Ethnic diversity and conflict

Published in:
Journal of Institutional Economics

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Ethnic Diversity and Conflict

MICHAEL BLEANEY[1]

School of Economics, University of Nottingham, Nottingham, UK

(michael.bleaney@nottingham.ac.uk)

ARCANGELO DIMICO[2]

Queen's Management School, Queen's University Belfast, Belfast, UK

(a.dimico@qub.ac.uk)

Word Count: 8530

Abstract

We argue that the reason why it has proved hard to determine whether negative effects on economic performance and conflict are more strongly associated with polarized rather than fractionalized societies is because the distinction between polarization and fractionalization is only relevant for societies with ethnic diversity above a certain threshold. In addition high levels of ethnic fractionalization at a country level are generally associated with regional concentration of minorities, and as a result many regions may have a very different ethnic composition from the national average, and in particular they may have much higher levels of ethnic polarization than the national level. Because of the very different ethnic composition of different regions in this situation, conflict is more likely to be confined to a limited geographical area.

Keywords: Fractionalization, Polarization, Ethnic Concentration, Conflict

JEL Code: O10, F50

[1] Michael Bleaney, University of Nottingham, School of Economics, University Park, Nottingham, NG72RD. email: michael.bleaney@nottingham.ac.uk

[2] Arcangelo Dimico (QUCEH, CHaRMS), Queen's University Belfast, Management School, Riddell Hall, Belfast, BT95EE, UK. email: a.dimico@qub.ac.uk

1. Introduction

There is considerable evidence that greater ethnic diversity is associated with weaker economic performance. Alesina *et al*. (2003), Easterly and Levine (1997) and Montalvo and Reynal-Querol (2005b) find a negative effect of measures of ethnic diversity on economic growth. Alesina *et al.* (1999) and Miguel and Gugerty (2005) show that local supply of public goods is adversely affected when the population is ethnically heterogeneous, Costalli, Moretti and Piscedda (2014) use a synthetic control method to assess the national-level economic impact of civil war in a sample of 20 countries.

A common explanation for the negative effect of ethnic diversity on economic performance is that more ethnically diverse societies suffer more serious internal conflicts. Barro (1991) shows that revolutions and coups have a strong negative effect on growth. The empirical association between ethnic diversity and armed conflict has been quite extensively investigated. Bleaney and Dimico (2011) present evidence of a significant correlation, although Hegre and Sambanis (2006) conclude that ethnic diversity is robustly correlated only with the probability of small-scale conflict[3]. There is some evidence that ethnic diversity makes conflicts harder to stop (Collier *et al*., 2004; Fearon, 2004), rather than that it makes them more likely to start (Collier and Hoeffler, 1998, 2004; Collier *et al*., 2009; Fearon and Laitin, 2003). Kanbur *et al.* (2011) provide an in-depth survey of the relationship between ethnic diversity and ethnic strife. Blattman and Miguel (2010) also cover the topic extensively[4].

A persistent – and so far unresolved – matter of debate has been whether it is ethnic diversity *per se* or a particular pattern of diversity that is important. Montalvo and Reynal-Querol (2005a) argue that conflict is associated with polarization between two sizeable ethnic groups rather than just the proliferation of groups as measured by a fractionalization

---

[3] There are a few studies which focus on linguistic rather than ethnic diversity (i.e., Wimmer and Cederman, 2009) which find either a significant effect of linguistic diversity on conflict or a negative relationship between linguistic polarization and conflict (Gardeazabal, 2011) .

[4] There are also studies which look at the impact of conflict on ethnic diversity (i.e., Rohner, Thoenig, and Zilibotti; 2013) and the impact of conflict linked to ethnicity on social conflict (i.e., Becchetti, Conzo, and Romeo; 2014).

index.[5] Some recent research has suggested that conflicts are particularly hard to stop in polarized societies (Montalvo and Reynal-Querol, 2010; Schneider and Wiesehomeier, 2010).

In this paper we investigate two possible reasons why it has proved hard to determine whether negative effects on economic performance are more strongly associated with polarized rather than fractionalized societies.

One is that, for purely mathematical reasons, the distinction between polarization and fractionalization is only relevant for societies with ethnic diversity above a certain threshold. Below this threshold the two measures are very highly positively correlated, and therefore statistically indistinguishable. If the major feature of the data is that ethnically homogeneous societies are significantly more peaceful than those with even quite a limited degree of heterogeneity, and beyond that the level of diversity makes little difference, then statistical results will be similar whether diversity is measured by fractionalization or polarization. The first hypothesis is therefore that ethnic diversity is largely uncorrelated with conflict, except that near-homogeneous societies are more peaceful than others. In other words the degree of ethnic diversity does not explain why one country with significant ethnic diversity is more prone to armed conflict than another country with significant ethnic diversity, but only why such countries are more conflict-prone than ethnically homogeneous countries.

The second hypothesis assumes that polarization is the critical aspect of ethnic diversity, but not necessarily just at the national level. In highly fractionalized societies, the national measure of polarization tends to be rather low, because of the limited population share of even the largest groups. However, if high levels of ethnic fractionalization are associated with regional concentration of minorities, as we show below to be typically the case, then many regions may have a very different ethnic composition from the national average, and in particular they may have much higher levels of ethnic

---

[5] A country is said to be polarized if the second largest group makes up a substantial share of the population and is not much smaller than the largest group.

polarization than the national level.[6] On the other hand, because of the very different ethnic composition of different regions, in this situation conflict is more likely to be confined to a limited geographical area. Although we do not have data on regional ethnic diversity, it is possible to test whether the geographical extent of conflict decreases at high levels of fractionalization, and whether minority groups are more likely to be engaged in rebellious activity when they are regionally concentrated.

Previous work that is quite closely related to this second hypothesis includes Toft (2003), Fearon and Laitin (2011) and Esteban and Ray (2011). As Toft notes, bargaining theory tells us that armed conflict is not usually the rational solution to disagreements, because it is so costly to both sides, but may occur under certain conditions, of which the relevant one in this case is the indivisibility of the prize. Toft argues that, where minorities are regionally concentrated rather than dispersed, they are much more likely to regard their territory as indivisible, which tends to create conflict with other groups in the locality. This theory suggests that regional concentration of minorities should be positively correlated with the probability of armed conflict. It say nothing about the overall intra-regional population distribution, and tends to suggest that fractionalization is more important than polarization, because higher fractionalization implies more minorities.

Fearon and Laitin (2011) argue that nearly a third of ethnic civil wars arise from migration, mainly by members of the dominant ethnic group, into areas historically inhabited by a minority group. They discuss a number of examples, mainly from Asia. This idea seems quite closely related to ours in that local polarization appears to be an important factor, but it is more restrictive in that they stress the element of migration.

Esteban and Ray (2011) hypothesize that either fractionalization or polarization can be relevant, depending on circumstances. They develop a theoretical model in which the relationship between the investment of resources in conflict and ethnic diversity depends on the nature of the prize for winning the conflict, and show that, in this model, ethnic

---

[6] Alternatively they may not, because some regions might be almost entirely inhabited by one ethnic group, but this case is most likely to occur when there is only one large minority (so that fractionalization is limited).

polarization stimulates conflict if the prize is a public good, whereas ethnic fractionalization (splintering into many groups) stimulates conflict if the prize is a private good.

The intuition for this result is as follows. The probability of winning increases with a group's population and its *per capita* investment. If the prize is a private good, the losers can be completely excluded and the winners take the entire prize. The *per capita* gain from winning is therefore inversely related to the winning group's size, so small groups are tempted to invest a lot in the conflict, even though their winning probability for a given *per capita* investment is low. In this case, ethnic fractionalization stimulates investment in conflict. If, on the other hand, the prize is a public good, which will be enjoyed also by the losers, the gain from winning consists, not in the quantity of the good obtained, but only in being able to choose one's preferred form of the public good. The *per capita* gain from winning is therefore independent of group size, and it only pays groups that can significantly improve their probability of winning at the margin to invest in conflict. This incentive is greatest for two roughly equally size groups (i.e. when polarization is high). Esteban *et al*. (2012) offer some supportive empirical evidence for this model.

This model suggests that either polarization or fractionalization may be associated with conflict, depending on circumstances. The purpose of the present paper is to offer a different reason why both polarization and fractionalization (measured at the national level) matter for conflict. As we show, ethnic groups are typically quite regionally concentrated. This is not surprising, since linguistic and cultural differences emerged historically mainly through geographical isolation. The implication of this is that the ethnic composition of regions of a country tends to differ quite markedly from the national average, and from each other. Conflict may therefore be quite localized, as is shown by geographically disaggregated data sets (i.e., Tollefsen *et al*., 2012). For example in Turkey conflict is mainly confined to areas with a substantial Kurdish-speaking population.

The hypothesis that we put forward here is that armed conflict is more common in polarized communities, whether they be polarized at the national level or at the local level.

5

Ethnic fractionalization is not intrinsically associated with conflict, but if ethnic groups are regionally concentrated (as they tend to be), fractionalized countries will tend to have pockets of local polarization. If the second largest group is a relatively small proportion of the national population, but a large proportion of the population of a region, the national measure of polarization will be relatively low, but the local measure of polarization in that region will be quite high, as will be the national measure of ethnic fractionalization (the standard measure of ethnic diversity, equal to one minus the Herfindahl index of concentration). This hypothesis would explain why ethnic fractionalization and polarization perform about equally well in predicting the presence of conflict in a country (i.e., in the model of Bleaney and Dimico, 2011): fractionalization picks up conflict associated with local (but not national) polarization, whilst both measures are low in ethnically homogeneous countries.

To test this hypothesis we use group-level data from the MAR dataset in order to have a direct test of the relationship between ethnic polarization at a subnational level (measured by the regional concentration of the group) and the probability of armed conflict, and also its geographical extent. Consistent with our hypothesis, we find that local concentration of groups increases with the level of national fractionalization, and because of that national fractionalization increases the probability of local pockets of polarization. Therefore countries with a higher level of national fractionalization tend to experience conflicts of smaller scale compared to ethnically polarized countries, which explains why ethnic fractionalization is more strongly correlated with conflict in datasets with a low death threshold (i.e. the UCDP/PRIO with a 25 battle-related death threshold) than in those with a relatively high death threshold (i.e. the Correlates of War with a 1,000 death threshold). Our explanation of such an effect relates to the probability that ethnic fractionalized countries are characterized by pockets of local polarization, which then tend to increase the probability of local and less extensive conflicts. This explanation is confirmed by the group-level analysis where we show a significant relationship between the national level of fractionalization and the probability of regional concentration of groups (and therefore the level of subnational polarization) which then has a negative effect on the

6

geographical extent of the conflict. These results are also consistent with evidence from Indonesia by Barron *et al.* (2009) who show a positive relation between local conflicts and ethnic clustering.

The paper is organised as follows. In Section Two we test the first hypothesis: that the correlation between conflict and measures of ethnic diversity is most evident towards the homogeneous end of the spectrum. In Section Three we show theoretically that ethnic fractionalization plus regional concentration of groups gives rise to pockets of local polarization. Section Four provides evidence that minority groups are indeed regionally concentrated. Section Five explores data from the Minorities at Risk Project on the regional concentration of minorities and their participation in rebellious activities. Regional concentration is associated with greater incidence of conflict, but less geographical extent. These results are consistent with our hypothesis that local polarization makes a country susceptible to local conflicts. Section Six concludes.

2. Conflict and Ethnic Diversity

The most commonly used measure of ethnic diversity is ethnic fractionalization, which is equal to one minus a Herfindahl index of concentration:

$$F = 1 - \sum_{i=1}^{n} \pi_i^2 \tag{1}$$

where π_i represents the population share of group i and n is the total number of groups. This measure has a minimum of zero, when there is just one group, and a maximum of one. In practice it is strongly negatively correlated with the share of the largest group.

Figure 1: Ethnic Fractionalization vs Ethnic Polarization
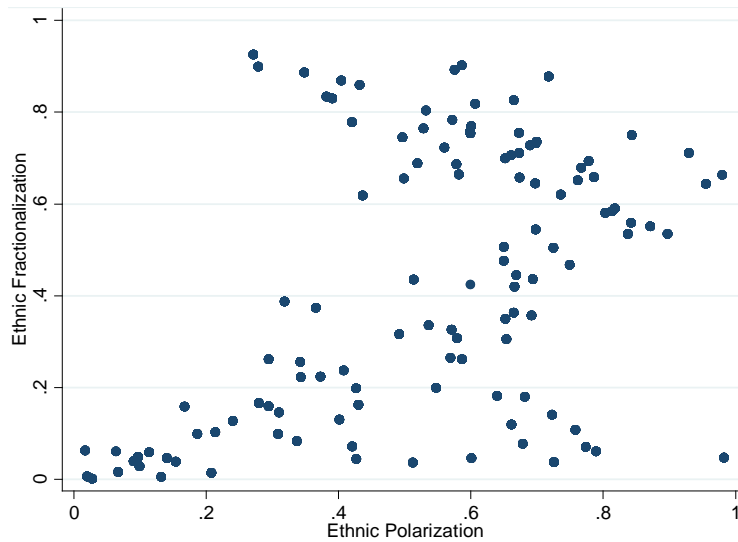
Esteban and Ray (1994) argue that conflict is most likely when the largest group is faced by a substantial minority group. Based on this, Montalvo and Reynal-Querol (2005a) use the following index of ethnic polarization, which is a particular case of the class of indices proposed by Esteban and Ray (1994):

$$P = 1 - \sum_{i=1}^{n} 4\pi_i (0.5 - \pi_i)^2 = 4\sum_{i=1}^{n} \pi_i^2 (1 - \pi_i) \tag{2}$$

Polarization reaches a maximum of one when there are just two equally sized groups, and is in practice strongly positively correlated with the share of the second largest group. As shown in Figure 1, polarization diminishes at high levels of fractionalization, and also as the share of the dominant group becomes very large. Polarization and fractionalization are strongly positively correlated when the two largest groups form a large share of the population, but tend to be negatively correlated if the share of the two largest groups is relatively low.[7]

Table 1. Conflict and Ethnic Diversity (Country-Year Data over the perido 1945-2004)

---

[7] If there are only two groups, P is precisely equal to twice F. If there are *n* groups of equal size, then as *n* tends to infinity, F tends to one but P tends to zero. In our sample the share of the two largest groups averages 81%. For cases where the share exceeds 81%, the correlation between F and P is 0.61, but for cases where the share is less than 81%, the correlation is only 0.09.

	Dependent Variable: Incidence of Conflict (UCDP/PRIO)					
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Full Sample		If Ethnic Frac >0.5		If Ethnic Frac < 0.5	
GDP p.c. (t-1)	-0.100***	-0.142***	-0.140***	-0.141***	-0.107***	-0.101***
	(-4.13)	(-6.01)	(-3.62)	(-3.73)	(-3.41)	(-3.17)
Population (t-1)	0.0715***	0.0896***	0.0600**	0.0504	0.113***	0.119***
	(3.24)	(3.97)	(2.01)	(1.63)	(3.24)	(3.34)
Ethnic Fractionalization	0.626***		0.389		1.725***	
	(5.58)		(1.17)		(5.22)	
Ethnic Polarization		0.726***		-0.479		1.166***
		(5.90)		(-1.61)		(5.49)
Mountainous Terrain	0.241*	0.0575	0.127	0.155	0.174	0.0743
	(1.89)	(0.45)	(0.79)	(0.94)	(0.79)	(0.33)
Onshore Oil Fields	0.124**	0.0477	0.187**	0.210**	-0.0327	-0.0656
	(1.96)	(0.75)	(2.22)	(2.44)	(-0.30)	(-0.59)
Anocracy (t-1)	0.108	0.124*	0.0242	0.0235	0.189	0.170
	(1.44)	(1.68)	(0.26)	(0.25)	(1.45)	(1.29)
Post-Cold War Dummy	-0.150	-0.130	-0.226	-0.236	-0.00811	-0.0621
	(-0.84)	(-0.73)	(-0.96)	(-1.01)	(-0.03)	(-0.23)
Time Trend	0.0115***	0.0115***	0.0171***	0.0172***	0.00680	0.00714
	(3.87)	(3.91)	(4.05)	(4.07)	(1.55)	(1.62)
Time Trend*Cold War Dummy	-0.00978**	-0.00925**	-0.0134**	-0.0135**	-0.00516	-0.00647
	(-2.16)	(-2.08)	(-2.19)	(-2.20)	(-0.75)	(-0.96)
Incidence (t-1)	2.028***	2.024***	1.963***	1.961***	2.085***	2.076***
	(25.73)	(25.72)	(19.86)	(19.82)	(15.97)	(15.83)
Incidence (t-2)	0.969***	0.971***	0.955***	0.950***	0.942***	0.939***
	(12.10)	(12.11)	(9.56)	(9.51)	(7.13)	(7.07)
Constant	-2.324***	-2.222***	-1.796***	-1.105**	-2.870***	-2.979***
	(-6.81)	(-6.64)	(-3.38)	(-2.37)	(-5.86)	(-5.87)
Observations	6,756	6,756	3,310	3,310	3,446	3,446

Estimation method: probit. Robust z-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1
The dependent is a binary variable coded 1 if for any country-year there is conflict and 0 if in such a country-year there is not conflict. Source of data: UCDP/PRIO

Table 1 shows the results of estimating a probit model for the incidence of conflict, based on the specification of Bleaney and Dimico (2011). Models 1 and 2 are identical except that Model 2 includes ethnic polarization in place of ethnic fractionalization. Both

measures have highly significant positive coefficients in the full sample. Then we divide the sample according to whether ethnic fractionalization is greater than 0.5 (Models 3 and 4) or less than 0.5 (Models 5 and 6). At higher levels of ethnic diversity (Models 3 and 4), neither fractionalization nor polarization is statistically significant, but at lower levels (Models 5 and 6), they are almost as statistically significant as in the full sample. These results dramatically confirm our first hypothesis, that polarization and fractionalization perform similarly in these empirical tests because the correlation is strongest at low levels of ethnic diversity where the two measures are highly positively correlated. In the remainder of the paper, we investigate the apparent statistical insignificance of ethnic variables at higher levels of diversity.

3. Ethnic Fractionalization and Polarization

The available data sets refer only to polarization (or fractionalization) at the national level. The relationship between this number and local polarization can be complex. Polarization at the local level will depend not just on the national population shares of each ethnic group but also on how geographically concentrated they are. If two large groups are geographically separated, it is possible that polarization could be low at the local level but high at the national level. On the other hand, if ethnic fractionalization (F) is high, in which case polarization (P) is necessarily quite low, there could be quite strongly polarized regions of the country. This is particularly the case as ethnic minorities are often geographically concentrated. For example, suppose that there are ten groups each representing exactly 10% of the population. Then $F = 0.90$ and $P = 0.36$, which is not very high. If, however, these groups are unequally distributed across the country, so that in each region there are only five groups each representing 20% of the population, then local polarization is 0.64 in every region, which implies that the average local polarization is considerably greater than the national polarization figure.

Thus the combination of high ethnic fractionalization and an uneven regional distribution of ethnic groups is likely to result in relatively high polarization levels in at least some regions. If polarization at the local as well as the national level increases the incidence of conflict, then high ethnic fractionalization will also tend to be associated with an increased incidence of local conflicts. We test this hypothesis using data on the proportion of the country affected by conflict in any given year.

To investigate the relationship between average local polarization and ethnic fractionalization more formally, consider cases where at least one ethnic group is not represented in any given region. We analyse two possibilities: (1) where the regional population shares are formed from the national population shares by adding the missing group's shares entirely to that of one other group; and (2) where the share of the missing group is redistributed equally amongst all the represented groups.

In the first case, as shown in the Appendix, the merging of any two groups with shares π_i and π_j will change polarization by

$$\Delta P = 4\pi_i \pi_j [2 - 3(\pi_i + \pi_j)] \tag{3}$$

which is positive if the two groups' shares sum to less than two-thirds, but negative if they sum to more than two-thirds. Once the number of groups exceeds three, and if the two largest groups make up less than two-thirds of the population, then all regions will have greater polarization than the national measure. Thus in this case it is highly likely that average local polarization exceeds national polarization when ethnic fractionalization is high.

Table 2. National and Local Polarization

Number of equal-sized	National fractionalization	National polarization	Regional polarization index (equal in every region)

groups	index	index		
			Each group	Each group
3	0.67	0.89	1	0
4	0.75	0.75	0.89	1
5	0.80	0.64	0.75	0.89
6	0.83	0.56	0.64	0.75
7	0.86	0.49	0.56	0.64
8	0.89	0.44	0.49	0.56

Note. It is assumed that a group is evenly distributed across regions where it is represented

Table 2 shows some numerical examples where there are n regions and n equally sized groups within the national population, but the groups are regionally concentrated to the extent of each being absent from either one or two regions. It is assumed that the groups that are present within a region are equally represented within the regional population. To take an example, when $n = 5$ and there are two groups absent from each region, the three groups that are present each make up one-third of the regional population. The table shows that as long as $n \geq 4$, so that ethnic fractionalization is relatively high, regional polarization exceeds national polarization.

In the second case, it is shown in the Appendix that elimination of group j and redistribution of its share equally amongst the n remaining groups results in a change of polarization of:

$$\Delta P = 4am^2[3F - 1 - a(1 + m + 3m^2) + a^2m(1 + 3m + 2m^2)] \tag{4}$$

where $m = n^{-1}$, $a = n\pi_j$ and F is the index of ethnic fractionalization defined in equation (1). The parameter a reflects the relative size of the excluded group j (which is equal to the average of the others when $a = n/(n+1)$). Equation (4) is more likely to be positive when F is high, but it is also negatively related to the size of the eliminated group (a). For $a=1$, equation (5) reduces to

$$\Delta P = 4m^2(3F - 2 + 2m^3) \tag{5}$$

This is always satisfied if $F > \frac{2}{3}$, which is likely if $n \geq 3$ and the largest group's share is less than 50%. Thus, when regional shares are constructed by setting one group's share to zero and adding the remainder equally to all the other groups' national shares, average local polarization will exceed national polarization when F is high.

On the other hand, if there are only two groups (so $F \leq 0.50$), deviations of regional population shares from the national population shares imply that the population-weighted average of regional polarization is *below* the national figure, although regional polarization will be greater than the national figure in some regions.[8]

In the absence of regional data on ethnic composition, we capture this effect by using ethnic fractionalization as a proxy for the difference between average regional polarization (*LP*) and national polarization (*NP*). As Table 2 suggests, this deviation tends to be negative when F is low and positive when F is high.

We hypothesize that the incidence of conflict is increasing in *NP* and in *LP*, but that the proportion of the country affected is decreasing in *LP*, for given *NP*, because if *NP* is low, some regions may be very unpolarized, or have ethnic composition very different from those where there is conflict, and so remain unaffected. Thus we expect the extent of conflict to be negatively related to the degree of ethnic fractionalization.

4. Regional Concentration of Minority Groups

The MAR Project was initiated by Ted Gurr and is based at the University of Maryland's Center for International Development and Conflict Management (CIDCM). It "*tracks politically-active ethnic groups throughout the world from 1945 to the present. identifying where they are, what they do, and what happens to them. MAR focuses*

---

[8] For example, if there are two equally sized regions, then average regional polarization is smaller than national polarization by $2x^2$, where x is the deviation of a group's regional population share from its national average. This is the result of the non-linearity of the polarization index, which becomes more sensitive to variations in the size of the largest group as that size deviates from 0.5.

specifically on ethnopolitical groups, non-state communal groups that have "political significance" in the contemporary world because of their status and political actions. Political significance is determined by the following two criteria:

- *The group collectively suffers, or benefits from, systematic discriminatory treatment vis-a-vis other groups in a society*
- *The group is the basis for political mobilization and collective action in defense or promotion of its self-defined interests*" (MAR Codebook, pg 1)

This source provides information on geographical concentration of minority groups (by administrative region), and also for 1985 onwards the extent to which each group was involved in rebellious activity. The MAR data define minorities as either (1) widely geographically dispersed, (2) unevenly distributed with less than 50% in any one administrative region, (3) concentrated with between 50% and 75% in one region, or (4) highly concentrated with more than 75% in one region. Table 3 uses MAR data to show that in 83 cases out of 113, the second largest population group is at least 50% concentrated in one administrative region (and in 45 cases the population is more than 75% concentrated in one region).[9] These figures are not distorted by cases where the population of the second largest group is small – in fact concentration is more likely when the population share is larger, as is shown by the figures for mean population share. For example the mean shares of the second group are 0.2 and 0.19 respectively for the 83 countries in which either the majority of the population or the total share of the population of groups is concentrated in one region, but below 0.19 and 0.13 if the population is defined as widely dispersed or primarily urban but otherwise dispersed. A similar pattern is evident in the case of the third and fourth largest groups. Out of all the minorities in the data set, 74.7% have at least 50% of their population concentrated in one administrative region of the country.

Table 3: Minority Groups' Population Share by Regional Concentration in 2000

---

[9] Table 3 uses the variable coded "GROUPCON" in the MARGen data set.

Population Shares for Second Largest Group Given the Level of Concentration in 2000

Spatial Distribution of This Group	Freq.	Mean Share
Widely Dispersed	11	0.197
Primarily Urban otherwise Dispersed	19	0.126
Majority Concentrated in One Region	29	0.200
>75% Concentrated in One Region	54	0.192
	113	0.183

Population Shares for Third Largest Group Given the Level of Concentration in 2000

	Freq.	Mean Share
Widely Dispersed	9	0.037
Primarily Urban otherwise Dispersed	17	0.066
Majority Concentrated in One Region	27	0.082
>75% Concentrated in One Region	50	0.095
Total	103	0.082

Population Shares for Fourth Largest Group Given the Level of Concentration in 2000

	Freq.	Mean Share .
Widely Dispersed	6	0.030
Primarily Urban otherwise Dispersed	14	0.037
Majority Concentrated in One Region	18	0.042
>75% Concentrated in One Region	44	0.064
Total	82	0.052

In Tables 4 and 5, as an example, we show the level of regional concentration of groups for four highly fractionalized countries (Ghana, Nigeria, Angola, and Ethiopia), and for four highly polarized countries (Burundi, Rwanda, Latvia, and Estonia) respectively. Table 4 shows that groups in highly fractionalized countries are likely to be concentrated in one region. On the other hand, Table 5 shows that groups in highly polarized countries tend to be highly dispersed.

Table 4: Concentration of Minority Groups for Selected Fractionalized Countries (1985-2005)

Country	Fractionaliz.	Polariz.	GroupName	Group Sh.	Spatial Distribution	Share in Reg.
Ghana	0.846	0.50	Mossi-Dag.	0.160	Concentrated in Region	>75%
			Ewe	0.130	Concentrated in Region	>75%
			Ashanti	0.280	Concentrated in Region	>75%
Nigeria	0.8	0.61	Yoruba	0.200	Concentrated in Region	>75%
			Ibo	0.170	Majority in one Region	50-75%
			Ijaw	0.040	Concentrated in Region	>75%
			Hausa	0.290	Concentrated in Region	>75%
			Ogani	0.005	Concentrated in Region	>75%
Ethiopia	0.76	0.63	Amhara	0.264	Majority in one Region	50-75%
			Tigreans	0.097	Concentrated in Region	>75%
			Eritreans	0.75	Concentrated in Region	>75%
			Nilo-Saharan	0.016	Concentrated in Region	>75%
			Somalis	0.057	Concentrated in Region	>75%
			Afars	0.043	Concentrated in Region	>75%
			Oromo	0.400	Majority in one Region	50-75%
Angola	0.76	0.70	Mbundu	0.357	Concentrated in Region	>75%
			Bakongo	0.133	Concentrated in Region	>75%
			Cabinda	0.017	Concentrated in Region	>75%

Table 5: Concentration of Minority Groups for Selected Polarized Countries (1985-2005)

Country	Fractionaliz.	Polariz.	Group	Groups'	Spatial Distribution	Share in
Rwanda	0.18	0.36	Tutsis	0.163	Widely dispersed	No region
Burundi	0.33	0.64	Tutsis	0.153	Widely dispersed	No region
Latvia	0.58	0.86	Russians	0.341	Widely Dispersed	No region
Estonia	0.51	0.85	Russians	0.294	Primarily Urban (disp.)	<50%

Table 6: Regional Concentration of Groups and Ethnic Fractionalization in 2000 (Cross Minority Group Data)

	Dependent Variable: Concentration of Minority Groups		
Estimation Method: OLS	Model 1		
Ethnic Fractionalization (EF)	1.100***		
	(2.64)		
Ethnic Polarization		-0.124	
		(-0.35)	
Share of the Largest Group			-1.304***
			(-2.99)
Share of the Second Largest Group			-1.418*
			(-1.89)
Western Countries	-0.535	-0.684**	-0.571*
	(-1.62)	(-2.13)	(-1.67)
East Europe	-0.215	-0.313	-0.222
	(-0.85)	(-1.23)	(-0.84)
Latin America	0.222	0.257	0.329
	(0.83)	(0.95)	(1.21)
Sub-Saharan Africa	0.0751	0.371	0.0400
	(0.29)	(1.61)	(0.15)
Asia	-0.105	-0.102	-0.162
	(-0.41)	(-0.40)	(-0.61)
Constant	1.566***	2.173***	3.167***
	(5.44)	(7.54)	(8.14)
Observations	275	275	275
R-squared	0.119	0.089	0.122

The dependent variable is coded GROUPCON in the MARGen data set. It takes the value 0, 1, 2 or 3, with 3 representing the highest level of concentration. Robust t-statistics in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$

In Table 6 we provide evidence of the relationship between national ethnic diversity and regional concentration of minorities. In Model 1 the dependent variable is the degree of regional concentration of the group as specified above. The regressors are the national ethnic fractionalization index and dummies for different regions of the world. Only ethnic fractionalization is statistically significant (at the 1% level), and its coefficient is positive.

In Model 2 we replace ethnic fractionalization with ethnic polarization and the coefficient is insignificant. In Model 3 we finally replace indices of fractionalization and polarization with population shares of the largest and second largest groups. Regional concentration of minorities is higher when the shares of the two largest groups are smaller. These results imply that, when ethnic fractionalization is relatively high, local polarization may be considerable even if the national polarization index is relatively low.

5. Regional Concentration and Extent of Conflict

In this Section we turn to the relationship between regional concentration of minority groups and extent of conflicts. Table 7 provides data on the scale of the rebellion by each minority group in a given year, from the same source. There are only 1258 episodes in which a group is involved in a rebellion. Out of these 1258 episodes, there are 579 observations in which the rebellion is quite local (political banditry, campaign of terrorism, and local rebellion). In 170 episodes there has been civil war, and in 112 episodes there has been large-scale guerrilla activity. Finally there are also 395 episodes in which the scale of the guerrilla activity is either intermediate or small.[10] The data fit particularly well with our analysis given that these categories are also coded taking into consideration the geographical extent of the conflict. For example a local rebellion is defined as an armed attempt to seize power in a locale. Small-scale guerrilla activity is coded if there are 1) fewer than 1000 armed fighters; 2) sporadic armed attacks (less than six reported per year); or 3) attacks in a small part of the area occupied by the group, or in one or two other locales. On the other hand large-scale guerrilla activity is coded if there are 1) more than 1000 armed fighters; 2) frequent armed attacks (more than six per year); or 3) attacks affecting a large part of the area occupied by the group. A civil war affects the entire country.

---

[10] The relevant variable in the MARGen data set is coded "REB".

Table 7: Severity of Rebellion by Minority Groups

Rebellion	Freq.	Percent	Cum.
None reported	3,743	74.85	74.85
Political banditry, sporadic terrorism	307	6.14	80.98
Campaigns of terrorism (more than 6 events)	138	2.76	83.74
Local rebellions	134	2.68	86.42
Small-scale guerrilla activity	220	4.40	90.82
Intermediate guerrilla activity	175	3.50	94.32
Large-scale guerrilla activity	112	2.24	96.56
Civil war	163	3.26	99.82
Protacted Civil War	9	0.18	100.00

In order to test whether greater regional concentration of minority groups is associated with a higher probability of conflict and whether these conflicts are more likely to be of limited extent, we estimate a two-stage Heckman model where at the selection stage (Panel B) the dependent variable is a minority group's involvement in rebellion, and at the outcome stage (Panel A) the dependent variable is a measure of the scale of the rebellion (1 = political banditry up to local rebellions; 2 = guerrilla activity of any scale; 3 = civil war). As a measure of local distribution of the group we rescale the variable for the regional concentration of groups from the MARGen dataset on a 0-1 basis. As a result the variable is zero if the group is geographically dispersed and 1 if at least a minority of less than 50% of the population is concentrated in one region[11]. We also enter a full set of explanatory variables which includes national population (in logs), national GDP per capita (in log), dummies for anocracy, new independent states, and political instability which Sambanis (2004) finds to be significant predictors of conflict; controls for oil dependency from Fearon and Laitin (2003), alluvial diamond mines as in Lujala et al. (2005), and the relevant minority group's population share from the MAR project. Year dummies are also used to capture changes in the probability of conflict over time (i.e. after the cold war)[12]. The Heckman procedure controls for the sample selection bias (Heckman, 1979) related to the

---

[11] We also try different coding rules are results are largely confirmed.
[12] Table A1 in the Appendix 2 reports descriptive statistics.

extent of conflict, given that the latter is only observed for countries/groups which experience a conflict. By controlling for the correlation between unobservables in the selection equation (probability of conflict) with unobservables in the outcome equation (extent of conflict) the Heckman estimator controls for such a sample selection bias. Indeed, suppose that a conflict tends to have a smaller extent in countries which are more prone to conflict (i.e. countries which have experienced conflicts for longer periods like Sri-Lanka and Ethiopia) and that this is in some way correlated with the distribution of groups. Then if we do not control for this correlation between probability and extent of conflict there is a risk of over-estimating the extent of conflict. Practically, the distribution of groups will capture the correlation between distribution of groups and conflict rather than the correlation between distribution of groups and extent of conflict.

Results are shown in Table 8 for three alternative exclusion restrictions (variables that appear in the selection equation only): (a) a measure of the average level of precipitation (Model 1) which according to Miguel et al. (2004) affects conflict only through GDP growth and more specifically through commodity shocks (Ciccone, 2011), (b) a measure of the share of population in the tropical area (Model 2) from Sachs (2003) given that environment disease factors affect institutions (Acemoglu et al. 2002) and therefore the probability of conflict through state capability (Fearon and Laitin, 2003), (c) a dummy for pre-colonial origin (Model 3) which records whether the group has been living in the current region (area) since the eighteenth century (before colonization)[13]. Even though variables above represent plausible exclusion restrictions the potential non-excludability of these variables should bias results towards zero, given the negative correlation between the probability and the extent of conflict. Therefore in the worst case scenario there may be a downward bias (biased toward zero) of the variable of interest.

In Table 8 the inverse Mills ratio is significant for two out of three models and as argued above the negative coefficient captures the negative relationship between extent and probability of conflict. In Model 1 the variable for the regional concentration of groups is significant at the 1% level with positive coefficients in the selection equation, indicating

---

[13] This variable is coded as "TRADITN" in the MARGen data set.

that even with a moderate degree of regional concentration rebellion is more frequent. In the outcome equation for Model 1, the concentration dummy has a significant negative coefficient, indicating that with greater concentration rebellions are more localized. In Model 2 we replace precipitation with the share of the population in tropical area as an exclusion restriction and results are quite similar. The concentration dummy increases the probability of conflict but has a negative effect on the extent of rebellion though the effect is not significant due to the fact that the mills ratio is also not significant causing problems of sample selection bias. Finally in Model 3 we use the pre-colonial origins of the group as an exclusion restriction and results still hold with the dummy for regional concentration having a significant positive effect on the probability of conflict and a negative effect on the severity of rebellion.

The message of these results is, therefore, that more regionally concentrated minorities tend to engage in rebellions, and also that these rebellions are likely to be localized. This is consistent with our hypothesis about the importance of local polarization, which is likely to come about in ethnically fractionalized countries when groups are regionally concentrated. These results are also consistent with Toft's (2003) analysis of the impact of settlement patterns (i.e., regional concentration of groups) on the probability of ethnic violence when she argues that regional concentration of groups is likely to increase the probability of ethnic violence because of the higher capability related to the ability to mobilize fighters. She does not fully explore the impact of the distribution of groups on the extent of conflict. However her idea that groups concentrated in a historical homeland are more likely to be involved in conflict because of demands for independence seems consistent with our analysis related to the extent of conflict.

Table 8 – Severity of Rebellion by Minority Groups – 2-Step Heckman Estimator
(Minority Group Year Data over the period 1985-2005)

Estimation Method: Heckman Estimator	Panel A: Severity of the Rebellion			Panel B: Incidence of Rebellion		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Group Share of National Population	-3.232***	-1.193	0.266	0.777***	0.800***	-0.0767
	(-4.12)	(-1.49)	(0.93)	(4.24)	(4.35)	(-0.44)
Regional Concentration of Groups	-3.032***	-0.930	-0.821***	0.797***	0.781***	0.947***
	(-4.00)	(-1.20)	(-2.75)	(10.59)	(10.02)	(12.83)
GDP per Capita (lagged)	0.694**	-0.144	-0.272***	-0.324***	-0.299***	-0.306***
	(2.45)	(-0.49)	(-3.38)	(-13.85)	(-11.06)	(-13.34)
National Population (lagged)	-0.542***	-0.144	0.0187	0.155***	0.164***	0.0471***
	(-3.89)	(-0.94)	(0.54)	(9.40)	(9.20)	(3.01)
Oil Dependency	2.116***	1.065**	0.414***	-0.462***	-0.438***	-0.199***
	(5.15)	(2.37)	(3.82)	(-7.28)	(-6.99)	(-3.86)
Anocracy (Lagged)	0.542***	0.285*	0.276**	-0.102*	-0.0907	-0.180***
	(3.50)	(1.84)	(2.12)	(-1.85)	(-1.62)	(-3.52)
New Independent State	4.204***	2.436***	2.284***	-0.593***	-0.483**	-0.618***
	(5.26)	(3.25)	(4.68)	(-3.04)	(-2.44)	(-3.32)
Political Instability	-0.220	-0.167	0.0242	0.0533	0.110*	0.00251
	(-1.64)	(-1.03)	(0.19)	(0.86)	(1.74)	(0.04)
Diamond Mines	0.420**	-0.0487	-0.0430	-0.155***	-0.234***	-0.249***
	(2.46)	(-0.21)	(-0.38)	(-2.78)	(-3.80)	(-4.29)
Mountaious Terrain	-0.112**	0.0660	0.0702*	0.0508***	0.0375*	-0.0144
	(-1.99)	(1.36)	(1.80)	(2.67)	(1.90)	(-0.83)
Precipitation				-0.00004		
				(-1.38)		
Tropical Area					0.0537	
					(0.88)	
Pre-Colonial Origins						0.770***
						(11.23)
Inverse Mills Ratio	-6.111***	-2.396	-1.768***			
	(-4.45)	(-1.64)	(-4.53)			
Cut-Off 1	-9.135***	-5.721***	-4.007***			
	(-6.57)	(-3.85)	(-6.73)			
Cut-Off 2	-7.714***	-4.409***	-2.514***			
	(-5.57)	(-2.99)	(-4.31)			
Constant				-0.0381	-0.351	0.620**
				(-0.13)	(-1.00)	(2.24)
Observations	949	866	901	3,719	3,509	4,391

The dependent variable in Panel A is the severity of the rebellion recoded as 1 = political banditry up to local rebellions; 2 = guerrilla activity of any scale; 3 = civil war (see page 15). The dependent variable in Panel B is a dichotomous variable for whether such a group-year is involved in a conflict. Source: MARGen data set. Robust z-statistics in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Table 9– Marginal Effects

	Marginal Effects		
	Local Rebellion	Guerrilla	Civil War
Regional Concentration of Groups	0.275***	-0.065***	-0.210**
	(4.00)	(-2.80)	(-2.38)

In Table 9 we show marginal effects for the concentration variable in Model 3 above. For groups which are regionally concentrated the probability of a local rebellion increases by almost 27.5 percent. However the probability of guerrilla and civil war for groups that are regionally concentrated decreases by almost 6.5 percent and 21 percent respectively. Therefore the regional concentration of groups (which we use as a proxy of local polarization) on the one hand increases the probability of localized conflicts (local rebellion) and on the other hand decreases the probability of more widespread conflicts (i.e. guerrilla or civil war) which seems to be in line with our predictions

6. Conclusions

The relationship between ethnic diversity and conflict is still a matter of debate. In this paper we have first of all shown that this is because the relationship is far stronger at relatively low levels of diversity, where the two measures are so highly correlated as to be virtually indistinguishable. Then we have offered an explanation for the weakness of the relationship at higher levels of diversity.

Our central point is that minorities tend not to be evenly distributed across the country, so there can be a big difference between regional and national measures of diversity. Our theory is that what matters is polarization, either at the national or the local level. Because the regional distribution of ethnic groups within a country can be very uneven, an ethnically fractionalized country may be quite polarized in some localities, even

if its national polarization measure is not particularly high. Because of this, it may be difficult to pick up a relationship between conflict and national polarization measures at the more ethnically diverse end of the spectrum. The data strongly support this hypothesis: there is a strong correlation between the probability of conflict and either measure of ethnic diversity up to a measure of fractionalization of 0.5, but none at all above that level.

Using information on the geographical concentration of minorities and their involvement in rebellious activity from the data set of the Minorities at Risk Project we find (a) that in ethnically fractionalized countries minority groups are more likely to be regionally concentrated, implying pockets of local polarization; and (b) that minorities whose populations are geographically concentrated are more likely to be involved in rebellious activity, although this activity is also more likely to be small-scale and local. This evidence on minority groups' involvement in rebellion is consistent with the hypothesis that polarization at the local level makes local conflict more likely. The association between high ethnic fractionalization and significant polarization at the regional level explains why conflict is more likely in these cases than would be predicted on the basis of a relatively low level of national polarization.

Appendix 1

To investigate the relationship between *average* local polarization and ethnic fractionalization in the presence of regional concentration of groups more formally, consider cases where at least one ethnic group is not represented in any given region. We analyse two possibilities: (1) where the regional population shares are formed from the national population shares by adding the missing group's share entirely to that of one other group (which is equivalent to the merging of two groups); and (2) where the share of the missing group is redistributed equally amongst all the represented groups.

(1) Merging of two groups

If groups *i* and *j* merge, then from equation (2) the change in the polarization index is:

$$\Delta P = 4[(\pi_i + \pi_j)^2(1 - \pi_i - \pi_j) - \pi_i^2(1 - \pi_i) - \pi_j^2(1 - \pi_j)] = 4\pi_i\pi_i[2\text{-}3(\pi_i + \pi_i)] \qquad \text{(A1)}$$

This expression is positive if the two groups' shares sum to less than two-thirds, but negative if they sum to more than two-thirds. Thus regional concentration of groups tends to increase local polarization up to the point where the merged group reaches a two-thirds majority locally. In particular, once the number of groups exceeds three, and if the two largest groups make up less than two-thirds of the population, then *all* regions will have greater polarization than the national measure. Note that this case of more than three groups with the two largest groups' shares representing less than two-thirds implies a high degree of ethnic fractionalization.

(2) Elimination of one group with its share equally divided between the remaining groups

In the second case (the elimination of group j and redistribution of its share equally amongst the n remaining groups), the algebra is more involved. Let the eliminated group be group 1, and let there be n other groups.

Writing $m = 1/n$, then

$$\Delta P = 4 \sum_{i=2}^{n+1} [(\pi_i + m\pi_1)^2 (1 - \pi_i - m\pi_1) - \pi_i^2 (1 - \pi_i)] - 4\pi_1^2 (1 - \pi_1) \tag{A2}$$

Expanding this gives

$$\Delta P = 4 \sum_{i=2}^{n+1} (2m\pi_i\pi_1 - 3m\pi_i^2\pi_1 - 3m^2\pi_i\pi_1^2 + m^2\pi_1^2 - m^3\pi_1^3) - 4\pi_1^2 (1 - \pi_1) \tag{A3}$$

Substituting

$$\sum_{i=2}^{n+1} 1 = 1/m, \quad \sum_{i=2}^{n+1} \pi_i = 1 - \pi_1, \quad \sum_{i=2}^{n+1} \pi_i^2 = 1 - F - \pi_1^2,$$

we obtain

$$\Delta P = 4m\pi_1(3F - 1) - 4(1 + m + 3m^2)\pi_1^2 + 4(1 + 3m + 2m^2)\pi_1^3 \tag{A4}$$

Substituting $\pi_1 = am$ yields the following expression:

$$\Delta P = 4am^2[3F - 1 - a(1 + m + 3m^2) + a^2 m(1 + 3m + 2m^2)] \tag{A5}$$

26

where $m = n^{-1}$, $a = n\pi_{\hat{j}}$ and F is the index of ethnic fractionalization as defined in equation (1) before the elimination of group 1. The parameter a reflects the relative size of the excluded group j (which is equal to the average of the others when $a = n/(n+1)$). Equation (A5) is more likely to be positive when F is high, but it is also negatively related to the size of the eliminated group (a). For $a=1$, equation (A5) reduces to

$$\Delta P = 4m^2(3F - 2 + 2m^3) .$$
(A6)

This is always satisfied if $F > \frac{2}{3}$, which is likely if $n \geq 3$ and the largest group's share is less than 50%. Thus, when regional shares are constructed by setting one group's share to zero and adding the remainder equally to all the other groups' national shares, average local polarization will exceed national polarization when F is high.

Appendix 2: Descriptive Statistics, Variables and Sources of Data

Table A1: Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Extent of Conflict	5,366	.4045844	.7947459	0	3
Group Population Share	6,752	0.116299	0.147400	0.0004	.87
Regional Concentration of Groups	6,845	.8701242	.3361913	0	1
GDP per Capita (lagged) – in log	5,962	8.157372	1.042435	5.139058	10.44478
National Population (lagged) – in log	5,954	9.849109	1.511517	5.667917	14.06172
Oil Dependency	6,273	.241033	.4277444	0	1
Anocracy (Lagged)	6,273	.1895425	.3919702	0	1
New Independent State	5,976	.0140562	.1177326	0	1
Political Instability	4,879	.1768805	.3816067	0	1
Diamond Mines	5,976	.2121821	.4088873	0	1

Group Concentration, Extent and Incidence of Conflict: Minority at Risk Dataset

Ethnic Fractionalization, Largest Ethnic Group, Second Largest Group, Mountainous Terrain, the Ratio of Oil Export to Manufacturer Exports: Fearon and Laitin (2003). Fearon and Laitin (2003) provide data on oil exports for the period 1945-1999. We use data from the Energy International Administration to calculate the ratio of oil exports to merchandise exports for the period 2000-2003.

GDP per Capita and Population: Penn World Table 6.2.

Alluvial diamonds: Lujala *et al.* (2005).

Ethnic Polarization: Montalvo and Reynal-Querol (2005a). The polarization data are not available for all countries, so we impute the missing numbers using data on the share of the second largest group provided by FL. The correlation between polarization and the share of the second largest group is 0.77.

Regional Concentration and Rebellious Activity of Minority Groups: Minority at Risk Project (MAR).

Democracy: Marshall M. G. and K. Jaggers (2002). We use a dummy which is equal one if the Polity2 index is above zero in order to capture whether a country is democratic or not, and the Polity2 index in order to control for the extent of democracy.

References

Alesina, A., R. Baqir, and W. Easterly (1999), 'Public Goods and Ethnic Divisions', *Quarterly Journal of Economics, 114 (4): pp. 1243–1284*

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg (2004), **'**Fractionalization', *Journal of Economic Growth,* 8: 155-194.

Barro R. J. (1991), 'Economic Growth in a Cross-Section of Countries', *Quarterly Journal of Economics*, 106 (2): 407-433.

Becchetti, L., P. Conzo, and A. Romeo (2014), **'**Violence, Trust, and Trustworthiness: Evidence from a Nairobi Slum', *Oxford Economic Papers, 66 (1): 283-305*

Blattman, C., and E. Miguel (2010) , 'Civil War', *Journal of Economic Literature*, 48(1): 3-57

Bleaney, M.F., and A. Dimico (2011), 'How Different are the Correlates of Onset and Continuation of Civil Wars?', *Journal of Peace Research* 48(2): 145-155

Ciccone, A. (2011), 'Economic Shocks and Civil Conflict: A Comment', *American Economic Journal: Applied Economics, 3(4): 215-27.*

Collier P., and A. Hoeffler (1998), 'On Economic Causes of Civil War', *Oxford Economic Papers*, 50(4): 563-573

Collier P., and A. Hoeffler (2004), 'Greed and Grievance in Civil War', *Oxford Economic Papers*, 56 (4): 563-595.

Collier P., A. Hoeffler, and M. Söderbom (2004), 'On the Duration of Civil War', *Journal of Peace Research*, 41(3): 253-273.

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

Costalli, S., L. Moretti, and C. Pischedda (2014), 'The Economic Costs of Civil War: Synthetic Counterfactual Evidence and the Effects of Ethnic Fractionalization'. *HiCN Working Paper 184*

Easterly W., and R. Levine (1997), 'Africa's Growth Tragedy: Policies and Ethnic Divisions', *Quarterly Journal of Economics*, 112(4): 1203-50

Esteban J., and D. Ray (1994), 'On the Measurement of Polarization', *Econometrica*, 62(4): 819-51.

Esteban J., and D. Ray (2011) "Linking Conflict to Inequality and Polarization, *American Economic Review*, 101: 1345-74

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

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

Fearon J. D. (2004), 'Why Do Some Civil Wars Last So Much Longer than Others?', *Journal of Peace Research*, 41 (3): 275-301.

Fearon J. D., and D. D. Laitin (2003)**, '**Ethnicity, Insurgency, and Civil War', *The American Political Science Review*, 97 (1): 75-90.

Fearon J. D., and D. D. Laitin (2011)**, '**Sons of the Soil, Migrants, and Civil War', *World Development*, 39 (2): 199-211.

Gardeazabal, J. (2011), 'Linguistic polarization and conflict in the Basque Country', *Public Choice 149:405–425*

Goldstone J. A., R. Bates, T. Robert Gurr, M. Lustik, G. Monty, G. Marshall, J. Ulfelder, and M. Woodward (2010), 'A Global Model for Forecasting Political Instability', *American Journal of Political Science*, 54(1): 190–208

Heckman J. (1979), 'Sample Selection Bias as a Specification Error', *Econometrica*, 47: 153-161

Hegre, H. and N. Sambanis (2006), 'Sensitivity Analysis of the Empirical Literature on Civil War Onset', *Journal of Conflict Resolution,* 50 (4): 508-535

Hudson M. C., C. L. and Taylor (1972), *'World Handbook of Political and Social Indicators',* Second Edition, Yale University Press

Kanbur R., P. K. Rajaram, and A. Varshney (2011)', Ethnic Diversity and Ethnic Strife. An Interdisciplinary Perspective', *World Development*, *39(2)*: *147-158*

Lujala P., N. P. Gleditsch, E. and Gilmore (2005), 'A Diamond Curse? Civil War and a Lootable Resource' *Journal of Conflict Resolution*, 49 (4): 538-56

Marshall M. G., and K. Jaggers (2002), 'Polity IV Project: Political Regime Characteristics and Transitions 1800-2002'

Miguel E., and M. G. Gugerty (2005), 'Ethnic Diversity, Social Sanctions, and Public Goods in Kenya', *Journal of Public Economics, 89 (11-12): 2325–2368*

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

Minority at Risk Project. (2009) "Minority at Risk Dataset", College Park, MD: Center for International Development and Conflict Management; Retrieved from: www.cidcm.umd.edu/mar/data.aspx

Montalvo J., and M. Reynal-Querol (2005a), 'Ethnic Polarization, Potential Conflict, and Civil Wars'. *American Economic Review*, 95(3): 769-813.

Montalvo J, and M. Reynal-Querol (2005b), 'Fractionalization, Polarization and Economic Development', *Journal of Development Economics*; 76(2): 293-323.

Montalvo J, and M. Reynal-Querol (2010), 'Ethnic Polarization and the Duration of Civil Wars', *Economics of Governance* 11(2): 123-143

Peace Research Institute Oslo, (2007)**, '**Armed Conflict Dataset: Version 4-2007"' www.prio.no/CSCW/Datasets/Armed-Conflict/UCDP-Prio/

Rohner, D., M. Thoenig and F. Zilibotti (2013), 'Seeds of Distrust: Conflict in Uganda', *Journal of Economic Growth 18: 217-52*

Sambanis N. (2004)**, '**What is Civil War? Conceptual and Empirical Complexities of an Operational Definition', *Journal of Conflict Resolution*; vol. 48(6): 814-59

Schneider, G., and N. Wiesehomeier (2010),' 'Diversity, Conflict and Growth: Theory and Evidence', *Diversity* 2, 1097-1117.

Singer D. J., and M. Small (1982), '*Resort to Arms: International and Civil Wars, 1816-1980',* Beverly Hills, CA: Sage

Singer D. J., and M. Small (1994)**,** *Correlates of War Project: International and Civil War Data, 1816-1992: A Statistical Handbook*', New York: John Wiley

Toft, M. (20030, '*The Geography of Ethnic Violence: Identity, Interests, and the Indivisibility of Territory',* Princeton: Princeton UP.

Tollefsen, A. F., H. Strand, H. and Buhaug (2012) 'PRIO-GRID: A Unified Spatial Data Structure'. *Journal of Peace Research*, vol. 49 (2): 363-75.

Weidmann, N. B., J. K. Rød, and L. E. Cederman (2010), ' Representing Ethnic Groups in Space: A New Dataset,' *Journal of Peace Research*, vol. 47 (4): 491-9

Wimmer, A., and L. E. Cederman (2009), '"Ethnic Politics and Armed Conflict: A Configurational Analysis of a New Global Data Set', *American Sociological Review*, *74*(2): 316–337