Unpacking Ethnicity

Exploring the Underlying Mechanisms Linking Ethnic Fractionalization and Civil Conflict

Hilde A. Bakkan, Tor G. Jakobsen & Jo Jakobsen

Abstract

In the study of civil conflict, many regard ethnic fractionalization as a measure of grievance. Yet, there is also a rational-actor explanation as to why ethnicity might cause rebellion. A third possibility is that ethnic fractionalization in itself causes civil war. Using a novel approach, we provide an empirically-based answer to the question of which of these three mechanisms is most prominently at play. Firstly, we present our updated index of ethnolinguistic fractionalization. We have re-calculated the index so that it corresponds with today’s ethnic compositions, using interpolation to fill in values between 1990 (the last recordings of the old measure) and our own data from 2011. Secondly, we establish empirically that there is indeed a positive (though slightly curvilinear) link between ethnicity and civil conflict. Thirdly, we present two factor analyses, both of which demonstrate that ethnicity in itself is a mechanism that explains the outbreak of civil conflict.

Keywords

Ethnicity, civil war, factor analysis, opportunity, grievance, identity
Introduction

Collier and Hoeffler’s (2004) renowned article “Greed and Grievance in Civil War” tested the explanatory power of indicators related to opportunity and grievance, presenting *ethnic fractionalization* as a measure of the latter. This followed Fearon and Laitin (2003), who also placed this measure in the grievance category. Yet, there is also a rational-actor approach as to why ethnicity might be a source of rebellion. Gagnon (1994–1995), for example, argues that conflicts taking place along ethnic lines are not caused primarily by ancient hatred, but rather by determined political actors who consciously use ethnicity as an instrument to create violent conflict. A third, less explored possibility is that ethnic fractionalization *in itself* causes civil war. The literature on civil conflict, taken as a whole, has still not investigated this issue thoroughly. Using a novel approach, we therefore seek to provide an empirically-based answer to the question of which of these three mechanisms is most prominently at play. The paper asks the following research question: *What mechanism links ethnic fractionalization to the onset of civil conflict?*

The text proceeds as follows: Firstly, we present our updated index of ethnolinguistic fractionalization (ELF). This index reflects the probability that two randomly drawn individuals, belonging to the same country, are members of different ethnolinguistic groups. We have re-calculated the index so that it corresponds with today’s ethnic compositions, using interpolation to fill in values between 1990 (the last recordings of the old measure) and our own data from 2011. Secondly, we establish empirically that there is indeed a positive (though slightly curvilinear) link between ethnicity and the onset of civil conflict – though only at a relatively low level of conflict intensity. Thirdly, we present two factor analyses, both of which demonstrate that ethnicity in itself is a mechanism that explains the outbreak of civil conflict, working through the mechanism of identity.
1. Theory

There exist sundry definitions of ethnicity, which demonstrates that it is a difficult notion of which to obtain a clear grasp (Horowitz 2000). Ethnicity as a term hails back to the Greek word *ethnos*, and it was originally used to describe a kinship group whose members are linked together by bonds of blood (Wolff 2006). Ethnicity is based on the belief in a common heritage – viz., a belief in the existence of group characteristics that are held to be innate. The salience of ethnic identity becomes particularly recognized when ethnic identity has conspicuous consequences collectively for a group in its relation to the state and to other ethnic groups. In other words, ethnicity is a phenomenon that is likely to be highly significant when it forms a basis for people’s security, material well-being, status, or access to political power (Gurr 2000).

There are basically three different perspectives seeking to explain the underlying mechanisms that drive ethnic groups into mobilizing for violence: primordialism, instrumentalism, and constructivism. Firstly, the primordialists argue that the main characteristics of a group are permanently fixed. Each individual is born with a specific ethnic identity that cannot change; ethnic identities remain the same over time (Fenton 2010). Following this line of thinking, a person’s ethnic identity represents a phenomenon that is deeply integrated in human experience and history. Hence we cannot reject that it exists, either subjectively or objectively; it must rather be seen as a vital part of life or, specifically, of the relationship between groups and individuals (Wolff 2006, 33). In general, therefore, civil war is a result of ethnic differences, with other factors – like economic or political differences – playing a far less significant role (Blimes 2006, 537; Østby 2008).

The instrumentalists, for their part, criticize primordialism for posing a simplistic explanation that centers on “ancient hatred” as a cause of civil war, claiming that they overlook deeper problems that are the true reasons for conflict (Blimes 2006). They hold
instead that ethnicity is merely a tool that an individual or a group uses in order to achieve a particular goal (Blimes 2006, 537; Gagnon 1994–1995). Ethnicity is not in itself the source of violent conflict; it only becomes associated with war when leaders endeavor to demonstrate that ethnicity is tied to a history of conflict, which promotes acute social uncertainty and fear of what the future might bring. Only then is it that ethnicity emerges as a major fault line along social fractures (Lake and Rothchild 1998, 7).

Constructivism represents the third perspective on the connection between ethnicity and violence. This school of thought aims to bridge the primordialist and instrumentalist views. A constructivist views ethnicity as a result of common tangible features. Ethnicity is neither fixed, nor is it completely malleable. It is in part inherited and in part chosen (Østby 2008). The significance and content of ethnic-group identity can and do fluctuate over time, usually as a response to changes in the group’s political and social environment (Gurr 2000, 4; Wolff 2006, 34). Constructivism is at present the dominating school of thought within conflict studies.

1.1 The empirical connection between ethnicity and conflict

According to Collier and Hoeffler (2004) and Fearon and Laitin (2003), there is only a scant connection between the level of fractionalization within a country and the risk of civil war. Contrary to this, Sambanis (2001) found that ethnic fractionalization is significantly and positively correlated with the likelihood of civil war onset, a result supported by Regan and Norton (2005). Hegre and Sambanis (2006) found a link between ethnicity and low intensity conflicts, though there was no such connection when full-scale civil wars were used as the dependent variable. Collier, Hoeffler, and Rohner’s (2009) combined measure of ethnic fractionalization and religious fractionalization – which they named “social fractionalization” – was also found to be associated with civil war onset. The common denominator of all these
studies is that they employed a version of the ethnolinguistic fractionalization index (ELF) (Department of Geodesy and Cartography 1964). The ELF Index reflects the probability that two randomly drawn individuals, belonging to the same country, are members of different ethnolinguistic groups. The groups are defined by their roles, their origins, and their relationship to other groups (Taylor and Hudson 1972, 215). The ELF Index is based on the formula for the Herfindahl Index as well as data from *Atlas Naradov Mira* (Department of Geodesy and Cartography 1964) and other sources (Fearon and Laitin 2003). It ranges from 0 (ethnolinguistic homogeneity) to 1 (maximum heterogeneity or fractionalization).

Though overall findings are still inconclusive, the lion’s share of researchers find ethnic fractionalization to be a significant indicator of civil war onset. Of the above-mentioned scholars, only Collier and Hoeffler (2004), and in part Fearon and Laitin (2003), employ ethnic fractionalization as an indicator of grievance.

Others have found violence to be less pronounced in homogeneous and heterogeneous (or fractionalized) societies, juxtaposed with so-called polarized societies (Montalvo and Reynal-Querol 2005; see also Ellingsen 2000). *Figure 1* depicts three ideal-type versions of a state’s ethnic composition, which are possibly related to armed rebellion in a non-linear fashion. The best way to control for this when employing ethnic fractionalization as a measure of ethnicity, is to model a curvilinear relationship.

---

1 Fearon and Laitin filled in missing values based on sources such as the CIA World Factbook, Encyclopedia Britannica, the Library of Congress Country Studies, and other country-specific sources (Fearon and Laitin 2003, 78).
2. Three explanations of the link between ethnicity and civil conflict

The present literature tends to view ethnicity as a measure of grievance, although some argue that it should instead be considered an indicator of opportunity. Alternatively, following the primordialist concept of ethnicity, one could claim that ethnicity is a phenomenon which directly and in its own right causes civil conflict. In such a conception, ethnicity might usefully be named identity.

2.1 The grievance explanation

As stated earlier, in their oft-cited article Collier and Hoeffler (2004) included ethnic fractionalization as a measure of grievances. Toft (2002, 84) argues that ethnicity is a fundamental factor in people’s identity, just as it forms a vital common basis for group identity. On the other hand, it does not necessarily suffice to have a common identity; a shared grievance must also be present (Murshed and Gates 2005, 122). Based on the theory of
relative deprivation, grievance is triggered by any mismatch between prevailing expectations and actual conditions. Relative deprivation can be linked to ethnicity through a feeling of disaffection held in common by group members. Our first hypothesis is therefore as follows:

H1: *Ethnicity is an indicator of grievance.*

### 2.2 The opportunity explanation

The opportunity approach to civil war views ethnic groups as an extraordinary form of interest groups where group members have common preferences with regard to (virtually) all types of public policies. Ethnic groups are thereby conceptualized as political coalitions whose *raison d’être* is to fend for their members’ (or leaders’) economic interests (Fearon 2006, 861ff). Ethnic groups must be considered rational actors who make decisions on whether to enter or initiate armed conflict based on cost-benefit assessments and utility-maximization reasoning (Jacoby 2008, 126; Oyefusi 2008; Toft 2002, 85). Wealth and resources are viewed as critical for group survival. In short, ethnic groups are basically seen as rational actors who rebel when they have the opportunity to acquire resources and wealth as a result of the rebellion. From this argument stems our second hypothesis:

H2: *Ethnicity is an indicator of opportunity.*

### 2.3 A measure of its own: identity

A third explanation sees shared ethnic identity as a sufficient mobilization factor in itself (Harff and Gurr 1993). The common identity “separates us from them,” and it includes relations within, across, and outside the boundaries that distinguish between different groups, providing these boundaries with meaning (Tilly and Tarrow 2007, 79). Ethnicity – being a
function of the fundamental characteristics of language, religion, and heritage – has enormous power when it becomes the predominant basis of identity (Wolff 2006, 31). Vitally, ethnicity can be seen as an identity-associated feature – and a mechanism in its own right that can cause civil conflict directly and independent of any relationship to or existence of either grievance or opportunity. Based on this argument, our third and last hypothesis is as follows:

H3: *Ethnicity in itself is an indicator for civil conflict.*

Our three hypotheses are illustrated in *Figure 2.*

**Figure 2. Overview of the theoretical framework**

```
Ethnicity →? Greivance →? Opportunity → Civil Conflict

Identity →?
```

### 3. Research design

This paper makes two important contributions to the literature on civil conflict. First and foremost, by means of both a logistic regression model and exploratory factor analysis, we test – and unpack – the relationship between ethnicity and civil conflict. Our second contribution consists of the construction of an updated version of the ethnolinguistic fractionalization (ELF) Index. The newest data included in Fearon’s (2003) version of the ELF Index were from early 1990. The world – as well as many of its constituent states – has since certainly experienced major demographic changes. Following the works of Alesina et al.
(2003) and Fearon (2003), our updated version of the ELF Index is coded using secondary data. Calculations are based on the Herfindahl Index procedure:

\[ ELF = 1 - \sum_{i=1}^{n} s_i^2, \text{ where } s_i \text{ is the share of group } i \ (i=1,\ldots,n) \]

We calculated an updated ELF value for all countries in the world, in the main using information from *Ethnic Groups Worldwide* (Levinson 1998), *CIA World Factbook*, the *Joshua Project*, and *Library of Congress*. Basing the coding on several sources, each of which can rightfully be regarded as trustworthy and authoritative, increases our confidence in the reliability of our measure. Instances of significant inter-source variance were, in any case, rare. In general, the most recent information was applied.

The ELF Index has, understandably, been criticized for the fact that it is a constant measure whose value for a given country does not change over time (Kalyvas 2008; Laitin and Posner 2001). The new version rectifies this. Our ELF Index is based on the method of interpolation between the values presented by Fearon (2003) and the updated ones from 2011.

---

2 See https://www.cia.gov/library/publications/the-world-factbook/

3 See http://www.joshuaproject.net

4 See http://www.loc.gov

5 Yet some variance in coding practice between these different sources does apply. For example, the CIA World Factbook often refers to a given percentage that depicts the sum of several ethnic groups. The Democratic Republic of Congo (DRC) is one case in point. The CIA World Factbook presents one joint number (45%) for the three largest ethnic groups in that country. Considering that the DRC consists of around 200 ethnic groups, using the CIA World Factbook would, in this instance, understate the degree of ethnic fractionalization. In such cases, the Joshua Project, which retains more detailed information, has been used instead.
which we coded. The resulting variable – ETHFRACPOLATED – has been kept constant for the period up to and including 1990 (6 time units in the dataset, 1985–1990), using Fearon’s data. The values between 1990 and 2011 are interpolated. Thereby we account for temporal changes in the composition of ethnic groups.

In addition to running a factor analysis, using the variable ETHFRACPOLATED as a measure of ethnicity, it is of interest to test a variable that captures any curvilinear relationship. We therefore created such a variable – ETHFRACWAR – that is pinpointed to capture the part of ethnicity that is associated with conflict. Thereafter, we performed a logistic regression analysis with CIVIL CONFLICT ONSET as the dependent variable, and ETHFRACPOLATED, as well as its squared term, as independents. From this model, the vertex of the relationship between civil conflict onset and ethnicity was calculated. As shown in Figure 3, the vertex is reached when the level of fractionalization is equal to 0.78.

**Figure 3. Likelihood of civil conflict onset, by level of ethnic fractionalization**

---

6 Interpolation is a method that draws a straight line between two known points. In our case this means the old ELF values and the new ones, which are used as the basis upon which to calculate, in a linear fashion, the values between these two time points.
We thereafter recoded the variable based on the vertex, so as to ensure that the highest values correspond to the highest probability of civil conflict. Thus, the ELF value 0.78 equals 1.00 after recoding; the values 0.77 and 0.79 both equal 0.99; 0.76 and 0.80 now equal 0.98; and so on. Both ETHFRACIPOLATED and ETHFRACWAR are included in the two separate factor analyses exhibited in the results section.

4. Methods and variables

In order to shed light on our three hypotheses, we performed both a logistic regression analysis (i.e., the binary time-series cross-section method, see Beck, Katz, and Tucker 1998) and two factor analyses. Our models include 175 states over the time period 1985–2010, which yields a total of 4,471 units (country-years). The dependent variables are CIVIL CONFLICT and CIVIL WAR ONSET.\textsuperscript{7} UCDP/PRIO defines civil conflict onset as (the first year of) a conflict with more than 25 yearly battle-related deaths (Gleditsch et al. 2002). The definition of civil war uses the threshold of 1000 battle-related deaths in a given year. Furthermore, a new conflict/war onset is only registered if there are 2 or more years since the last observed conflict/war. For temporal control we have constructed the variable PEACE YEARS, which measures the number of years since the last recording of conflict in a given country. Together with this variable, we included a natural cubic spline with three knots as there is no theoretical reason for assuming that there is a linear impact of time on the probability of conflict. The estimated spline coefficients can be used to trace out the path of duration dependence (Beck, Katz, and Tucker 1998, 1270). In order to clarify the empirical association of our measure of ethnicity, we also included a handful of variables that can be

\textsuperscript{7} Both variables are from UCDP/PRIO Armed Conflict Dataset, version 4-2012.
viewed as indicators of grievance, as well as variables that proxy opportunity. These are described in the two subsequent sub-sections.

4.1 Theoretical grievance variables

POPULATION: The empirical literature has identified the size of a country’s population as one of the strongest predictors of civil war onset (Blattman and Miguel 2010; Dixon 2009; Jakobsen, de Soysa, and Jakobsen 2013). And even if this relationship is still rather under-theorized, it is possible that population size and grievances are positively related – as Collier and Hoeffler (2004) claim that population pressure may result in political violence. Because of its skewness, we follow others and include a log-transformed measure of population in both the regression analysis and the factor analysis (Dixon 2009). In the regression models, POPULATION is lagged one year, due to its dependence on time. Data are from the World Bank (2013).

POPULATION GROWTH: The yearly rate of increase in a state’s population is also included as a grievance variable. The same theoretical justification as for POPULATION applies here: If the level of grievance is positively associated with population size, it must also rise when the population grows, possibly due to population-related pressure (Abouharb and Kimball 2005, 750).

POLITY: Democracy has also proven to be a particularly robust predictor of civil conflict, and it is widely seen as a measure of grievance (Collier and Hoeffler 2004; Smith 2009). As many others, we use the Polity scale, which runs from -10 (fully institutionalized autocracy) to +10 (fully institutional democracy). Data are from the Polity IV Project.\(^8\) In the factor

analysis, the original variable is included. In the regression analysis, though, we use a recoded version with three categories. Values below -5 are referred to as AUTOCRACY, and values above 6 are coded as DEMOCRACY. The reference category – ANOCRACY – represents the values between -4 and 6. This dummy coding applies because others have demonstrated that there is an inverted U-shaped relationship between regime type and the onset of civil war (Hegre et al. 2001).

MILITARY SPENDING: Countries with a substantial military budget (relative to GDP) do not necessarily properly attend to the welfare of their citizens, which might increase sentiments of grievance among people. This could particularly be the case in the aftermath of a civil war, as expectations among the public of improvements in state services are likely to be disappointed if the government continues to prioritize the military, which, in turn, would increase the risk of a renewal of the civil war (Murshed 2002). Moreover, high military spending can also be interpreted as an indication of ill intentions on the part of the government (Collier et al 2003, 87). MILITARY SPENDING measures military expenditures as a percentage share of a country’s GDP, with data from the World Bank (2013).

EMPOWERMENT RIGHTS: This variable is an additive index that controls for the government’s respect for (different categories of) human rights. The variable is a 15-point scale, running from 0 (government does not respect human rights) to 14 (the government fully respect human rights) (Cingranelli and Richards 2010). According to Gurr (1983), repressing citizens is a strategy often used by the ruling elites to ensure their power, which in turn may

---

9 The variable has been dummy coded in the regression analysis in order to demonstrate the inverted U-shape between regime type and civil war onset that has been documented by Hegre et al (2001).

10 The measure includes all current and capital expenditures on the armed forces. Civil defense and current expenditures for previous military activities are excluded (World Bank 2013).
certainly cause grievance among the populace. EMPOWERMENT RIGHTS is included in the both the regression and the factor analysis; in the former it is lagged one year due to its time dependence.

PHYSICAL INTEGRITY RIGHTS: This is also an additive index, measuring the protection from extrajudicial killings, torture, and disappearance, as well as imprisonment for political beliefs. A low score on this 9-point scale indicates a lack of such protection, and the presence of inequality between groups and individuals within the country in question, which should be related to grievance (Cingranelli and Richards 2010). The same argument as for the preceding variable applies here: The ruling elite uses these methods not least in order to secure their grip on power (Gurr 1983).

4.2 Theoretical Opportunity Variables

GDP PER CAPITA: The size of a country’s economy is one of the most commonly-used variables in this field of study, with the literature robustly demonstrating a negative relationship between GDP per capita and civil war onset (Dixon 2009; Hegre and Sambanis 2006; Ward, Greenhill, and Bakke 2010). Others have judged this measure to belong to the opportunity dimension (Collier and Hoeffler 2004); Fearon and Laitin (2003, 75), for example, use it as a proxy for police and counterinsurgency strength/weakness. Due to the positive skewness of the variable, we log transformed it. This variable is also included in the logistic regression analysis, where we lag it one year. Data are from the World Bank (2013).

FUEL EXPORT: The positive relationship between civil war and natural-resource wealth – usually proxied by fuel exports – is one of the most robust findings in the empirical literature (Dixon 2009; Le Billon 2001; Murshed and Tadjoeddin 2009; Ross 2003; Snyder and
Bhavnani 2005). Theory informs us that the variable proxies opportunity; groups that get access to natural resources thereby likely acquire the financial means to initiate or sustain the rebellion (Ross 2004; Ross 2006, 265). We employ a dummy variable based on data from World Bank (2013). In countries where the export of fuel exceeds 1/3 of total merchandise exports, the variable has been given the value 1, otherwise 0. In the regression analysis, this variable is lagged one year, due to time dependence.

BUREAUCRATIC QUALITY: A competent bureaucracy should more easily be able to police the national territory and provide services that decrease the risk that citizens will choose to rebel (Fearon and Laitin 2003; Oyefusi 2008). Furthermore, a strong bureaucracy might also act as a check on rebellion (DeRouen and Sobek 2004). Contrarily, given that providing sound institutions is costly, a frail bureaucracy is associated with financially weak states, which in turn is related to a (relatively) high opportunity for rebellion (Fearon and Laitin 2003). BUREAUCRATIC QUALITY is measured on a 4-point scale, with data from the International Country Risk Guide (ICRG 2011). High scores indicate that the bureaucracy of the country in question has the expertise and strength required to continue functioning without any interruption to the provision of government services, and without any extreme changes in policy. Low scores, for their part, signify that the bureaucracy is prone to malfunctioning, in particular if the government changes (ICRG 2011).

LAW AND ORDER: This variable is correlated with the quality of the bureaucracy. In order for the legal system to be functioning well, it needs both legitimacy and necessary funding (DeRouen and Sobek 2004); LAW AND ORDER thus also yields information about a country’s wealth. The two aspects of law and order are coded separately, and together they make up a 6-point scale, with data from the ICRG (2011). The “law” component signifies the
impartiality and strength of the legal system, while “order” reflects the degree to which the law is commonly observed. A high score on LAW AND ORDER indicates that the state has the necessary capacity to control what happens within its borders, which should reduce the opportunity for rebellion.

AGRICULTURE: Developing-country economies are usually dominated by the primary sector. Poor conditions for agriculture, we reckon, have severe implications for socio-economic development and conflict. Indeed, most armed conflicts are located in regions that are heavily dependent on agriculture. Agriculture can be deemed a proxy for capacity because countries that are dependent on farming often have a “backward” economy, with the link between agriculture and armed conflict being poverty (de Soysa and Gleditsch 1999). This variable is measured as the size of the agricultural sector (including hunting, forestry, fishing, livestock production, and cultivation of crops) relative to GDP (World Bank 2013).

LIFE EXPECTANCY: A country’s life expectancy is affected by several factors. These include infant and birth mortality rates, which should be positively related to civil war onset not least given their association to level of development, which in turn proxies economic capacity (wealth) and opportunity for rebellion (Dixon 2009).\(^\text{11}\) Our variable measures life expectancy at birth, with data from the World Bank (2013).

RURAL POPULATION: In the process of recruitment, rebels often find it helpful to deny government troops access to rural areas. Insurgency groups typically establish themselves along the rim of the country’s borders, especially if the country is poor and/or has dysfunctional institutions. Importantly, people in rural areas might have a relatively low level

\(^{11}\) Due to too many missing observations, we could not use a more specific measure of mortality rates.
of loyalty or allegiance to the state, which increases rebel groups’ opportunity to recruit (Weinstein 2005). RURAL POPULATION measures the percentage of the population that live in rural areas (World Bank 2013).

EDUCATIONAL SPENDING: Previous studies have included measures of the level of education (Dixon 2009). EDUCATIONAL SPENDING, being an economic variable, is included as a measure of opportunity. The variable measures public spending on education as a percentage of GDP (World Bank 2013).

MOUNTAIN: Empirical findings suggest that countries with a relatively rough, mountainous terrain have a higher risk of civil war, all else being equal (Collier and Hoeffler 2004; Fearon and Laitin 2003). This is likely so because a country’s terrain helps determine military visibility, and hence opportunity for rebellion (Collier et al. 2003, 71). We therefore include the measure MOUNTAIN in the regression analysis, with data from Collier and Hoeffler (2004). Values for each country are kept constant over the entire time period.

5. Regression analysis

We first present the results from our regression analysis. The purpose of this is to get a clear view of the empirical relationship between ethnicity – proxied by our updated ELF Index – and conflict, before we in the next section investigate more specifically which mechanisms account for any such relationship.

There is no absolute agreement on whether or not ethnic fractionalization leads to civil war. Some find little or no such link (Collier and Hoeffler 2004; Fearon and Laitin 2003). Others do, however (Collier, Hoeffler, and Rohner 2009; Sambanis 2001). What all these studies have in common is that they have used a dependent variable measuring civil war onset
with a threshold of 1,000 battle-related deaths, which really signifies high-intensity conflict.

Scholars who instead use a threshold of 25 battle-related deaths – or similar measures capturing low intensity-conflicts as well – all find a significant relationship between ethnicity and conflict initiation (Ellingsen 2000; Hegre and Sambanis 2006; Regan and Norton 2005).

Table 1. Logistic regression analysis, civil conflict/war onset, 1985–2010

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 1 Conflict</th>
<th>Model 2 Conflict</th>
<th>Model 4 War</th>
<th>Model 5 War</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>0.315***</td>
<td>0.364***</td>
<td>0.123</td>
<td>0.129</td>
</tr>
<tr>
<td>GDP per capita</td>
<td>-0.286**</td>
<td>-0.370***</td>
<td>-0.496***</td>
<td>-0.511***</td>
</tr>
<tr>
<td>Fuel exporter</td>
<td>0.803**</td>
<td>0.823***</td>
<td>0.643</td>
<td>0.634</td>
</tr>
<tr>
<td>Mountain</td>
<td>0.004</td>
<td>0.001</td>
<td>0.010</td>
<td>0.010</td>
</tr>
<tr>
<td>Autocracy</td>
<td>-0.341</td>
<td>-0.234</td>
<td>0.068**</td>
<td>-1.054**</td>
</tr>
<tr>
<td>Democracy</td>
<td>0.658*</td>
<td>0.601*</td>
<td>-0.193*</td>
<td>-1.248*</td>
</tr>
<tr>
<td>Empowerment rights</td>
<td>-0.091**</td>
<td>-0.083**</td>
<td>-0.998</td>
<td>-0.101</td>
</tr>
<tr>
<td>EthfracIpolated</td>
<td>1.665***</td>
<td>7.712***</td>
<td>0.101</td>
<td>-0.819</td>
</tr>
<tr>
<td>EthfracIpolated2</td>
<td>-6.374***</td>
<td>-7.878</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temporal dependence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peace years</td>
<td>0.093</td>
<td>0.112</td>
<td>0.302*</td>
<td>0.305*</td>
</tr>
<tr>
<td>Spline(1)</td>
<td>0.002</td>
<td>0.002</td>
<td>0.021**</td>
<td>0.021**</td>
</tr>
<tr>
<td>Spline(2)</td>
<td>0.001</td>
<td>0.001</td>
<td>-0.025**</td>
<td>-0.025**</td>
</tr>
<tr>
<td>Spline(3)</td>
<td>-0.001</td>
<td>-0.001</td>
<td>0.019*</td>
<td>0.020*</td>
</tr>
<tr>
<td>Constant</td>
<td>-7.043***</td>
<td>-8.401***</td>
<td>-2.898</td>
<td>-1.933</td>
</tr>
<tr>
<td>Pesudo LL</td>
<td>-321.361</td>
<td>-317.774</td>
<td>-130.516</td>
<td>-130.492</td>
</tr>
<tr>
<td>Observations</td>
<td>2734</td>
<td>2734</td>
<td>2754</td>
<td>2754</td>
</tr>
<tr>
<td>Years</td>
<td>1985-2010</td>
<td>1985-2010</td>
<td>1985-2010</td>
<td>1985-2010</td>
</tr>
</tbody>
</table>

Notes: ◁Lagged one year; Log transformed; Anocracy as reference category; probability of F-value; *significant at 10%, **significant at 5%, ***significant at 1%; the units are clustered by country using Huber-White robust standard errors (White 1980).
Given these variations in terms of measures as well as results, we employ both of these dependent variables in our models. We do expect the low-threshold variable (CONFLICT) to show a significant relationship with the updated ELF Index. Our expectations are less certain as regards the dependent using 1,000 battle-related deaths as the threshold (WAR).

The findings from our regression models are presented in Table 1. The results in Model 1, where civil conflict is the dependent variable, show that there is indeed a highly significant relationship between the updated ethnic fractionalization variable – ETHFRACIPOLATED – and the risk of civil conflict onset. This is perfectly in line with much of the empirical literature (Ellingsen 2000; Hegre and Sambanis 2006; Regan and Norton 2005). An empirical relationship between ethnicity and civil conflict onset is thereby established, which creates the foundation to investigate further ethnicity as an indicator of civil conflict onset. In Model 2 we also test for a curvilinear association between ethnic fractionalization and civil conflict onset. The result for ETHRACIPOLATED2 shows that there is in fact such a link between these two variables – even in the presence of standard controls – taking the shape of an inverted U. This means, first, that the likelihood of civil conflict erupting is lowest in countries that are (almost purely) homogenous. Secondly, the risk increases as the level of fractionalization rise – but only up to a certain point. When we move past the vertex, above which countries are highly ethnically fractionalized, the likelihood of civil conflict onset actually decreases. This lends support both to those who argue that the size of ethnic groups matter greatly to the risk of rebellion (Ellingsen 2000; Montalvo and Reynal-Querol 2005) and to the contention that ethnicity is an indicator of opportunity.

This relationship is taken into account in the subsequent section’s factor analysis, where we code the variable ETHFRACWAR based on the binary relationship between ethnic fractionalization and civil conflict onset.
In Models 3 and 4 the dependent variable is civil war, measured at a high threshold of casualties. Here, neither ETHFRACIPOLATED nor ETHFRACIPOLATED2 is significant. This bolsters the findings of Collier and Hoeffler (2004) and Fearon and Laitin (2003), neither of whom found any significant impact of ethnic fractionalization on the likelihood of (high-intensity) civil war. We can therefore conclude that no empirical relationship exists between ethnic fractionalization and high-intensity civil war. Yet, our second conclusion is that there does exist such a relationship – and a fairly strong one to boot – between ethnic fractionalization and civil conflict, with the latter concept also including wars at a fairly low level of intensity. This result gives us a clear reason to attempt to identify the specific mechanism(s) that provide(s) this link. In the section that follow we do exactly this, seeking to establish empirically whether ethnicity is an indicator of grievance – as claimed by some (Collier and Hoeffler 2004; Fearon and Laitin 2003) – or an indicator of opportunity, or, as the third possible option, a determinant of civil conflict in its own right.

6. What mechanism links ethnicity and civil conflict?

In order to answer the research question – through what mechanism does ethnic fractionalization work as an indicator for civil conflict onset? – an exploratory factor analysis is performed, where we use an oblique promax rotation.\(^\text{13}\) Two separate models are presented in Tables 2 and 3, respectively; the difference between the two models is accounted for by the variable measuring ethnic fractionalization. Table 2 includes ETHFRACIPOLATED, whereas Table 3 contains ETHFRACWAR; except for this, the two factor analyses are identical. Each model identifies four different factors, each of which is given a name based on the theoretical content of the variables loading on the factor in question.\(^\text{14}\)

\(^\text{13}\) Factor scores under 0.35 have been omitted from the tables.

\(^\text{14}\) All together, the four factors explain 0.733 of the variance in Table 3, and 0.709 in Table 4.
Table 2. Factor analysis, ETHFRACIPOLATED, 1985–2010

<table>
<thead>
<tr>
<th>Variables</th>
<th>Wealth</th>
<th>Empowerment</th>
<th>Grievance</th>
<th>Identity</th>
<th>Uniqueness</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP per capita</td>
<td>0.940</td>
<td></td>
<td></td>
<td></td>
<td>0.095</td>
</tr>
<tr>
<td>Agriculture</td>
<td>-0.886</td>
<td></td>
<td></td>
<td></td>
<td>0.238</td>
</tr>
<tr>
<td>Bureaucratic quality</td>
<td>0.881</td>
<td></td>
<td></td>
<td></td>
<td>0.220</td>
</tr>
<tr>
<td>Life expectancy</td>
<td>0.820</td>
<td></td>
<td></td>
<td></td>
<td>0.171</td>
</tr>
<tr>
<td>Rural population</td>
<td>-0.813</td>
<td></td>
<td></td>
<td></td>
<td>0.289</td>
</tr>
<tr>
<td>Law and order</td>
<td>0.720</td>
<td></td>
<td></td>
<td></td>
<td>0.366</td>
</tr>
<tr>
<td>Educational spending</td>
<td>0.512</td>
<td>0.428</td>
<td></td>
<td></td>
<td>0.422</td>
</tr>
<tr>
<td>Military</td>
<td></td>
<td>-0.825</td>
<td></td>
<td></td>
<td>0.361</td>
</tr>
<tr>
<td>Fuel export</td>
<td></td>
<td>-0.805</td>
<td></td>
<td></td>
<td>0.391</td>
</tr>
<tr>
<td>Polity</td>
<td></td>
<td>0.804</td>
<td></td>
<td></td>
<td>0.204</td>
</tr>
<tr>
<td>Empowerment rights</td>
<td></td>
<td>0.751</td>
<td></td>
<td></td>
<td>0.205</td>
</tr>
<tr>
<td>Population*</td>
<td></td>
<td></td>
<td>-1.011</td>
<td></td>
<td>0.152</td>
</tr>
<tr>
<td>Physical integrity rights</td>
<td></td>
<td></td>
<td></td>
<td>0.636</td>
<td>0.231</td>
</tr>
<tr>
<td>EthfracIpolated</td>
<td></td>
<td></td>
<td></td>
<td>0.714</td>
<td>0.274</td>
</tr>
<tr>
<td>Population growth</td>
<td></td>
<td></td>
<td></td>
<td>0.542</td>
<td>0.390</td>
</tr>
<tr>
<td><strong>Eigenvalue</strong></td>
<td><strong>5.974</strong></td>
<td><strong>2.314</strong></td>
<td><strong>1.346</strong></td>
<td><strong>1.000</strong></td>
<td><strong>N=1185</strong></td>
</tr>
</tbody>
</table>

The variables that load on the first factor, both in Table 2 and Table 3, are measures that theoretically belong within the opportunity explanation of civil war onset. GDP PER CAPITA and AGRICULTURE have the highest loadings on this factor. Theoretically, these variables relay information about a country’s riches, hence we label this factor *Wealth*. With respect to *Wealth*, one difference between Table 2 and Table 3 is represented by EDUCATION SPENDING, which only loads on *Wealth* in Table 2. This has little or no bearing on the theoretical identification of the factor, however, because it loads only weakly compared to other typical opportunity proxies. The same goes for POPULATION GROWTH, which is associated with *Wealth* in Table 3 but not in Table 2.
Table 3. Factor analysis, ETHFRACWAR, 1985–2010

<table>
<thead>
<tr>
<th>Variables</th>
<th>Wealth</th>
<th>Empowerment</th>
<th>Grievance</th>
<th>Identity</th>
<th>Uniqueness</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP per capita</td>
<td>0.927</td>
<td></td>
<td></td>
<td>0.096</td>
<td></td>
</tr>
<tr>
<td>Agriculture</td>
<td>-0.950</td>
<td></td>
<td></td>
<td>0.225</td>
<td></td>
</tr>
<tr>
<td>Bureaucratic quality</td>
<td>0.737</td>
<td></td>
<td></td>
<td>0.328</td>
<td></td>
</tr>
<tr>
<td>Life expectancy</td>
<td>0.933</td>
<td></td>
<td></td>
<td>0.203</td>
<td></td>
</tr>
<tr>
<td>Rural population</td>
<td>-0.896</td>
<td></td>
<td></td>
<td>0.280</td>
<td></td>
</tr>
<tr>
<td>Law and order</td>
<td>0.617</td>
<td></td>
<td></td>
<td>0.401</td>
<td></td>
</tr>
<tr>
<td>Educational spending</td>
<td></td>
<td>0.493</td>
<td></td>
<td>0.610</td>
<td></td>
</tr>
<tr>
<td>Military</td>
<td>-0.828</td>
<td></td>
<td></td>
<td>0.360</td>
<td></td>
</tr>
<tr>
<td>Fuel export</td>
<td>-0.811</td>
<td></td>
<td></td>
<td>0.391</td>
<td></td>
</tr>
<tr>
<td>Polity</td>
<td>0.807</td>
<td></td>
<td></td>
<td>0.205</td>
<td></td>
</tr>
<tr>
<td>Empowerment rights</td>
<td>0.755</td>
<td></td>
<td></td>
<td>0.207</td>
<td></td>
</tr>
<tr>
<td>Population</td>
<td></td>
<td>-0.948</td>
<td></td>
<td>0.244</td>
<td></td>
</tr>
<tr>
<td>Physical integrity rights</td>
<td></td>
<td>0.663</td>
<td></td>
<td>0.232</td>
<td></td>
</tr>
<tr>
<td>Ethfracwar</td>
<td></td>
<td></td>
<td></td>
<td>0.985</td>
<td>0.027</td>
</tr>
<tr>
<td>Population growth</td>
<td>-0.517</td>
<td></td>
<td></td>
<td></td>
<td>0.558</td>
</tr>
<tr>
<td><strong>Eigenvalue</strong></td>
<td>6.230</td>
<td>2.347</td>
<td>1.343</td>
<td>1.071</td>
<td>N=1185</td>
</tr>
</tbody>
</table>

Those who argue that ethnic fractionalization is an indicator of the opportunity for civil conflict onset basically view ethnic groups as an extraordinary form of interest groups where group members have common preferences with regard to (virtually) all types of public policies (Fearon 2006). Ethnic groups must hence be conceptualized as rational actors whose decisions on whether to enter or initiate armed conflict are based on cost-benefit assessments and utility-maximization reasoning (Toft 2002). Findings that suggest that an ethnic group typically must be of a certain size in order for rebellion to be practically possible, further underpins the argument for ethnicity being an indicator of opportunity (Ellingsen 2000; Montalvo and Reynal-Querol 2005; see also the result on ETHFRACIPOLATED2 in Table 1).
As Tables 2 and 3 both demonstrate, though, ethnicity does not load on Wealth. Therefore, we reject this paper’s hypothesis H₂, which states that ethnicity is a mechanism of opportunity. Ethnic fractionalization, we conclude, is not a mechanism of opportunity, and therefore it does not belong to the realm of rational actor theory.

The second and third factors, both of which can be deemed components of the grievance dimension, are named Empowerment rights and Grievance, respectively. For factor number two this is so considering the grievance-based theoretical arguments that are associated with the variables MILITARY, POLITY, and EMPOWERMENT RIGHTS, all of which load strongly on Empowerment. FUEL EXPORT also loads on this factor, which at first glance presents a slight problem considering that it is often viewed as an indicator of opportunity. Still, FUEL EXPORT can certainly also be seen as a grievance indicator. A state or government will typically be unwilling to give up, or it will seek to acquire, stronger control over, areas where there are valuable natural resources. This may, in turn, cause sentiments of grievance among those who are affected by the objectives of the state. In particular, the discovery of oil, gas, diamonds, or other valuable resources often leads to the more or less forced removal of those living in the area (Toft 2002).

The third factor – Grievance – has high loadings of variables whose theoretical contents are clearly linked to grievance-based explications. In Table 2, this factor is made up of POPULATION and PHYSICAL INTEGRITY RIGHTS; whereas in Table 3, EDUCATION SPENDING also loads on this factor, though not as markedly.

Explanations of civil war that center on grievance are rooted in relative deprivation theory. The deprived-actor school assumes that deprivation causes anger, and that it therefore has a direct effect on the likelihood of any individual contemplating participating in violent action (Davies 1962; Graham and Gurr 1969; Gurr 1968, 1970). Such an explanation is linked to ethnicity through the argument that grievance is strongly associated with ethnic identity. If
a common grievance is shared among group members – that is, if it follows ethnic lines – the group more easily musters the coherence and sense of common purpose necessary to start fighting collectively for what each individual feels it is entitled to. Between-group inequality, in general, might also represent such grievance (Eckstein 1989; Midlarsky 1988; Sen 1973); for example, ethnic groups that are politically discriminated against are more likely to turn to armed rebellion than are other groups (Cederman, Wimmer, and Min 2009, 2010). Still, on this score the results presented herein lends no support to Collier and Hoefller (2004), who used ethnic fractionalization as a mechanism of grievance. According to our results, their assumption was wrong. Ethnicity, as it is, does not load on either of the grievance factors. Consequently, we must reject hypothesis H1.

As it turns out, it is the fourth factor that gives an answer to our overriding question of what mechanism provides the link between ethnic fractionalization and the initiation of civil conflict. In Table 2, ETHFRACIPOLATED loads on factor four, together with EDUCATION SPENDING and POPULATION GROWTH. Notably, it is the former that exhibits the highest loading on this factor; therefore, and considering the theoretical explanation of ethnicity that holds it to be an indicator of identity, this factor is named Identity. Also noteworthy, in Table 3 only one variable loads on factor four, namely ETHFRACWAR, which substantiates our inclination to view Identity as a factor in its own right.

We therefore conclude that ethnic fractionalization in itself is an indicator of civil conflict onset – based on the vitality of ethnic identity. According to Wolff (2006), ethnic identity has enormous power when it becomes the predominant constituent part of an individual’s personality. Ethnicity, to be sure, rests upon an amalgamation of several factors that are key to most individuals, such as language, religion, heritage and territory (Toft 2002). The latter of these – territory – forms an important part of the story of why ethnic identity can be considered as an indicator in its own right. This is the case considering that a substantial
part of ethnic-group identity rests upon the territory where they live (Toft 2002). Territory, to be sure, is typically vital for the survival of the group. Therefore, if an ethnic group sense that their domain is under threat, the salience of their shared identity becomes more apparent as the group might, should, or must mobilize in order to protect their territory and secure their sheer survival and identity (Toft 2002). According to primordialists, ethnic identity in itself suffices to spur a group to mobilize for armed rebellion, due to the fundamental characteristics of ethnicity, “ancient hatred,” and congenital inter-group differences. Not only do primordialists argue that ethnic identity is a sufficient factor in this respect, they also contend that it is virtually a necessary one; (just about) only identity can explain why civil wars break out (Blimes 2006).

Figure 4. Argument and results

In summary, as depicted in Figure 4, the results lend support to hypothesis H₃: *Ethnicity is in itself an indicator of civil conflict*. Or, to be more exact, ethnicity works as an indicator *through the mechanism of identity*. Neither those who operationalize ethnic fractionalization as an indicator for grievance, nor those who claim that it is closely associated with opportunity, receive any backing from the analysis herein.

Conclusion

The first notable contribution of our paper consisted of the presentation of an updated version of the ELF Index. That index reflects the probability that two randomly drawn individuals,
belonging to the same country, are members of different ethnolinguistic groups; thus it yields the level of ethnic fractionalization within a country. By way of interpolation, we were able to capture temporal changes in the ethnic composition of states. The resulting variables were then – in a logistic regression analysis covering the period 1985–2010 – shown to be significantly positively, though slightly curvilinearly, related to the outbreak of civil conflict. No such significant relationships were found when we employed a measure of high-intensity civil war as the dependent variable.

Secondly, we have also demonstrated that ethnic fractionalization, in its own right, is an indicator of civil conflict onset, due to the salience of ethnic identity. This means that Collier and Hoeffler (2004) and Fearon and Laitin (2003) are in the wrong when they employ ethnic fractionalization as an indicator of grievance. It follows that further research should include identity in any causal model that conceptualizes civil conflict onset. On the other hand, we do not claim that ethnic identity is sufficient for civil wars to be initiated. Primordialists, as it is, are unable to explain why not all ethnic groups fight each other; the variance must thus have other sources than only group identity. On this score, we follow Gurr (2000, 66), who claims that “[i]t is simplistic to argue that one kind of motivation is primary and the others subsidiary.” Civil conflict must rather be seen as a result of grievance, opportunity, and identity.
References


