PRIO25A
POP	0.189***	0.0965**	0.123**	0.121**	0.0695	0.0694	0.0721
	(0.0670)	(0.0470)	(0.0510)	(0.0503)	(0.0484)	(0.0484)	(0.0504)
POL	4.406***	3.369***	3.771***	3.811***	4.046***	4.072***	3.638***
	(1.648)	(1.121)	(1.091)	(1.145)	(1.211)	(1.210)	(1.281)
FRAC	1.519***	1.047***	0.646**	0.625**	0.750***	0.761***	0.739**
	(0.346)	(0.239)	(0.279)	(0.288)	(0.273)	(0.273)	(0.288)
GINI	-2.896	-2.542**	-2.015	-2.059	-2.272	-2.315*	-2.399
	(1.847)	(1.092)	(1.312)	(1.312)	(1.393)	(1.379)	(1.459)
LGC15	-2.142	-2.422	-1.345	-1.321	-1.330	-1.327	-1.441
	(2.257)	(1.504)	(1.698)	(1.689)	(1.823)	(1.826)	(2.081)
LAG		1.694***	1.643***	1.642***	1.596***	1.596***	1.627***
		(0.119)	(0.126)	(0.127)	(0.131)	(0.131)	(0.139)
GDPPC			-0.211***	-0.215***	-0.256***	-0.247***	-0.180**
			(0.0635)	(0.0660)	(0.0704)	(0.0733)	(0.0775)
OILDIAMOND				0.0326	0.0295	0.0222	-0.0658
				(0.111)	(0.116)	(0.117)	(0.121)
MOUNT					0.00336	0.00330	0.00337
					(0.00218)	(0.00219)	(0.00240)
NCONT					0.512**	0.520**	0.540***
					(0.205)	(0.205)	(0.202)
DEMOC						-0.0382	0.0106
						(0.137)	(0.181)
EXCONS							-0.0650
							(0.217)
AUTOCR							0.0589
							(0.162)
POLRIGHT							0.143
							(0.188)
CIVLIB							0.112
							(0.214)
_cons	-4.408***	-3.220***	-1.877**	-1.813**	-0.837	-0.889	-1.548
	(1.092)	(0.762)	(0.843)	(0.841)	(0.854)	(0.862)	(0.980)
N	1248	1113	1089	1089	1089	1089	981

Standard errors in parentheses
* p¡0.1, ** p¡0.05, *** p¡0.01

5.5. Instrumental Variable Approach

In the baseline, I use the local-global complementarity variable from DEA. But the paper reports a suspicion of endogeneity between the population spatial mixing and the public good. Individuals coming from the same ethnolinguistic group would cluster to feel safer when facing bad public goods provision. This implies that public goods provision has an impact on spatial mixing while we try to evaluate the opposite. Even though this presence of endogeneity is challenged with solid arguments it could not be eliminated and led to the construction of an instrument for the local-global complementarity. As in their paper, they were evaluating for endogeneity with public goods, this instrument can still be used to analyze the link with civil conflicts.

The instrument relies on languages used in neighboring countries. For each cell of a country, the closest cell in a neighboring country is taken, and if a language is spoken there and, in the country, then the language is assigned to the cell. If a language is not spoken in any of the closest cells in the neighboring countries, it is assumed that it is spoken in all cells of the country. Results of the use of this instrumental variable can be found in Table (9).

Throughout, the instrumental local-global complementarity is highly significant and positive and has a larger value than in the baseline. POL is also highly significant and positive; FRAC and GINI are respectively positive and negative but are only significant in the first two columns.

When analyzing the marginal effects of ethnic polarization and instrumental local-global complementarity on conflict, this time the latter has a larger impact. Results show that both coefficients have a large positive impact on conflict and are robust, but the instrumental LGC value is larger. This analysis implies through the instrument that, an increase in interactions with other groups at a local level will increase the overall antagonism felt towards those same groups in the society at large, increasing the intensity of conflict.

Table 9: Instrumental LGC5

	(1) PRIO25A	(2) PRIO25A	(3) PRIO25A	(4) PRIO25A	(5) PRIO25A	(6) PRIO25A	(7) PRIO25A
IVLGC5	5.199***	3.374**	4.374***	4.355***	4.468***	4.532***	4.100**
	(1.284)	(1.525)	(1.626)	(1.627)	(1.658)	(1.667)	(1.745)
POP	0.235***	0.132***	0.163***	0.161***	0.108**	0.109**	0.110**
	(0.0366)	(0.0450)	(0.0467)	(0.0474)	(0.0512)	(0.0512)	(0.0536)
POL	3.278***	2.643**	3.100***	3.140***	3.270***	3.309***	3.022**
	(0.890)	(1.063)	(1.086)	(1.096)	(1.161)	(1.163)	(1.288)
FRAC	1.232***	0.839***	0.290	0.273	0.388	0.407	0.408
	(0.194)	(0.236)	(0.269)	(0.277)	(0.283)	(0.283)	(0.305)
GINI	-3.205***	-2.923**	-2.282	-2.309	-2.478	-2.563	-2.680
	(1.037)	(1.383)	(1.666)	(1.668)	(1.731)	(1.734)	(1.864)
LAG		1.699***	1.623***	1.622***	1.571***	1.571***	1.608***
		(0.105)	(0.106)	(0.106)	(0.107)	(0.107)	(0.114)
GDPPC			-0.264***	-0.267***	-0.310***	-0.291***	-0.232***
			(0.0555)	(0.0569)	(0.0602)	(0.0654)	(0.0803)
OILDIAMOND				0.0283	0.0304	0.0137	-0.0611
				(0.112)	(0.113)	(0.116)	(0.129)
MOUNT					0.00457*	0.00447*	0.00469*
					(0.00259)	(0.00259)	(0.00282)
NCONT					0.516***	0.534***	0.564***
					(0.173)	(0.175)	(0.188)
DEMOC						-0.0849	-0.0240
						(0.122)	(0.146)
EXCONS							-0.0430
							(0.188)
AUTOCR							0.0641
							(0.161)
POLRIGHT							0.177
							(0.230)
CIVLIB							0.0393
							(0.210)
_cons	-5.304***	-3.915***	-2.176**	-2.120**	-1.126	-1.249	-1.827*
	(0.618)	(0.751)	(0.847)	(0.876)	(0.959)	(0.977)	(1.088)
N	1248	1113	1089	1089	1089	1089	981
First-Stage F-Stat.	353.76	271.37	219.85	192.2	156.53	143.62	99.61

Standard errors in parentheses
* p¡0.1, ** p¡0.05, *** p¡0.01

6. Conclusion

This paper examines the impact of local interactions on civil conflict in 5 km by 5 km cells through the period 1960-2008. Firstly, I implemented a variable created by DEA, the local-global complementarity, in the framework from EEA. This variable evaluates the probability that when three individuals are randomly drawn with 1 and 2 from the same cell and 3 from the society at large, 2 and 3 belong to the same group and 1 to a different one. The size of the cells, 5 km by 5 km was chosen because considered to be reasonable to capture the local component of interactions. Then to define groups and the distance between them, I followed the literature e.g., Fearon (2003a) and Desmet et al. (2012) which used the linguistic distances as proxies. Linguistic differences vary following the level of aggregation chosen on the language tree from the Ethnologue. An intermediate level was considered to be the most relevant by Desmet et al (2012) and hence, used in this work.

When proceeding to the empirical analysis, my results show that an increase in local interactions with other groups increases the overall antagonism felt towards those groups in the society in general which increases conflict intensity. But since the local-global complementarity coefficient was not significant in most of the specifications, those results cannot be considered highly robust.

Through the different extensions, I challenge the results of the baseline analysis. The alternative measures of conflict intensity were less relevant and the groupings from the Ethnologue were less robust since they were lacking the intergroup distance. Then, I questioned the cells' size chosen by DEA by using 10 km by 10 km cells. However, larger cells were not producing significantly different results consolidating the idea that 5 km by 5 km was the right size to assess local interactions. Next, I compute the same regressions but change the linguistic aggregation level from the Ethnologue. I use level 2 which is very coarse and considers very different languages as part of the same group and level 15 which is the most precise level differentiating dialects. Both levels were not producing significant results thus I consider that an intermediate level of linguistic aggregation stays more relevant.

In the last extension, I use an instrument for the local-global complementarity elaborated by DEA. This instrument relies on the closest cells of the neighboring countries to define the shares of speakers of each language in each cell. The results with the instrumentalized variable are far more significant and consistent with local interactions increasing conflict. Furthermore, the LGC effect on conflict is even higher than the effect of polarization suggesting that local interactions impacting overall antagonism are useful in the explanation of conflict intensity and should not be omitted.

In this paper, local interactions increase conflict but with the level 15 of linguistic aggregation, the LGC had a negative sign even though the coefficient was not significative. More research is needed on the different linguistic aggregation level to observe at what level does the local-global complementarity coefficient have a negative value and with what controls.

Appendix

Analysis without ethnic polarization

In the next table, the ethnic polarization has been taken out of the equation. Fractionalization is significant in all the regressions while LGC is significant in most of them, but both still have their positive coefficient. GINI is not significant in most of the specifications while lagged conflict and per capita GDP are still highly significant and respectively positive and negative. Without ethnic polarization, we can observe that the local-global complementarity coefficient has a higher positive value and is more significant than in the baseline.

Table 10: Without Polarization

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	PRIO25A	PRIO25A	PRIO25A	PRIO25A	PRIO25A	PRIO25A	PRIO25A
POP	0.255***	0.155***	0.186***	0.187***	0.143***	0.143***	0.133***
	(0.0636)	(0.0439)	(0.0496)	(0.0489)	(0.0485)	(0.0486)	(0.0506)
FRAC	1.488***	0.997***	0.522*	0.539*	0.635**	0.646**	0.613**
	(0.368)	(0.251)	(0.296)	(0.298)	(0.279)	(0.280)	(0.295)
GINI	-1.414	-1.322*	-0.392	-0.381	-0.369	-0.401	-0.804
	(1.172)	(0.794)	(0.876)	(0.879)	(0.901)	(0.907)	(1.067)
LGC5	2.093	1.626	2.644**	2.653**	2.952**	3.002**	2.485**
	(1.688)	(1.126)	(1.173)	(1.176)	(1.193)	(1.194)	(1.210)
LAG		1.727***	1.669***	1.670***	1.624***	1.624***	1.664***
		(0.116)	(0.119)	(0.119)	(0.121)	(0.121)	(0.126)
GDPPC			-0.224***	-0.221***	-0.256***	-0.246***	-0.191***
			(0.0592)	(0.0598)	(0.0599)	(0.0648)	(0.0720)
OILDIAMOND				-0.0282	-0.00670	-0.0160	-0.0697
				(0.105)	(0.109)	(0.109)	(0.117)
MOUNT					0.00592***	0.00587***	0.00543**
					(0.00226)	(0.00226)	(0.00233)
NCONT					0.428**	0.436**	0.463**
					(0.193)	(0.193)	(0.194)
DEMOC						-0.0447	-0.0200
						(0.141)	(0.181)
EXCONS							-0.0250
							(0.218)
AUTOCR							0.0419
							(0.158)
POLRIGHT							-0.0143
							(0.219)
CIVLIB							0.191
							(0.238)
_cons	-5.485***	-4.206***	-2.788***	-2.836***	-2.064***	-2.136***	-2.450***
	(1.051)	(0.718)	(0.770)	(0.753)	(0.755)	(0.760)	(0.928)
N	1248	1113	1089	1089	1089	1089	981

Standard errors in parentheses
* p¡0.1, ** p¡0.05, *** p¡0.01

References

Banks, A. S. (2008). Cross-National Time-Series Data Archive (CNTS) 1815–2007. Jerusalem, Israel: Databanks International. (last accessed in May 2022)

Collier, P. A. (1998). On economic causes of civil war. *Oxford Economic Papers, 50* (4), pp. 563–573.

Collier, P. A. (2004). Greed and Grievance in Civil War. *Oxford Economic Papers, 56* (4), pp. 563-95.

Desmet, K. G.-O. (2020). The geography of linguistic diversity and the provision of public goods. *Journal of Development Economics, 143*.

Desmet, K. I.-O. (2009). Linguistic Diversity and Redistribution. *Journal of the European Economic Association 7* (6), pp. 1291-318.

Desmet, K. I.-O. (2012). The Political Economy of Linguistic Cleavages. *Journal of Development Economics 97* (1), pp. 322-38.

Duclos, J.-Y. J. (2004). Polarization: Concepts, Measurements, *Econometrica, 72* (6), pp. 1737-72.

Esteban, J. D. (2011). Linking Conflict to Inequality and Polarization. *American Economic Review, 101* (4), pp. 1345–74.

Esteban, J. L. (2012). Ethnicity and Conflict: An Empirical Study. *American Economic Review, 102*, 1310-42.

Fearon, J. D. (2003 a). Ethnic and Cultural Diversity by Country. *Journal of Economic Growth, 8* (2), pp. 195-222.

Fearon, J. D. (2003 b). Ethnic and Cultural Diversity by Country: Dataset. *Journal of Economic Growth*. (last accessed in May 2022)

Fearon, J. D. (2003). Ethnicity, Insurgency, and Civil War. *American Political Science Review, 97* (1), pp. 75–90.

Freedom House. (2011). Freedom in the World 1973–2012. (last accessed in May 2022), on ” http://www.freedomhouse.org

Gleditsch, N. P. (2002). Armed Conflict 1946–2001: A New Dataset. *Journal of Peace Research 39*, pp. 615-37. (last accessed in May 2022)

Greenberg, J. H. (1956). The Measurement of Linguistic Diversity. *Language, 32* (1), pp. 109-15.

Laitin, D. D. (2000). What Is a Language Community? *American Journal of Political Science, 44* (1), pp. 142–55.

Lewis, P. M. (2014). *Ethnologue: Languages of the World* (éd. 17th). Dallas: SIL International. (last accessed in May 2022)

Lewis, P. M. (ed. 2009). *Ethnologue: Languages of the World* (éd. 16th). Dallas: SIL International. http://www.ethnologue.com/

Miguel, E. S. (2004). Economic Shocks and Civil Conflict: An Instrumental Variables Approach. *Journal of Political Economy, 112* (4), pp. 725–53.

Montalvo, J. G.-Q. (2005). Ethnic Polarization, Potential Conflict, and Civil Wars. *American Economic Review, 95* (3), pp. 796-816.

ORNL, O. R. (1997). *Landscan*. (last accessed in May 2022), on Landscan: https://landscan.ornl.gov/

PolityIV. (1800-2009). Polity IV Project: Political Regime Characteristics and Transitions. (last accessed in May 2022), on https://www.systemicpeace.org/polity/polity4.htm

Rummel, R. J. (1963). Dimensions of Conflict Behavior Within and Between Nations. *General System Yearbooks, 8* , pp. 1-50.