Master's thesis

NTNU
Norwegian University of
Science and Technology
Faculty of Social Sciences
and Technology Management
Department of Sociology and Political Science

Hilde Anderssen Bakkan

Unpacking Ethnicity and Civil Conflict

Exploring which Underlying Mechanism Links Ethnic Fractionalization and Civil Conflict

Master's thesis in Political Science

Trondheim, spring 2013

NTNU – Trondheim
Norwegian University of
Science and Technology

Hilde Anderssen Bakkan

Unpacking Ethnicity and Civil Conflict

Exploring which Underlying Mechanism Links Ethnic Fractionalization and Civil Conflict.

Master's thesis in political science

Trondheim, spring 2013

Norwegian University of Science and Technology

Faculty of Social Science and Technology Management

Department of Sociology and Political Science

NTNU – Trondheim
Norwegian University of
Science and Technology

Aknowledgement

First, I would linke to thank my supervisor, Tor Georg Jakobsen, for useful advise and comments during the process. I would also like to thank Tanja Ellingsen, for helpful guidance concerning stucture and theory.

I also would like to thank Ida Baalsrud, Mia Dimmeldal, and Silje Sæther for discussion, support, and good spirits. Lastly Arne, for his patience, and for helping me construct the models in the thesis.

Trondheim, June 2013

Hilde Anderssen Bakkan

Contents

Figures

Tables

1 Introduction

"In the aftermath of elections in 2007, at least 1,200 people died after politicians whipped up ethnic tensions that triggered violence on an unprecedented scale" (Allen 2012). The tension within Kenya resulted in civil war along ethnic lines, and is one of several examples where ethnic dimensions play a role in the onset of civil conflict. In studies of civil conflict onset, ethnicity has an important function for two reasons; first ethnic groups make collective action easier to achieve. Secondly they are uniquely visible, and therefore more constraining (Kalyvas 2008: 1044). Most wars in developing countries have an ethnic dimension, based on the fact that there are conflict along ethnic lines (Murshed and Gates, 2005: 122; Buhaug, 2006: 697).

When Collier and Hoeffler (2004) wrote their hotly debated article "Greed and Grievance in Civil war"[1], they tested opportunity and grievance mechanisms for civil war onset.[2] Constructing their two separate models, they operationalized their measure for ethnicity, ethnic fractionalization, as an indicator of grievance. Ethnic fractionalization was also argued to be and indicator for grievance by Fearon and Laitin (2003). The aim and main contribution of this thesis is to answer: *Through what mechanism*[3] *ethnic fractionalization work as an indicator for civil conflict onset?* Is it right to use ethnic fractionalization as an indicator for grievance, as done by Collier and Hoeffler (2004), and Fearon and Laitin(2003) when modeling civil war onset or is there other alternatives? According to Ellingsen (2011: 100) the literature on civil war onset have a tendency to treat ethnic factors as a grievance mechanism, but "it is possible that the ethnic composition of a country actually is more related to feasibility and opportunity factors than what is argued". Whether ethnic fractionalization is an indicator of grievance or opportunity for civil war onset has not yet been investigated by researchers, it has been operationalized as a measure for grievance, but might as well be an indicator in itself.

In their article Collier and Hoeffler (2004) and to some extent Fearon and Laitin (2003) conceptualized civil war as a result of either opportunity or grievance, all of them finding that

[1] This article has recieved some criticism, and their model has in later years been developed (Collier, Hoeffler and Rohner 2009).
[2] Whether the reason for civil war onset is opportunity or grievance is not the main focus in this thesis.
[3] With mechanism is meant the three links between ethnicity, and civil war onset: grievance, opportunity and identity.

opportunity mechanisms best can explain civil war onset. By limiting civil war to be a result of either grievance of opportunity, other significant mechanisms might not be taken into account when making a causal model for civil war onset. This would be the case if ethnic fractionalization is an indicator for civil war onset in itself. An implication of this is that the model of civil war must be widened, also including identity as a mechanism for civil war. With starting point in the conceptualization of civil war as a result of either mechanism , and the link between ethnicity and civil conflict establish a theoretical framework.

In order to answer the research question three mutually exclusive hypotheses are presented. Introducing the arguments for the separate hypotheses, I first argued that ethnicity is an indicator for grievance. Second, I argue that it is an indicator for opportunity, and last that it is an indicator in itself. These arguments are presented based on a theoretical framework that is presented earlier in Chapter 2 where I first establish the central concepts civil conflict and ethnicity are defined. Following this, the underlying theory concerning the contemporary debate of whether civil war is a cause of grievance or opportunity is presented. Although the focus of this thesis is not whether civil war is a result of opportunity or grievance factors, an overview over the two directions is needed in order to understand how ethnicity can work as an indicator of either grievance or opportunity. After this is the connection between ethnicity and civil war is reviewed, separating between theoretical and empirical connections. A theoretical connection is presented based the school of primordialism, instrumentalism and constructivism. The empirical connection summarizes the empirical findings on the link between ethnicity and civil war in earlier research. Following this the three theoretical arguments mentioned over are presented leading up to the hypotheses. In all, three hypotheses are presented: $H_{1:}$ *Ethnic fractionalization is an indicator of opportunity*, H_2: *Ethnic fractionalization is an indicator of grievance*, and H_3: *Ethnic fractionalization is an indicator in itself.*

After building the theoretical framework the research design is demonstrated. First, in chapter 3 the second contribution to the research field is presented, an updated index of ethnolinguistic fractionalization (ELF). This index is used to measure the level of ethnic fractionalization and is the most frequently used variable when looking at the relationship between ethnicity and civil war (Kalyvas 2008). The ELF-index gives the probability of two randomly drawn people in a country coming from different ethnolinguistic groups (Fearon and Laitin 2003). The old index have values from early 1990, the new index is interpolated

between the values of the old and new one between 1990 and 2011. Following, the different statistical methods that are used in the thesis are presented stating with logistic regression with time-series cross-sectional data, the principle for this method and its challenges is demonstrated. Thereafter, the statistical method used to answer the research question, exploratory factor analysis, is presented. Last in this Chapter 3 the variables are described, starting with the dependent variables civil war, and civil conflict onset. Thereafter the other variables that are used in the analysis are introduces. These are sorted into either variables that proxies the grievance conceptualizing of civil war, or opportunity. After presenting the different variables results from the empirical analysis and discussion of the results follows in Chapter 4.

Based on the empirical analysis using logistic regressions it is established an empirical link between ethnicity and onset of civil conflict, this relationship is curvilinear taking the shape of an inverted-u. No significant relationship between ethnicity and civil war onset is found.[4] Following, results from two separate factor analyses are demonstrated, one based on the original interpolated ELF-index, the second taking into account the curvilinear relationship between ethnic fractionalization and civil war based on a recoded variable. Both factor analyses show that ethnicity is a mechanism of civil conflict onset in itself, confirming hypothesis H_3. In other words, hypotheses H_1 and H_2 are rejected, concluding that ethnicity is a mechanism for civil conflict onset in itself due to the salience of identity.

As a result of this ethnic fractionalization cannot be operationalized as a grievance indicator, as has been done by Collier and Hoeffler (2004), and Fearon and Laitin (2003). Neither is it an opportunity indicator as indicated by Ellingsen (2011), it is a mechanism in itself. Future research must take into account the aspect of identity when constructing a causal model explaining civil conflict onset. Lastly, civil onset as a combination of grievance, opportunity, and identity, not see it as a result of an "either-or" term as Collier and Hoeffler (2004) do in their article. Given that ethnic fractionalization lead to civil conflict due to the salience of identity, it is argued that identity alone cannot explain onset, based on some logical faults in the primordial argument. Civil conflict onset must therefore be seen as a result of grievance, opportunity, and identity as argued by Gurr (2000), and Ellingsen (2000).

[4] Civil conflict and civil war is defined in the next chapter.

2 Theory

First and foremost this part of the thesis will focus on establishing a link between different theoretical approaches, and ethnicity as an indicator of these in relation to civil conflict onset. In order to do this it is important to have an understanding of the central concepts civil conflict, civil war, and ethnicity. First the notions of civil conflict and civil war are established. Following this, ethnicity is defined. Thereafter the theoretical foundation for conceptualizing civil war as a result of grievance or opportunity is laid down. The connection between ethnicity and civil war is then presented, starting with the theoretical link as seen by primordialists, instrumentalists and constructivists, followed by the empirical link based on earlier research. Here different measures on ethnicity, in addition to the ELF-index are shortly presented, and results when using these measures. Based on the theoretical foundation ethnicity is argued to be an indicator of for civil conflict onset through either grievance, opportunity or an indicator itself due to the salience of identity is presented. To each of these arguments a hypothesis is presented, in all there are three mutually exclusive hypotheses.

2.1 Defining Civil Conflict

In the research on civil war there are mainly two dominant definitions, where the main difference is the fatality threshold. Collier et al. (2003) define civil war as taking place when "an identifiable rebel organization challenges the government military and the resulting violence results in more than 1,000 combat related deaths, with at least 5 percent of the casualties on each side" (Collier et al. 2003: 54). Within the datasets generating with a threshold of 1000 battle related deaths there are some differences regarding coding rules. Some register war only if there is more than 1000 battle related deaths within a year[5], others include the entire course of war.[6]

The second level of threshold is 25 battle related deaths within a year. This threshold is much lower and is referred to as civil conflict.[7] In this thesis there are two dependent variables, one based on the definition of civil war, the other one is based on the definition used by UCDP/PRIO[8]. They define an armed conflict as a "contested incompatibility that concern

[5] This is the coding rule applied for the dependent variable measuring war in this thesis.
[6] See Collier and Hoeffler (2004), Fearon and Laition (2003), Sambanis (2004), and the Correlates of War (COW).
[7] In the thesis the terms civil war and civil conflict is used interchangeably, this is also the case in the literature in the research field. In cases where it is important to note the difference it will be specified.
[8] Harbom and Wallesteen (2012); Gleditsch et al. (2002)

government or territory or both where the use of armed force between two parties result in at least 25 battle related deaths in one year. Of these parties, at least one has to be the government state" (Themnér and Wallensteen 2012: 572).

It is registered that over the last decades the nature of civil war has changed. Throughout the 1990s the number of civilian casualties reached 90 percent or more of the total number of deaths, a high proportion of these are children and women (Chesterman 2001: 2). But civil war is not only destructive in that a number of people that are killed', other parts of society are also affected, like the economy, infrastructure and general development. An additional effect can also be the spread of insecurity throughout the region, which would increase the risk of civil war in neighboring countries. A war-thorn country is also a country where opportunities for human trafficking and terrorism are more likely and society as a whole is ruined. But, despite the grave consequences of it, civil war still occurs, often along ethnic lines (Murshed and Gates 2005).

2.2 Defining Ethnicity

"Ethnicity is like family or marriage; everybody knows what it means but nobody can define it" (Smith, quoted in Alonso 1994: 379). There are innumerable definitions of ethnicity, which demonstrates that it is a difficult notion to get a clear grasp of (Horowitz 2000). According to Fenton (2010: 3) ethnicity is "something 'out there' which correspond to what observers call 'ethnicity'". Ethnicity as a term goes back to the Greek word "ethnos" and it was used to describe a kinship group linked together by bonds of blood (Wolff 2006). Ethnicity is based on a belief in a common heritage, which usually brings with it traits that are held to be innate (Horowitz 2000). These features are also found in the German national romantic conception of a state, often referred to as *das Volk*. The German understanding of a state emphasize on a common language, culture, and history, features that can be seen as innate. Religion has in later years also been regarded as an important feature (Ellingsen 2000).

Based on the view of ethnicity as "something 'out there'", and the link to kinship, common heritage, language, culture and also religion an ethnic group is in this thesis defined as a group made up by individuals who share common traits such as language, religion or race, a belief in a common heritage, and an association with a given territory (Toft 2002: 87). The salience of ethnic identity becomes recognized when ethnic identity has a collective consequence for a

group in its relation to states and other ethnic groups. In other words, ethnicity is likely to be highly significant when it is a basis for people's security, material well-being, status, or access to political power (Gurr 2000).

2.3 Grievance and Opportunity in Civil Conflict

As mentioned in the introduction the contemporary debate on whether civil war is a result of opportunity or grievance factors was introduces when Collier and Hoeffler (2004) presented their much debated article "Greed and Grievance in Civil War". In the article they present two models, one includes opportunity factors, and the other one includes grievance factors. The results from their analysis show that the opportunity model gives a better explanation for the occurrence of civil war compared to grievance related factors. Fearon and Laitin (2003) also use a similar distinction, and do also conclude that opportunity structures best explain civil war onset. Both articles use ethnic fractionalization as a measure on ethnicity, and use it in their analysis as an indicator of grievance. The aim of this thesis is to examine t*hrough what mechanism*[9] *ethnic fractionalization work as an indicator for civil conflict onset?*. Below the underlying theory of grievance and opportunity is presented. This is important because it demonstrates the conceptualizing of civil war as a result of either grievance or opportunity (Collier and Hoeffler 2004; Fearon and Laitin 2003), and give a basis before relating ethnicity to either of the concepts later in this chapter.

2.3.1 Grievance and Relative Deprivation

As mentioned Collier and Hoeffler (2004) found little support for grievance as an explanatory factor for civil war onset in their analysis. However there is no consensus among scholars that grievance does not lead to rebellion and civil war (Collier 2007). Grievance is a result of deprivation in an individual sense, which in turn establishes itself collectively. It is linked to political and material deprivation, and is often aggregated within specific homogeneous group, ethnic groups being one example of such a group. In other words it is the individual penury that is the starting point for collective mobilization (Regan and Norton 2005: 323). One way of explaining this individual grievance is by applying the relative deprivation theory onto the development of frustration and grievance.

[9] With mechanism is meant the three links between ethnicity, and civil war onset: grievance, opportunity and identity.

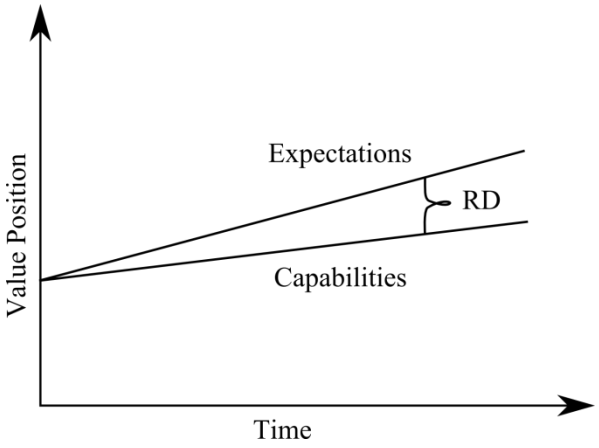
Relative deprivation was first used in a systematic way in the 1940s by Stouffer et al. (1949) who studied of the *American Soldier*. The aim of the study was to explain the feelings of an individual who did not get the conditions or status that he himself thought he was entitled to. The basic thought is that "persons may feel that they are deprived of some desired state or thing, in comparison with some standard, or with the real or imagined condition of other people" (Williams 1975: 355). The deprived actor school assumes that anger is a result of deprivation, and that it has a direct effect on the likelihood that a person will participate in violent action. One of the earliest contributions in newer time is Davies (1962), he made a substantial contribution in explaining the link between depression and violence. He argues that it is the "dissatisfied state of mind rather than tangible previous of 'adequate' or 'inadequate' supplies of food, equality, or liberty which produces the revolution" (Davies 1962: 6).

Gurr (1970) also had a significant influence on the theory of relative deprivation with the book *Why Men Rebel.* He states that the extent, and intensity of the deprivation is the basis for potential collective violence. The potential chain is "first the development of discontent, second the politicization of that discontent, and finally its actualization in violent action against political objects and actors" (Gurr 1970: 12f).

Graham and Gurr (1969) argue that it is the progressive form for deprivation that is the most dangerous one.[10] As shown in Figure 1 the distance between expectations and capabilities are the same in the beginning, both are increasing, but the capabilities cannot keep up with the expectations. The difference between the two is referred to as relative deprivation. Relative deprivation can be seen as a subjective characteristic that occur when a person does not get what he thinks he has the right to, the perception of being deprived of something you feel you are entitled to (Ellingsen 2000)

[10] Other forms of relative deprivation are decremental deprivation, aspirational deprivation and persisting deprivation.

Figure 1. Progressive deprivation



Source: Graham and Gurr (1969)[11]

It is also another approach that links grievances to the onset of civil conflict, inequality. Structural inequality within a state has been linked to several incidences of political violence, and revolutionary movement in the developing world. The inequality generates a sense of frustration for those that are not included (Regan and Norton 2005: 320). The distinction between relative deprivation and inequality may seem unimportant at first, but it is a difference that is worth noticing. Relative deprivation rests upon a set of psychological processes, and frustration is based on a gap between a person's individual expectations, and reality. Inequality on the other hand is based on individual expectations compared to *others* within society (Gurr 1970, Regan and Norton 2005). In literature on civil war onset. inequality between groups is referred to as horizontal inequality (Østby 2008). Common for both relative deprivation and inequality is that they can be used to argue that ethnicity function as a mechanism for civil war onset through grievance, this is done later.

2.3.2 Opportunity and Rational Actor

Opportunity as a part of the "greed" versus grievance debate has its roots in economic theory. Rebellion is a result of the circumstances, and the opportunities that make it possible. Arguing that civil war is a result of opportunity is developed from the rational actor theory. Rational actor theory in research on civil war onset was taken into use as a reaction to the grievance-based explanations that dominated during the 1960s (Jacoby 2008: 124). The theory argues that participation in protest groups, rebellion and social movements is not a result of psychological states or traits, but a result of "rational decision processes whereby people

[11] Relative Deprivation (RD).

weight the cost benefits of participation" (Klandermans 1984: 583). Olson (1965) argues that individuals do not participate in collective action based on grievance, but based on the possible gain from participation. It is the possible gain that must be seen as the trigger factor (Ellingsen 2000). The rational individual can only be stimulated to participate based on a "*separate and 'selective' incentive"* (Olson 1965: 51, emphasize his). In order to decide whether or not to participate each individual do a cost-benefit analysis. There must be a personal gain for a person to take part in the group. If the cost is too high compared to the benefit, the rational actor will not participate (Jacoby 2008: 126). The incentive must be selective, those who participate in a given group or organization, and works for the groups interests must be treated differently compared to those that do not participate (Olson 1965: 51). In other words the personal benefit can only be given to an individual if a person participates in the war; if people can get the benefit anyway it will be much more rational to be a free-rider.

> A rational individual who does not like the cost (loss of life, limb or family members), and in the absence of negative incentives, would be better off not to participate, whether or not other individuals do. Those who ultimately engage in revolutionary action do so on the basis of rational (cost-benefits) analysis. Individuals engage in revolutionary activity because they derive some personal gain (Conteh-Morgan 2004: 96, emphasize his).

At the start of the greed versus grievance debate, Collier and Hoeffler (2004), and Fearon and Laitin (2003) argued that rebellion was driven by greed, and the opportunity for personal benefits when participation. Collier and Hoeffler (2004) only focus on economic opportunity in their article. A similar grasp is done by Fearon and Laitin (2003), by removing the materialist factors from grievance, the materialist factor is the main theoretical argument for the opportunity factor. It is the opportunity to pursue material good that is the focus in the opportunity argument.

For rebel organizations victory in civil war is rare and not seen as an expected outcome for a rational actor. A rational actor will therefore not participate in rebellion based on the reward the person gets when they win the war it is the possibility of personal gains during the war that intrigues the individual to participate (Collier, Hoeffler, and Rohner 2009). Collier, Hoeffler, and Sambanis (2005) model civil war as an industry that produces rewards for those who participate in it. The explanation of civil war as a result of greed or feasibility assumes that rebels act in a pursuit of material gain based on self-interests, this is an argument that has

its roots in the rational actor theory (Regan and Norton 2005). It is only occasionally that rebel groups can be sustained both financially and military, based on this opportunity for personal gain is a significant factor. The financing of a rebellion is often associated with the availability of natural resources (Murshed and Tadjoeddin 2009: 88). The opportunity to loot natural resources gives rebel groups access to wealth, and a desire for wealth is motivated by some form of greed (De Soysa 2002: 396).

Collier, as one of the most influential researchers arguing for opportunity as a cause of war criticized the grievance explanation of civil war in several ways. According to Collier (2007: 198) the popular assumption that grievance is a cause of war has been shaped by the dissertation that a conflict in itself generates. Grievance is neither a cause of conflict, nor is it an unintentional by-product of it. Dismissing the grievance approach is done based on a believe that apparent grievance and desire for power is found more or less in all types of societies, and that it is the opportunity of feasibility of rebellion that distinguished the societies experiencing civil war from those that do not (Collier 2007). This demonstrates the "either-or" perspective that is used by Coller and Hoeffler (2004). But as mentioned before this is just a backdrop for the main focus in this thesis. It is not the debate of whether it is grievance or opportunity that can explain civil war onset that is center of the thesis, but through what mechanism for civil war onset that ethnicity functions as an indicator. In the following section I look closer at the connection between ethnicity, and civil war, both theoretically and empirically.

2.4 Connections between Ethnicity and Civil Conflict

The presentation of grievance and opportunity above is only linked to civil war onset, and not ethnicity In order to understand the relevance of the research question: *Through what mechanism does ethnic fractionalization work as an indicator for civil conflict onset?*, a connection between ethnicity and civil war must be demonstrated, this is done below. Fist, a theoretical relationship between ethnicity and civil war is presented through the three dominating perspectives: primordialism, instrumentalism and constructivism. Following, the empirical relationship between ethnicity and civil war is put forward based on a literary review on earlier research on this link. It is not only ethnic fractionalization that is used as a measure for ethnicity when researching the connection between ethnicity and civil war, a brief

presentation over other measurer of ethnicity, and empirical finding related to these are shortly presented.[12] But first, the theoretical link between ethnicity and civil war is presented.

2.4.1 The Theoretical Connection between Ethnicity and Civil Conflict

There are basically three different perspectives to understand the underlying mechanisms that drive ethnic groups to mobilize for violence: primordialism, instrumentalism, and constructivism. The primordialists see the characteristics of a group to be fixed. Each individual are born into their ethnic identity. Ethnic characteristics cannot change, and therefore they remain the same over time (Fenton 2010). A persons' ethnic identity is deeply integrated in human experience and history, and it cannot be rejected that it exists, subjectively or objectively. It must therefore be seen as a part of life in the relationship between groups and individuals (Wolff 2006: 33). Following the primordialist view on ethnicity the common features of speech, custom, history, and so on, must have an ineffable, and at time coerciveness in, and of themselves. A shared ethnic identity can become overpowering (Geertz 1973: 259). According to primoardialist civil war is a result of ethnic differences, and not a result of other factors in society, like economic or political differences (Blimes 2006: 537; Østby 2011: 28). According to Vanhanen (1999) civil conflict onset in ethnically divided societies is a result of the assumption that ethnic nepotism is part of human nature. The primordial view on ethnicity has been criticized for not taking into consideration that ethnic groups change over time. Further, they fail to explain why conflict between ethnic groups only erupts in some countries, and not in others (Blimes 2006).

The instrumentalists, on the other hand, criticize primordialists for simplifying the explanation of "ancient hatred" as a cause of civil war, claiming that they overlooks deeper problems that are the true reasons for conflict (Blimes 2006). Instrumentalists hold the position that ethnicity is only a tool that a group or an individual uses in order to achieve a goal (Blimes 2006: 537; Østby 2011: 29). "Ethnicicty is […] a resource in the hands of leaders to mobilize and organize followers in the pursuit of other interests, such as physical security, economic gain, or political power" (Wolff 2006: 33). In other words, ethnicity is in itself not the source for violent conflict. It becomes prone to war when leaders demonstrate that it is tied to a history of conflict, acute social uncertainty, and fear of what the future might bring, only then

[12] By doing this I also touch upon a large debate in the research on civil war onset, how to measure ethnicity, due to the focus of the thesis and limitations to the length of the paper I will not go deeper into this debate. See Cederman and Girardin (2007), Fearon, Kasara, and Laitin (2007), and Cederman, Wimmer, and Min (2009) for further information on the debate.

is it that ethnicity emerge as a major fault line along social fractures (Lake and Rothchild 1998: 7). This means that ethnic identity can be used as a mobilization factor for collective violence by leaders, but not as a conflict factor in itself as the primordialists hold (Nordås 2012).

The third perspective on the connection between ethnicity and violence is constructivism. This school tries to bridge the primordialist and instrumentalist views. The debate linked to constructivism dominated in the 1980s and 1990s (Kanbur, Rajaram and Varshney 2011: 148). Due to its position as a bridge constructivism sees ethnicity as a result of common tangible features. Constructivists believe that ethnicity is neither fixed, nor is it completely open for change. Ethnicity is partly inherited and partly chosen (Østby 2011: 28ff). At birth several identity categories are given to each individual, only some of them being relevant at different times in life. The significance of ethnic identity may vary. Further, the significance and content of ethnic group identity can and do change over time, and usually it is a response to a change in the groups' political and social environment (Gurr 2000: 4; Wolff 2006: 34). In today's literature on civil war it is the constructivist that in large parts dominates. Table 1 presents a short overview over the three different perspectives introduced above.

Table 1. Three theoretical perspectives on ethnicity and civil war

Primordialists	Each individual are born into their ethnic identity. Ethnic identity is fixed.
Instrumentalists	Ethnicity is a tool that a group or an individual uses in order to achieve a goal, often associated with political leaders.
Constructivists	Ethnicity is neither fixed, nor is it completely open for change. Ethnicity is partly inherited and partly chosen.

Source: Fenton (2010), Østby (2011)

2.4.2 The Empirical Connection between Ethnicity and Civil Conflict

Looking at the empirical link between ethnicity and civil war the main question asked regarding ethnicity on a macro-level is whether or not ethnicity is significant for the likelihood of civil war onset or not (Kalycas 2008). Following is a review of the results when

researching the connection between ethnicity and civil conflict. First, the research that has used ethnic fractionalization as a measure for ethnicity are presented, the same measure used in this thesis. Following, other measures on ethnicity and belonging results are presented briefly.

As mentioned Coller and Hoeffler (2004), and to some extent Fearon and Laitin (2003), operationalize civil war as either explained by opportunity or grievance, and ethnic fractionalization is used as an indicator for grievance. Testing the relationship between ethnicity and civil war onset neither of them finds a significant relationship. In other words there is according to them no connection between the level of fractionalization within a country and the risk of civil war onset. Contrary to Collier and Hoeffler (2004), and Fearon and Laitin (2003), Sambanis (2001) found that an increasing level of ethnic fractionalization is significantly and positively correlated with the likelihood of civil war onset. Distinguishing between protest, rebellion, and civil war, Regan and Norton (2005) demonstrated that ethnic fractionalization is positively and significantly associated with rebellion and civil war, but not protest. Hegre and Sambanis (2006) also found a link between ethnicity and civil war in low intensity wars,civil conflict, but contrary to Regan and Norton (2005) they found no relationship between the two in full-scale civil war. Collier, Hoeffler and Rohner (2009) combine a measure of ethnic fractionalization and religious fractionalization, and called it social fractionalization. They found that social fractionalization increases the risk of civil war onset significantly. This is contrary to results from earlier analyses done by Collier and Hoeffler (2004) where they found no significant relationship between ethnic- or religious fractionalization and civil war onset. As a robust check of their analysis Coller, Hoeffler, and Rohner (2009) used civil conflict onset as the dependent varibale, doing so they found social fractionalization to be positively and significantly associated with the risk of civil conflict onset. Buhaug (2006) also a threshold of 25 battle related deaths and found that an increase in ethnic fractionalization increases the risk of civil conflict onset. Common for the results presented over is that they are all based on a version of the ELF-index. The results are inconclusive, but there is an overweight of researches who found ethnic fractionalization to have a significant indicator for civil war, and conflict onset.

Of the research mentioned above it is only Coller and Hoeffler (2004), and in part Fearon and Laitin (2003) that proxy's ethnic fractionalization as an indicator for grievance. Cederman, Wimmer, and Min (2009) also link ethnicity to grievance, but do not use ethnic

fractionalization as a measure for ethnicity. Contrary to the research presented, they argue that highly diverse societies are not more prone to conflict than other countries, but that ethnicity and civil war is linked together due to different ethnopolitical configurations. Armed rebellion is much more likely if a large proportion of the population is excluded based on their ethnic identity. Violence is also more likely if a large number of competing elites share power in a state that is segmented. It is grievance as a result of inequality that leads to civil war. In order to test their argument Cederman, Wimmer and Min (2009) have developed a dataset that "identifies all politically relevant ethnic categories around the world and measure access to executive-level state power for members of the ethnic categories in all year from 1945–2005" (Cederman, Wimmer, and Min 2009: 325), the dataset is called *Ethnic Power Relations* (EPR). The results of their analysis show that, based on an underlying structural factors, the onset of civil war is significantly correlated with the share of the population that are excluded on the basis of their ethnic background (Cederman, Wimmer and Min 2009; 2010).

Minorities at Risk (MAR) is another dataset looks at ethnic groups in relation to other ethnic groups. It is based on the assumption that minority ethnic groups are discriminated, which in turn results in grievance. It documents that status of ethnic and religious minority groups in all countries in the world after 1946. Using the data in research indicate that ethnonationalists groups are more likely than other groups to adapt strategies of rebellion (Gurr 2000: 232).

Like Cederman, Wimmer and Min (2009), Montalvo and Reynal-Querol (2005) argue that highly diverse countries not are more prone to conflict than other countries. Instead Montalvo and Reynal-Querol (2005) argues that countries that are polarized have a higher likelihood of civil war onset. Measuring ethnicity by using an index of ethnic polarization, often called the RQ-index[13], Montalvo and Reynal-Querol (2005) found that ethnic polarization has a positive effect on the likelihood of civil war onset. In other words countries that are homogenous or heterogeneous have a lover likelihood for civil war onset than polarized counties. That the level of ethnic polarization matters for the probability of civil conflict onset, demonstrates that the size of the ethnic group compared to one another also matter. This is also found by Ellingsen (2000).

[13] Inspired by the Esteban and Ray (1994)

Using a different variety of categories Ellingsen (2000) found that in countries where the largest ethnic group constitutes less than 80 percent of the total population, the risk of civil war onset is 1.3 times higher than in societies where the largest ethnic group constitutes over 80 percent of the population.[14] Further, the analysis give that in countries where the second largest ethnic group makes up 5 –20 percent of the population, the likelihood of civil war onset is higher than in countries where the second largest group is under 5 percent and, over 20 percent. Using a very similar measure on ethnicity as Ellingsen (2000), Hegre et al. (2001) found that ethnic heterogeneity increases the probability of civil war. The likelihood of onset is twice as high in countries where the largest ethnic group makes up only half of the population compared to countries where it constitutes more than 95 percent of the population. Finding that the level of polarization have an effect on the likelihood of civil war onset, and that the size of the ethnic group matter Montalvo and Reynal-Querol (2005), and Ellingsen (2000) implicate that ethnicity is linked to opportunity, this will be discussed later.

Above the empirical connection between ethnicity and civil war has been reviewed. Later in in the thesis the relationship between ethnicity and civil conflict is tested empirically tested once again mainly to test the new ELF-index, but also to test for a curvilinear relationship between ethnicity and civil conflict onset. Montalvo and Reynal-Querol (2005), and Ellingsen (2000) both found it to be less violence in homogenous and heterogeneous (fractionalized) countries. The best way to control for this is when using ethnic fractionalization as a measure for ethnicity is to control for a curvilinear relationship. Since the ELF-index only is a secondary contribution, and the empirical relationship between ethnicity and civil war is a minor part of the theses, hypotheses are not presented. Testing the empirical relationship between ethnicity and civil conflict helpful in order to demonstrate why it is important to look closer at ethnicity as an indicator for civil war onset, and what mechanism it works through. In the following chapter ethnicity as an indicator of either grievance, opportunity or identity is presented, resulting in three mutually exclusive hypotheses.

2.5 Three Explanations of the link between Ethnicity and Civil Conflict

Above the theoretical foundation for opportunity and grievance was presented. As pointed out by Ellingsen (2011) literature today has a tendency to look at ethnicity as a grievance indicator for civil war, but it can also be seen as an indicator for opportunity. In the following

[14] Two dependent variables war used, both gave the same results. Civil war onset is for the period 1946-1992, civil conflict onset for the period 1989-1992.

I argue that ethnicity can be seen as an indicator of grievance, opportunity or identity. Theory presented earlier in the chapter will be applied in order to build up separate arguments. Three separate hypotheses are presented, all of them being mutually exclusive. I start by arguing that ethnicity is an indicator for grievance, this is in line with Collier and Hoeffler (2004) and Fearon and Laitins' (2003) way of operationalizing ethnic fractionalization. Following this, I argue that ethnicity is an indicator of opportunity as indicated by Ellingsen (2011). Last I argue that ethnic identity is an indicator for civil war onset in itself, this is based on the primordialist conception of ethnicity, and the salience of territory for an ethnic group.

2.3.2 The Link between Ethnicity, Grievance, and Civil Conflict

In their attempt to make a clear distinction between proxies for opportunity and grievance Collier and Hoeffler (2004) included ethnic fractionalization as an indicator for grievance. Following I argue that ethnicity can be seen as an indicator of grievance

Toft (2002: 84) argues that the link between ethnicity and grievance is founded in an individual's ethnic identity, along with a common group identity. Ethnicity is a fundamental factor in people's identity, and being a fundamental factor it is adequate in the process of grievance based mobilization (Ellingsen 2000). According to Murshed and Gates (2005: 122) it is not sufficient to only have a common identity, a shared grievance must also be present. Based on the relative deprivation theory, grievance is triggered by the expectations compared to the actual conditions. Relative deprivation can be linked to ethnicity through a common feeling of deprivation. It is the feeling of frustration that results in conflict and war, if the feeling of frustration occurs along ethnic lines ethnicity is an indicator for grievance

Approaching grievance through inequality, and discrimination it can also be argued that ethnicity is an grievance indicator. Here the focus is inequality between groups, also called horizontal inequalities (Østby 2008). Political discrimination is an example of inequality that results in grievances, and is helpful when explaining the link between ethnicity, grievance, and civil war. Ethnic groups that are politically discriminated, have a higher likelihood of turning to violent actions than ethnic groups that are not politically discriminated (Cederman, Wimmer, and Min 2009, 2010). In for example sub Saharan Africa, and South-, and Southeast Asia, ethnically based political parties are common. Access to political benefits is often structured along the same ethnic lines as the politics in the same countries (Fearon 2006). The

inequality between ethnic groups is linked to grievance, and "the relation between inequality and rebellion is indeed a close one" (Sen 1973: 1).

From an instrumentalist perspective on ethnicity can also be used when explaining the link between frustration, ethnicity, and civil war onset. Regan and Norton (2005) do this when they argue that rebel leaders use the common grievance within an ethnic group in order to mobilize them to participate in civil conflict onset. The rebel leaders mobilize ethnic groups in order to secure their own personal gain, rebel leaders can be seen as driven by opportunity whereas the larger ethnic group is driven by frustration. Based on the established link between ethnicity, grievance, and civil war onset demonstrated through deprived actor theory, horizontal inequalities, and an instrumentalist view on ethnicity, the first hypothesis in the thesis is presented:

H_1: Ethnicity is an indicator of grievance.

2.3.4 The Link between Ethnicity, Opportunity, and Civil Conflict

As the argument over show, ethnicity might very well be an indicator of grievance in line with the assumption done by Collier and Hoeffler (2004). Still arguing within the "useful conceptual distinction" (Collier 1999: 1) of grievance and opportunity, I argue that ethnicity also can be used as an indicator of opportunity.

The opportunity approach to civil war view ethnic groups as an extreme form of interest group where members have a common preference in all types of public politics. Ethnic groups are political coalitions that are present in order to take care of its members (or leader's) economic interests (Fearon 2006: 861f). Toft (2002: 85) concludes that ethnic groups must be seen as rational actors, they calculate, and maximize their utility before entering or initiating a civil war. In order to secure the survival of an ethnic group, resources and wealth is important, and ethnic groups can be seen as rational actors rebelling when they have the opportunity to secure or get access to resources and wealth. This is in line with Collier and Hoeffler (2004), and Fearon and Laitin (2003) that have a view on opportunity based on economical and material gains, greed are the motive for rebellion.

Ethnic groups can be seen in light of the rational actor theory. Ethnic groups and individuals belonging to the group calculate their collective interest before entering or starting a conflict

(Fenton 2010: 74). Their collective interest can have roots in securing financial powers through looting natural resources, or defending natural resources within the ethnic groups' territory (Murshed and Tadjoeddin 2009). When there is competition about scarce resources, ethnic identity become more important. In line with the rational actor theory a cost-benefit analysis is performed, and if the benefit is perceived as greater than the cost, given the probability of success, the likelihood of civil conflict increases.

> The rational choice theory on ethnic collective violence is based on the assumption that individuals have given goals, wants, tastes or utilities. Since all goals cannot be equally realized because of scarcity, individuals will choose between alternative courses of action so as to maximize these wants and utilities. The resulting action may be seen as the end product of two successive filtering devices. Individuals choose the most efficient way to realize their goal (Hechter, Friedman and Appelbaum 1982 415–16).

Taking into account the findings done by Montalvo and Reynal-Quearol (2005), and Ellingsen (2000) the size of an ethnic group is also important when arguing for ethnicity as an indicator for opportunity. If the largest ethnic group is under 80 percent there is a higher likelihood of civil war onset, than if the group is over 80 percent (Ellingsen 2000). If a country is close to homogenous a smaller ethnic group does a cost-benefit analysis, and concludes that they have no change to increase their well-being if starting a rebellion. On the other hand, if a country is fractionalized there will be several groups of some size, but not large enough to dominate, and calculate with victory. This demonstrates that there is a close link between ethnicity and opportunity, and that ethnicity can be seen as an indicator for opportunity. Based on this and an ethnic individual and group as a rational actor the second hypothesis is presented:

> *H_2: Ethnicity is an indicator of opportunity*

2.3.5 The Link between Ethnicity, Identity and Civil Conflict

The third link sees a shared ethnic identity as a sufficient mobilization factor for civil war onset. Both explanations for civil war onset presented over do not see ethnicity as a direct cause in itself, but below I argue that it is.

Moving outside of Collier and Hoefflers (2004) two models for civil war, it is argued that ethnic identity, and ethnic group identity is sufficient in order to mobilize for civil war (Harff and Gurr 1993). Identity "separates us from them", it includes relations within, across and

outside those boundaries giving meaning to both boundaries and relations (Tilly and Tarrow 2007: 79). Ethnicity as a cause of war in itself demonstrates the importance of ethnic identity for each individual. Ethnicity has an enormous power when it becomes a predominant identity (Wolff 2006: 31). It is the fundamental characteristics of ethnicity that makes it so powerful. These are characteristics such as language, religion, heritage and territory, and are the same characteristics that are central in the definition of ethnic groups (Ellingsen 2000). Ethnic groups already hold the advantage of common ideology, values and culture. This makes it easy to recruit participants for rebellion According to Harff and Gurr 1993) the strength of the group identity is one of the most important factors that can explain why ethnic groups participate in civil conflict. Ethnic bonds are very emotionally charged, and to some degree also irrational (Østby 2008).

The role of ethnic identity as a powerful mechanism for civil war onset in itself, is according to Toft (2002) linked to the importance of territory. Territory is not only linked to the identity of the ethnic group, but it is also an important factor in the survival of the group. The link between territory and ethnicity is shown empirically by Buhaug (2006) who find ethnic fractionalization to be significant for onset of territorial conflict, but not governmental conflicts.[15] It is only natural that ethnic groups live within the same territory and that territory function as a security for their survival and identity (Horowitz 2000: 201) If their territory is threatened, the salience of ethnic identity become apparent, and the group must protect their territory in order to secure their identity and survival (Toft 2002).

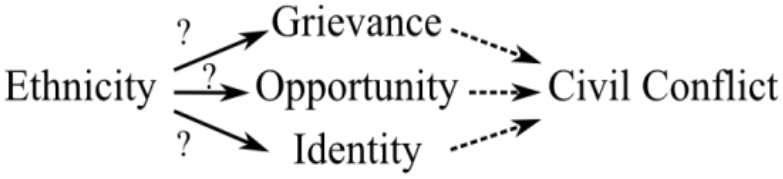
In other words ethnicity must be seen as a mechanism in itself based on identity, and not as a function of either grievance or opportunity. How much a person identifies with an ethnic group depends on the number of traits that are shared by group members (Harff and Gurr 1993). Arguing that ethnic identity in itself is a mechanism for civil war onset is based on a primordialist view on the connection between ethnicity and conflict. According to primordialists civil war is only a result of differences between groups, an no other factor in society, like political or economic differences. Conflicts involving ethnic groups are a result of "ancient hatred" (Blimes 2006; Østby 2011). The salience of ethnic identity is a direct mechanism for civil war onset. Based on this the last hypothesis is presented.

[15] A territorial conflict is defined to be involving demands for autonomy and secession. Governmental conflict are concerning the replacement of the central government, type of political system, or change in the composition of the government (Gleditsch et al. 2002: 619).

H_3: Ethnicity is an indicator for civil conflict in itself

In this chapter the theoretical framework is built up in order to answer the research question: *Through what mechanism does ethnic fractionalization work as an indicator for civil conflict onset?* Based on Collier and Hoeffler (2004), and Fearon and Laitin (2003) civil war can be conceptualized as either grievance or opportunity. Based on this, the concept of grievance and the deprived actor theory, followed by opportunity and the rational actor theory was presented. Thereafter the theoretical and empirical relationship between ethnicity and civil war was put forward. I argued ethnic fractionalization is an indicator of either grievance, opportunity or in itself, due to the salience of ethnic identity, using theory presented earlier in the chapter. This resulted in three mutually exclusive hypotheses. Model 1, below, demonstrates the theorized mechanisms between ethnicity, and civil war. As the model show it is only the arrows on the left that is important in order to answer the research question, ethnicity has above been theorized as an indicator of each of them:

Figure 2. Overview of theoretical framework



3 Research Design

In this chapter I first present the updated measure of ethnic fractionalization. A short chapter on methodology follows, before looking at logistic regression analysis using time-series cross-sectional data. Factor analysis is then presented, this is the method used to answer the research question. Thereafter, I present the dependent variables civil conflict and civil war. Last the variables used in the regression analysis and factor analysis is presented, theorized as either grievance variables or opportunity variables.

3.1. Updated measure on ethnicity

This thesis has two contributions to the research on civil war onset. For the first one of them the theoretical framework is presented over, looking at through what mechanisms ethnicity does works as an indicator for civil conflict onset. The second contribution is an updated version of the ethnolinguistic fractionalization (ELF) index. Measuring ethnicity with the ELF-index is the most commonly used measure in quantitative research (Collier 2001: 132; Kalyvas 2008). The index gives the likelihood of two randomly drawn individuals within a country belongs different ethnic groups, taking on values between 0 and 1. It is calculated based on the Herfindahl calculation formula[16], by subtracting the value of the Herfindahl formula from 1, the ELF-value is given as:

$$ELF = 1 - \sum_{i=1}^{n} s_i^2 \text{, where } s_i \text{ is the share of group i (i=1,…,n)} \quad (1)$$

The ELF index is originally based on a global survey of ethnic groups published in *Atlas Narodov Mira* (1964), and was first calculated by Taylor and Hudson (1972). In the period after Taylor and Hudson (1972) other variants of the index measuring fractionalization have been put forward. Alesine et al. (2003) distinguishes between ethnic, religious and linguistic groups. They use their list of distinctions between the three, and develop a fractionalization index for ethnic, religious, and linguistic fractionalization. It is calculated using the same formula as the ELF-index, but it is based on other data (Alesine et al 2003). Developing an

[16] The index was developed by the economists Herfindahl and Hirschman, it is originally named the Herfindahl-Hirschman index. It is a statistic measure of concentration. The index is commonly used to account for the number of firms in a marked, and by incorporating the relative size of the firm it gives the concentration. It is calculated as follows $HHI = \sum_{i=1}^{n} s_i^2$, where s_i is the share of firms *i* in the marked, and *n* is the number of firms. The calculation is done by squaring the marked shares of all the firms in the marked. Calculation the ELF-index it is the relative size of an ethnic group that is studied instead of firms (Kelly 1981).

index for cultural fractionalization, Fearon (2003) argues that it is important to measure the "ethnic distance" across groups. The distance is measured in three diagrams of families, of languages, and of different countries. Based on the data he collected a new ELF index with some additional characteristics was applied to research, the index was based on data from early 1990.

Since the early 1990s the world has changed, and it is about time that the measure on ethnic fractionalization is updated. As a contribution to the research field I have done this in association to the thesis. Similar to Alesine et al. (2003), and Fearon (2003) the new ELF-index is coded based on secondary data. Contrary to the earlier expansions, and recoding of an ELF-index, the updated version is calculated on no other features than the Herfindahl-index. An updated ELF-value has been calculated for all countries in the world, and has been done based on available information, mainly from Ethnic Groups World Wide, CIA World Factbook, Joshua Project and Library of Congress. Using the different sources as a control for one another, the most updated measures have been applied. Information from the different sources was mostly consistent with one another.[17]

Coding the composition for ethnic groups in all the countries around the world is based on an arbitrary process, it is difficult to answer the question "What are the ethnic groups in this country?" with a single right answer (Fearon 2003: 197). This gives room to the instrumentalist and constructivist view on ethnicity as a fuzzy, and situational character (Fearon 2003: 197). Or as Fenton (2008: 3) puts it ethnicity is "something 'out there'", which corresponds to what observers call ethnicity. Fearon (2003) points to the problems related to Hispanics as an ethnic group in the USA, why should they not be divided into Cuban Americans, Puerto Rican Americans, Mexican Americans, and so on? The same difficulty is applicable for Arabs from different countries. In the work of coding the updated measure used in this thesis, the term Arab and Hispanic has been applied. It has not been distinguished any further between the differences within these groups. The problem presented by Fearon (2003)

[17] CIA Factbook often referred to a given percentage, noting that this percentage of the population consisted of several ethnic groups. One example of this is Democratic Republic of Congo, data presented by CIA factbook summarize the three largest ethnic groups to make out 45 percent of the population, whereas the country as a whole have more than 200 ethnic groups. In these cases Joshua Project has been used listing several of the ethnic groups in the country, although the exact accuracy can be discussed it does undeniably give a much more appropriate ELF-value for the country.

illustrates that it is not straight forward to define and update the measures for ethnic fractionalization.

Some critique has been directed towards the use of the ELF-index in quantitative research. It has by the constructivist perspective on ethnicity been critiqued for being a constant measure, having the same value for a country over a time period of time (Laitin and Posner 2001; Kalyvas 2008). As an answer to this critique, the updated ELF-index is not seen as an independent index from the former ones, and the variable used in the empirical analyses is based on an interpolation[18] between the ones presented by Fearon (2003) and the updated ones from 2011. The variable ETHFRACIPOLATED has been kept constant up until 1990 (6 time units in the dataset), using the data from Fearon (2003). In the time period between 1990 and 2011 the values has been interpolated. By interpolating between the two time units the changes in the concentration of ethnic groups is reflected in the variable

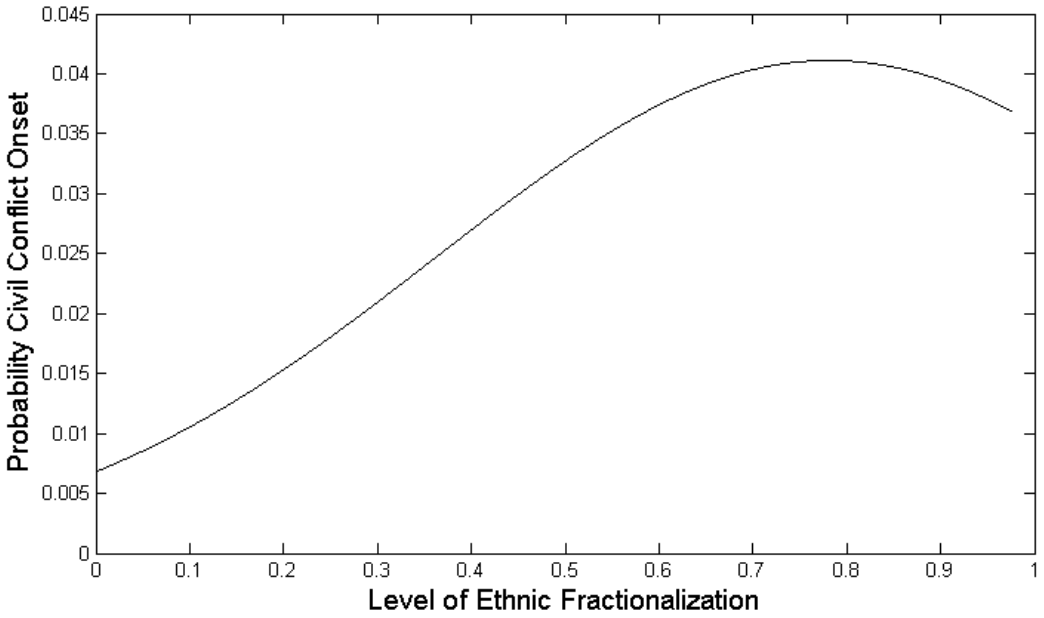
As an addition to a factor analysis using the variable ETHFRACIPOLATED as a measure of ethnicity, it is of interest to include a variable that link together ethnicity and civil conflict in a second factor analysis. Including such a variable gives a representation of the relationship between civil war and ethnicity in the factor analysis. Further, it functions as a robust check for the factor analysis that includes ETHFRACIPOLATED as a measure for ethnicity. The second variable measuring ethnicity is as mentioned pinpointed towards war, and is therefore called ETHFRACWAR. Based on results of former research on civil war onset it is reason to assume that there is a curvilinear relationship between ethnic fractionalization and civil war onset (De Soysa 2002; Motalvo and Reynal-Querol 2005). By doing a logistic regression analysis[19] with civil conflict onset as the dependent variable, ETHFRACIPOLATED and the squared term of ETHFRACIPOLATED the vertex of the relationship between civil conflict onset and ethnicity is calculated.[20] As shown in Figure 3 the curvilinear effect takes the shape of an inverted-u, with a vertex when the level of fractionalization is equal 0.78.

[18] Interpolation is a method that draw a straight line between the two known points, in this case the old ELF-values and the new ELF-values. The values at the different time units will be used to fill in the values in between the two points.

[19] Time-series cross-sectional data is used, this method is presented in chapter 3.3.

[20] See appendix A, Table A1 for the results from the logistic regression analysis.

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



Recoding ETHFRACIPOLATED, so that it is pinpointed toward that likelihood of civil conflict onset, ELF-values equal to 0.78 was coded as 1. Due to the curvilinear relationship, and the property of a second order function as being symmetric around the vertex, ELF-values equal to 0.77, and 0.79 were coded equal to 0.99, resulting in the new variable named ETHFRACWAR.[21] Both of the variables ETHFRACIPOLATED and ETHFRACWAR are included in two separate factor analyses presented in the next chapter. ETHFRACIPOLATED is first used to test the new ELF-index up against civil conflict, and to test for an empirical connection between ethnicity and civil conflict onset using well established variables in quantitative analysis on civil war onset as control variables. Before presenting the results from these analyses the different statistical methods that are used will be presented below, but first a short overview over the methodology.

3.2 Methodology

Concisely stated, *ontology* can be explained as the study of realtiy, while *espistemology* is the study of the nature of knowledge (Hay 2002; Moses and Knutsen 2007). The ontological position in this thesis is that there exists a real world that is independent of our perception of it. This view is in line with the positivist tradition, which state that the world is made up of

[21] 0.76 and 0.80 were coded as 0.98. 0.75 and 0.81 were coded as 0.88 and so on.

regularities. The positivist tradition gives a epistemological view that these regularities can be detected, and that the researcher can gather knowledge about the real world by observing it (Moses and Knutsen 2007). In order to gather knowledge about the world a method is needed. *Methodology* can according to Hay (2002: 63) be described as, "the choice of analytical strategy and research design which underpins substantive research", or as the logic behind the methods the researches chose. Following the positivist tradition, the statistical method for testing hypotheses is chosen in this thesis, this is the closest approach to experimental method Because of its ability to order and control temporal and casual relationships the experimental method is often seen as the best positivist method (Moses and Knutsen 2007).

The goal of scientific research is to draw conclusions that go beyond the collected data. By doing large-*N* studies it is possible to make generalizations about the casual effect of different phenomena. In this thesis a whole population of countries is examined, and generalized with *stochastic model theory*. In time-series analysis there are trends, this means that it is a stochastic variation around a non-constant mean. If one assumes that the model is deterministic, the variation around the mean is constant. Assuming that there is a stochastic trend in the data it allows for more flexibility in the estimation and form, which results in a model that is better fit to the data (Pevehouse and Brozek 2008. 461)

3.4 Logistic Regression with Time-Series Cross-Sectional Data

The first analysis presented in the next chapter, Analysis, is a logistic regression analysis[22] with the dependent variable civil conflict onset and civil war onset. These are both dichotomous variable with the categories being onset (=1) and no onset (=0)[23], based on this a maximum likelihood logistic model applied. According to Hamilton (1992) this approach gives a model that is easy to interpret. Time-series cross-sectional (TSCS) data is used to analyze data that has repeated observations, on the same fixed unit. In this thesis the repeated observations are yearly, and the units are countries (Beck 2011). Since the dependent variables are dichotomous logistic regression will be applied, the method can be referred to as binary time-series cross-section (BTCSC) method (Beck, Katz and Tucker 1998). Table 2, presented in the empirical analysis is a logistic regression with BTSCS-data. Using time series or, longitudinal data, the assumption of independent variables in logistic regression breached.

[22] Due to the secondary importance of this method in the thesis, the challenges assumptions for logistic regression will not be discussed. The focus is on the challenges when performing logistic regression with TSCS data.

[23] More info about the variable in section 3.3.1

Longitudinal is repeated measurements of different variables, and can be defined as data collected based on observed units of *n* number of variables over a period of time (Hox 2010: 109).

A challenge when using TSCS data is the temporal dependence, or autocorrelation (Midtbø 2012). The dependent variables in this analysis civil conflict/war onset are dependent on peace years, i.e. the number of years that has passed since the country last experienced civil conflict/war. Not addressing the established fact of temporal dependence would provide misleading results (Beck, Katz and Tucker 1998). Using the onset of civil conflict/war from the UCDP/PRIO dataset some of the problem is reduces by coding civil war onset instead of incidence (the year of ongoing conflict after year of onset is coded 0 for the dependent variable), this means dropping consecutive years where there are conflict/war from the analysis. But this does not solve the entire problem, peace periods will still be correlated over a period of time.

In research on civil war onset there are two methods that are frequently used in order to handle the problem of temporal dependence. One is to include the lagged dependent variable as in independent variable; the second is to use *natural cubic spline with three knots*[24]. The lagged dependent variable only control for time dependence one year back, while cubic spline will reflect that the effect of number of years since the last occurrence of conflict will decrease rapidly after few year, in a nonlinear function (Beck, Katz and Tucker 1998). Whereas Collier and Hoeffler (2004) and Fearon and Laitin (2003) chose to use a lagged version of the dependent variable in their articles. In this thesis natural cubic splines will be used to handle temporal dependence. This method is used by other scholars (e.g. Chiozza 2002; Walter 2004; Regan and Norton 2005; Buhaug 2006).

It is not only regarding the dependent variable that time must be taken into consideration when using BTSCS data, but also for independent variables that varies over time. Studding onset of civil war and using BTSCS data with yearly observation, it is less likely that it is the observations from the year of onset that best explain the conflict (Midtbø, 2012). For example it is not very logical that it was GDP per capita calculated at the end of the year conflict onset that influenced the onset of conflict. More likely is that GDP per capita the year before that

[24] See Appendix C, C5 for further information

again can help explain the onset of the conflict. In order to correct for this GDP per capita is lagged. In this thesis the time dependent variables used in the logistic analysis have only been lagged one year, but it is possible to lag the variables more than one year. Take for instance the onset of conflict in Mali 1994. By lagging GDP per capita one year it is the GDP per capita in 1993 that is matched with the time unit 1994 in the dataset. Further GDP per capita for 1994 will be matched with the time unit 1995. In order to ensure the robustness of the results presented, they have been clustered by country in the analysis. Cluster analysis is a way of getting more robust results in the form of robust standard errors when doing a logistic analysis with BTSCS data. The analysis is used to identify and group units that are similar to each other. In this case a unit (country-year) is also related to other units belonging to the same country (Hamilton 2013).

3.2 Factor Analysis

The thesis includes two separate factor analyses, they are distinguished by the variables ETHFRACIPOLATED and ETHFRACWAR. Factor analysis as a method is used in the thesis in order to answer the research question *Through what mechanism does ethnic fractionalization work as an indicator for civil conflict onset?* The goal of a factor analysis is to present a larger set of variables in terms of a smaller number of hypothetical variables (Kim and Mueller 1978a). There are two variants of factor analysis, namely confirmatory and exploratory. When using confirmatory factor analysis the researcher has an assumption of how many factors there will be before running the model, the goal here is to run an empirical check on the data that available.

In this thesis exploratory factor analysis will be applied, within this variant there is no assumption on the number of factors that lie underneath and explain the correlation between the different variables (Ulleberg and Nordvik 2001). The selection of variables to be included in the factor analysis is done based on theory, and the results in former research on civil war onset, and they have an adequate number of observations.[25] Although three possible links between ethnicity and civil war onset has been theorized, the theory and the variables give no clear indication as to how many factors they will be sorted into when doing the analysis. Which variables that are placed under which factors are based on the correlation between the

[25] The reason given for including the different variables is given when presenting the different variables, se section 3.6.2.

different variables, and the estimation of scores and loadings are done based on the estimation of covariance (Filzmoser et al. 2009). [26]

After selecting which variables to include in the factor analysis there are three important decisions to be taken when doing an exploratory factor analysis: 1) what factor extraction model to use, 2) the number of factors retained, and 3) what method is used to rotate factors (Conway and Huggcutt 2003: 150). First, there are different types of options when it comes to what type of extraction model to use. Here a principal axis factor analysis is done; this choice has been made based on the aim of the factor analysis, which is to understand the unobserved structure that explains the relationship among measured variables (Conway and Huggcutt 2003: 151).[27] The next decision is to decide how many factors to retain. Experience and research has shown that this is an extremely important step - different techniques leads to different results regarding the number of factors to keep. The most common technique is the eigenvalue specification. The factors that are retained have an eigenvalue over 1. Another way of deciding how many factors to keep is to do a scree-plot. This allows the researcher to examine the eigenvalues in a graph. Based on the graph one should keep all factors until the graph starts to flatten out (Kim and Muller 1978b: 43f). Both ways of retaining the number of factors are more or less the same. In this thesis both of the techniques has been applied. The use of a multiple of techniques gives the most robust selection of variables. Based on the screeplot[28] from the factor analysis in Table 3 the slope decreases after including three factors, but factor 4 have and eigenvalue over 1[29], and is therefore retained in the analysis as well.[30] After deciding how many factors to retain one can start looking at, and interpreting how the different variables load within the different factor.

The third and final stage is to rotate the results from the factor analysis. This is done in order to find a more interpretable result. It can be difficult to interpret the loading of the varibales without rotating the factor analysis. It is an important criterion for interpretation that the factor analysis has a "simple structure", this means that each factor is defined by a subset of

[26] Variables that had uniqueness over 0.7 were excluded from the analysis. Uniqueness give information about how much of the variables variance that is not explained by the factors included in the table. The closer the uniqueness is to 1, the less the variable belongs in the factor analysis (Treiblmaier and Filzmoser 2010).
[27] The alternative is principal factor analysis; this is used when the aim of the study is a pure reduction of the number of variables, without interpreting the results (Conway & Huffcutt 2003: 151).
[28] Se Appendix C, Figure C1.
[29] See Table 3 in section 4.2 for eigenvalues of the four factors retained.
[30] Screeplot and eigenvalues for Table 4 give the same number of factor. See Figure C2 for screeplot, and Table 4 for eigenvalues

measured variables that have high loading on the factor relative to the other variables in the analysis (Fabrigar et al. 1999: 281). The "simple structure" becomes more apparent when the factor analysis is rotated. There are two types of rotation that can be used, orthogonal and oblique rotation. Orthogonal rotation is based on uncorrelated factors, whereas oblique allows factor to be correlated – the last one being most likely in factor analysis done in the thesis. Ford et al. (1986), Fabrigar et al. (1999), and Conway and Huffcutt (2003) argue that oblique rotation gives the best quality on the rotation and the analysis. If factors are more or less uncorrelated the oblique rotation will have a low correlation between the factors instead of no correlation, and is therefore a safe choice in all cases, and is the one applied in this thesis (Conway and Huffcutt 2003: 152f). Within oblique rotation there are again different choices, in this thesis promax oblique rotation has been used, the rationale explanation behind this is that:

> the orthogonal solutions are usually close to the oblique solution, and by reducing the smaller loading to near-zero loading, one can obtain a reasonably good simple structure matrix. Then by finding the best fitting oblique factors for this target matrix, one obtains the desired solution (Kim and Mueller 1978b: 40).

After rotating the factor analysis the different variable loadings must be interpreted. The higher the loading of a variable, the more the variable has to say when identifying the hypothetical variable that is made out of the different factor scores. Naming to the new hypothetical variables is done based on the variables that load on the factor, the variables are already given theoretical meaning when deciding to include then in the analysis. The variables that have the highest loading have the highest influence when naming the factor. Again, this is done based on the theoretical operationalization of the variable.

3.6 Variables

The analysis is based on 175 states over the time period 1985– 2010, that includes a total of 4,471 units (country-years) that are included in the analysis. First, the dependent variable is presented, this is used in the logistic regression analyses that is included in the theses in order to test the new ELF-index, and establish an empirical relationship between conflict and ethnic fractionalization.

3.6.1 The dependent variables

The dependent variables in the logistic regression analysis are civil conflict onset and civil war onset.[31] Both are dichotomous variables that give information on civil conflict/war onset in the observed year, in a given country. They are equal 1 if there was an onset to conflict/war registered and 0 if it was not.[32] The difference between the variables is the threshold of battle related deaths. UCDP/PRIO defines civil conflict onset with a yearly threshold of more than 25 battle deaths. The definition of civil war is a threshold of at least 1000 battle related deaths in a given year. Further a new conflict/war is only registered if there is 2 years or more since the last observed conflict/war. It is worth emphasizing that the variable is only coded 1 in the year of civil conflict/war onset, although the conflict/war might continue several more years the next years are given the value 0. Since the onset variable only registers a conflict/war in the year it erupts it cannot be used to control for autocorrelation. In order to control for autocorrelation it is the incidence variable in the UCDP/PRIO -dataset that must be used. Here there are two options, either lagging the incidence variable one year, or using cubic splines. As already mentioned cubic spline is used in this thesis. The incidence variable gives information on whether or not there was an incidence of conflict/war in the country a given year, and by using this variable we get the information needed in order to control for autocorrelation.

3.6.2 Other variables

When building up a dataset to use in the thesis it was important to include variables that can be seen as indicators of either grievance or opportunity based on theoretical arguments, this is in order to have variables that are clearly associated with grievance or greed when recognizing the spate factors in the factor analysis. In order to get factors that are clearly separated as grievance or opportunity factors variables that are not so commonly used in research on civil war onset, but associated with either grievance of opportunity is also included. It is worth noticing that not all the variables are used in both the factor analysis and the logistic regression. In which analysis the variable is used is specified when presenting the variable.[33] If the variable is included in the regression analysis in additisn to the factor analyses this is specified. All variables except from MOUTIAN is used in the factor analysis.

[31] For descriptive statistic se Table B1 and Table B2 in Appendix B.
[32] Both variables are from UCDP/PRIO Armed Conflict Dataset version 4-2012
[33] See Table B3, B4 and B5 for descriptive statistics for these variables. For further information on the variables besides what is presented here, see the sources (World Bank 2013).

Theoretical grievance variables

POPULATION: is one of the most robust variables in research on civil war onset, although it is rather under-theorized (Dixon 2009; Jakobsen, De Soysa and Jakobsen 2013). Collier and Hoeffler (2004) state that population and grievance are related when they argue that grievance might rise with population. "Population-related pressures can create demands on the government to provide equal access to increasingly scarce resources and may, in fact, result in political violence, when citizens feel their demands are not met" (Abouharb and Kimball 2005: 750). The variable POPULATION gives the number of people living in a country (World Bank 2013a). Due to its robustness, and the common use of it, this variable is included in both the regression analysis and the factor analysis. In both the logistic regression analysis and the factor analysis the variable has been log-transformed, this is done in order to make the variable less squeed further the relative size is given in more meaningful units.[34] Since the variable also is time-dependent it has been lagged one year when used in the logistic regression analysis (Midtbø 2012).

POPULATION GROWTH: is included as a grievance variable based on the same argument as population. If the level of grievance rises with the increase in population it must also raise when the population grows. Population growth is "Population-related pressure", and does therefore result in grievance (Abouharb and Kimball 2005: 750). The variable gives the annual population growth in percentage, based on the total population (World Bank 2013b).

POLITY: Along with population democracy is also one of the variables that are most robust in research. The level of democracy or no democracy is by Collier and Hoeffler (2004) modeled as a factor of grievance. It is seen as a measure of grievance based in the qualities of a democracy compared to countries that are not democracies (Smith 2009). In order to measure the level of democracy the Polity scale is used. Polity is based on a scale from -10 to 10 (Polity IV).[35] In the factor analysis the variable is included as it is originally. In the regression analysis the variable has been recoded into a dummy variable 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.[36]

[34] Earlier research on causes for civil war also uses the log of a given states population (Dixon 2009).
[35] Countries that have -10 are seen as fully autocratic states, and countries with 10 are fully democratic states
[36] 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).

MILITARY SPENDING: Countries that have a high military spending, does not necessarily care much for the welfare of its citizens, and the grievance among people grows. This is also related to grievance after civil war, and the risk of a new civil war. After a war is ended countries keep spending a lot of money on the military instead of building up the country, this leads to grievance among people expecting a better state services after the war is over (Murshed 2002). another reason why military spending is seen as an grievance indicator is that the citizens of a country might see high military spending as an indication of a governments ill-intent (Collier et al 2003: 87). This variable measures military expenditures as a percentage share of countries GDP (World Bank 2013c).[37]

EMPOWERMENT RIGHTS: is an additive scale[38] controlling for the governments respect for different categories.[39] It is a 15 point scale where 0 is given to countries where the government has no respect for the rights, and 14 for countries where the government respect the rights (Cingranelli and Richards 2008). According to Gurr (1983) repression of citizens is a mechanism that is used by the ruling elite to ensure their power, this in turn may cause grievance among those that are not part of the ruling elite. This variable is included in the regression analysis as well as the factor analysis, in the regression analysis it is lagged one year due to its time dependence.

PHYSICAL INTEGRITY RIGHTS: is also an additive index, and it gives the rights not to be extrajudicially killed, tortured or disappeared, or imprisoned for political belief. Having a low score on this 9 point scale indicates that there is an inequality between groups and individuals within a country, which is seen as typical indicators for grievance (Cingranelli and Richards 2008). A low score means that the government do not respect any of the five indicator mentioned. The same argument can be applied here as for EMPOWERMENT RIGHTS; the ruling elite use these methods in order to secure their power (Gurr 1983).

[37] It includes all current and capital expenditures on the armed forces. Excluded are civil defenses and current expenditures for previous military activities (World Bank 2013)

[38] An additative scale is created by adding the individuals score for a number of different human right practices. It is created from disaggregated components, and there is no *a priori* assumption of patterns.

[39] The categories are: freedom of movement, freedom of speech, workers' rights, political participation and freedom of religion.

Theoretical Opportunity Variables

GDP PER CAPITA: is one of the most common variables when analyzing the risk of civil war onset. A robust finding is that GDP PER CAPITA is negatively associated with the risk of civil war onset (Dixon 2009). Collier and Hoeffler (2004), use GDP PER CAPITA in their opportunity model. The variable has by Jakobsen, De Soysa and Jakobsen (2013) been captured within a wealth factor for civil war onset, alongside other variables. Fearon and Laitin (2003: 75) use a low per capita income as a proxy for police and counterinsurgent weakness. GDP PER CAPITA is calculated based on the gross domestic product, and divided by the midyear population (World Bank 2013d). Due to the large difference in values in the variable it has been log transformed. This variable is also included in the logistic regression analysis. Therefore, in addition to being log-transformed it has also been lagged one year in order to control for the time dependence in the regression analysis.

FUEL EXPORT: researchers have found fuel export to be positively and one of the most robust variables associated with the risk of civil war onset (Dixon 2009). FUEL EXPORT represents the much debated link between natural resources and civil war onset. This is a variable that has been used as a proxy for ooportunity. Rebel groups sell contrabands to earn money, which give them opportunity to rebel (Ross 2004; Ross 2006: 265). In this paper the measure for fuel export is coded into a dummy variable based on data from World Bank (2013e). In countries where fuel export exceeds 1/3 of total merchandise export, the variable has been given the value 1, otherwise it is given the value 0. In the regression analysis this variable is lagged one year, due to its time dependence.

BUREAUCRATIC QUALITY: A bureaucracy with high quality is able to police national territory, and provide services that that makes the citizens less rebellious (Fearon and Laitin 2003). Further, countries that have a strong bureaucracy are able to hinder rebellion (DeRouen and Sobek 2004). A weak bureaucracy is associated with a financially weak states, this is due to the costs associated with a well-functioning bureaucracy (Fearon and Laitin 2003). Grounded on this theoretical arguments it is expected that bureaucratic quality is associated with a countries wealth. On the 4 point high scores is given to countries where the bureaucracy has the expertise and strength to keep govern without interruption in the government services and extreme changes in policy, and low scores in countries where this is not the case. In cases of governments change there tends to be a traumatic change in policy constructions and the administrative functions (ICRG 2011).

LAW AND ORDER: is closely associated with the quality of the bureaucracy. In order for the legal system to be functioning well it need legitimacy and the economic resources (DeRouen and Sobek 2004). LAW AND ORDER give information about a countries wealth. The two aspects of law and order are coded separately, and together they make up a 6 point scale. The component of law gives information about the impartiality and strength of the legal system, the order component is a valuation of how the law is observed. A high score on the law and order- scale is a logical indication of a state's capacity to control what happens within its borders (ICRG 2011).

AGRICULURE: In developing countries the primary sector of the economy is dominant. Poor conditions for agriculture hold severe implications for socio-economic development and avoid conflict. Most of the armed conflicts are focused in regions that are heavily dependent on agriculture. Agriculture can be seen as a proxy for capacity, because countries that are depending on agriculture often have a "backward" economy, and that the link that the link that work between agriculture is poverty. (De Soysa and Gleditch 1999). The variable gives agriculture as percentage of GDP; this includes hunting forestry, and fishing, livestock production, and cultivation of crops.[40,] (World Bank 2013f).

LIFE EXPECTANCY: Life expectancy in a country is affected by several factors, infant/birth mortality rates is one of these factors. Countries with a high infant/birth mortality rate have a higher risk of civil war onset, compared to countries that have lower infant/birth mortality rates. This is associated with the development in the country, which is highly associated with the economic capacities of a country (Dixon 2009).[41] Based on this LIFE EXPECTANCY is a variable that say something about a countries wealth. The variable gives life expectancy at birth, this calculation is based on different indicators (World Bank 2013g).

RURAL POPULATION: In the process of recruitment rebel groups finds it helpful to deny government troops access to rural areas. Today rebel groups establish themselves along the rim of the country borders. This is possible in poor countries, or countries with poor bureaucratic quality. People in rural areas might not feel the same type of loyalty to the state,

[40] This variable is not log-transformed because it is given in percentage (Midtbø 2012).
[41] Due to the problem of missing observations the variable measuring mortality rate was not used in the analysis.

which give rebel groups opportunity to recruit (Weinstein 2005). RURAL POPULATION give the percentage of the population that are living in rural areas (World Bank 2013h).

EDUCATIONAL SPENDING: Several studies have included variables that are linked to education in some matter (Dixon 2009). This is an economic variable and is therefore included as a measure for opportunity. Educational spending is calculated based on the public spending on education as a percentage of the total GDP in a given year (World Bank 2013i).

MOUNTAIN: In their research Collier and Hoeffler (2004), and Fearon and Laitin (2003) find that countries that have rough terrain have a higher risk of civil war onset. Based on these results the variable mountain is included in the regression analysis. The data for the variable is from Collier and Hoeffler (2004), and their mountain variable. Over the time period in their analysis the percentage of country territory that was covered of mountains where kept constant. Based on the constant property of mountains and the variable, the values for each country are kept constant over all time-unit periods of the analysis. Terrain is a simple factor for indicating military visibility (Collier et al. 2003: 71). This variable is only included in the regression analysis.

4 Analysis

This chapter presents the results from the empirical analysis. First the results from the regression analysis are put forward, the reason for including this analysis is to test for an empirical relationship between ethnicity, and civil conflict and war when using the updated ELF-index. Further, to test for a curvilinear relationship between ethnicity and civil conflict. The results show that the relationship between ethnicity and civil conflict onset is highly significant, also when controlling for the proposed curvilinear relationship. Second, the results from the factor analysis are presented. Two factor analyses are presented, one including the variable ETHFRACIPOLTAED, the other one includes ETHFRACWAR. Factor analysis is used in this thesis in order to answer the research question: *Through what mechanism does ethnic fractionalization work as an indicator for civil conflict onset?* The results support the assumption arguing that ethnicity is a mechanism for civil conflict onset in itself, due to the salience of ethnic identity.

4.1 Linking Ethnic Fractionalization to Civil Conflict Onset

As pointed out earlier in the thesis, there is no unitary agreement on whether or not ethnic fractionalization leads to civil war. Collier and Hoeffler (2004), and Fearon and Laitin (2003) find no significant link between ethnicity and civil war onset, whereas Sambanis (2001), and Collier, Hoeffler and Rohner (2009) find a significant relationship between the two. Common for all these scholars is that they have used a dependent variable measuring civil war onset with a threshold of 1000 battle related deaths. Scholars that use a threshold of 25 battle related death, or a similar measure with for low intensity conflicts all find a significant relationship between ethnicity, and civil conflict onset (Ellingsen 2000; Regan and Norton 2005; Buhaug 2006; Hegre and Sambanis 2006). In the following regression analysis I have used both dependent variables in order to see if the updated ELF-index gives the same results as other scholars. A significant relationship with a threshold of 25 battle related deaths can be expected based on earlier research, but with a threshold of 1000 battle related deaths it is more difficult to predict an outcome of the analysis due to the different findings done by other scholars.

Table 2. Logistic regression, coefficient (st.error), civil conflict/war onset

Variables	Model 1 Conflict	Model 2 Conflict	Model 4 War	Model 5 War
Population❖¤	0.315***	0.364***	0.123	0.129
	(0.113)	(0.108)	(0.119)	(0.129)
GDP per capita❖¤	-0.286**	-0.370***	-0.496***	-0.511***
	(0.145)	(0.153)	(0.178)	(0.186)
Fuel exporter❖	0.803**	0.823***	0.643	0.634
	(0.320)	(0.153)	(0.448)	(0.450)
Mountain	0.004	0.001	0.010	0.010
	(0.007)	(0.004)	(0.010)	(0.010)
Authocracy❖[a]	-0.341	-0.234	0.068**	-1.054**
	(0.345)	(0.344)	(0.510)	(0.511)
Democracy❖[a]	0.658*	0.601*	-0.193*	-1.248*
	(0.390)	(0.365)	(0.874)	(0.769)
Empowerment rights❖	-0.091**	-0.083**	-0.998	-0.101
	(0.037)	(0.038)	(0.068)	(0.068)
EthfracIpolated❖	1.665***	7.712***	0.101	-0.819
	(0.484)	(2.207)	(0.832)	(3.561)
EthfracIpolated^2 ❖		-6.374***		-0.788
		(2.211)		(3.904)
Temporal dependence				
Peace years	0.093	0.112	0.302*	0.305*
	(0.190)	(0.193)	(0.177)	(0.086)
Spline(1)	0.002	0.002	0.021**	0.021**
	(0.020)	(0.020)	(0.009)	(0.010)
Spline(2)	0.001	0.001	-0.025**	-0.025**
	(0.005)	(0.005)	(0.012)	(0.012)
Spline(3)	-0.001	-0.001	0.019*	0.020*
	(0.002)	(0.002)	(0.011)	(0.011)
Constant	-7.043***	-8.401***	-2.898	-1.933
	(2.117)	(2.121)	(2.846)	(2.925)
Pesudo LL	-321.361	-317.774	-130.516	-130.492
Observations	2734	2734	2754	2754
Years	1985-2010	1985-2010	1985-2010	1985-2010

Note:❖Lagged one year, ¤Log transformed, [a] Anocracy as reference category, [b]Probability of F-value, *significant at 10%, **significant at 5%, and ***significant at 1%. The units are clustered by country using Huber-White robust standard errors (White, 1980).

There are mainly two reasons for including the regression analysis in the thesis, first, in order to test for an empirical relationship between ethnicity, and civil conflict and war, when using the updated ELF-index. The second reason is to test for a curvilinear relationship between ethnicity and civil conflict and war. A significant curvilinear relationship between ethnic fractionalization and civil conflict would demonstrate that it is a higher likelihood for civil

conflict onset if the country is not highly fractionalized or homogenous. This is the same argument as presented by Ellingsen (2000) and Montavlo and Reynal-Querol (2005). The results from the regression analysis are presented in Table 2. Following, the control variables is commented on shortly before looking closer at relationship between ethnicity and civil war onset. The control variables that are included in the analysis are variables that are commonly used in research on civil war onset.

In Model 1 and Model 2 the dependent variable is civil conflict onset. The variables POPULATION, GDP PER CAPITA, FUEL EXPORT, and DEMOCRACY are variables that are robustly associated with civil conflict onset in earlier research. (Dixon 2009). The analysis shows that all these variables are significant when explaining civil conflict onset, but note that the direction of their influence varies. Democracy is part of a dummy set, and the results show that it is only democracy that is significant or civil conflict onset not autocracy. The variable PEACE YEAR with splines is as mentioned used to control for temporal dependence.[42] The results show that the number of years since the last conflict is not significant for the likelihood of civil conflict onset. Whether or not a country had civil war in previous years is normally a variable that is shown robust in earlier research, but here, with civil conflict onset as dependent variable, it is not (Buhaug and Sambanis 2006; Dixon 2009). The variable MOUTIAN is also not significant for civil conflict onset. This is contrary to Fearon and Laitin (2003) who found that rough terrain favors rebellion, but in accordance with Collier and Hoeffler (2004), who found no significant relationship. The last variable EMPOWERMENT RIGHTS is not commonly used in research on civil war onset. It is included in order to have a clear grievance measure in the analysis. The results show that an increase on the scale measuring empowerment rights has a negative and significant effect on the risk for civil conflict onset.

Having civil war as the dependent variable in Model 3 and Model 4, several of the control variables has lost its significance. Only GDP PER CAPITA and DEMOCRACY are still significant. ANOCRACY and the variables controlling for autocorrelation, PEACE YEARS whit splines are significant for civil war onset, and not civil conflict. All together this show that almost all the variables included in Table 2 is significant for either civil conflict onset, or civil war onset.

[42] As a test of roubsutness it has also been tested for temporal dependence by lagging the incidence variable, see Table A2 in Appendix A. This is the same as lagging the dependent variable, but due to the coding of the variables incidence must be used when UCDP/PRIO's dataset is used.

Further, the results of the regression analysis in Model 1, where civil conflict is the dependent variable, show that there is a significant relationship between the updated variable measuring ethnic fractionalization, ETHFRACIPOLATED, and the risk of civil conflict onset. This is the same result as found in much of the literature (e.g. Ellingsen 2000; Regan and Norton 2005; Buhaug 2006; Hegre and Sambanis 2006). Based on this result an empirical relationship between ethnicity and civil conflict onset is established. This creates the foundation to further investigate ethnicity as an indicator of regarding civil conflict onset. In Model 2 I have tested for a curvilinear relationship between ethnic fractionalization and civil conflict onset. The results show that there is a significant curvilinear relationship between ethnic fractionalization and civil conflict onset, taking on the shape of an inverted-u. By this it follows that the likelihood of civil conflict onset is lower in countries that are homogenous, but increases as the level of fractionalization rise. Moving past the vertex, closer to a heterogeneous country, the likelihood of civil conflict onset decreases.[43] Documenting a curvilinear relationship between ethnic fractionalization and civil conflict onset gives support to Montalvo and Reynal-Querol (2005) and Ellingsen (2000), who argue that the size of an ethnic group matters for the likelihood of civil conflict onset. Further, the result also add empirical support to the argument that ethnicity is an indicator of opportunity.

Looking at the results in Model 3 and Model 4, where civil war is the dependent variable, neither ETHFRACIPOLATED, nor ETHFRACIPOLATED^2 is significant for the likelihood of civil war onset. This gives support to Coller and Hoeffler (2004), and Fearon and Laitin (2003) who found no significant effect of ethnic fractionalization of the likelihood of civil war onset. Since an empirical relationship between ethnic fractionalization, and civil conflict onset is established based on the significance between the two variables, it is reason to look at through what mechanisms ethnic fractionalization does work as an indicator for civil conflict onset.[44] Below the research question is addressed directly, and it will be answered whether or not Collier and Hoeffler (2004), and Fearon and Laitin (2003) are right in operationalizing ethnicity as an indicator for grievance. Or if it is an indicator of opportunity, and even worse

[43] This relationship is taken into account in the factor analysis by coding the variable ETHFRACWAR. This variable is coded based on the binary relationship between ethnic fractionalization and civil conflict onset, and presented earlier in the thesis.

[44] The different results regarding ethnic fractionalization using respectively civil conflict, and civil war as dependent variables indicate that there are different mechanisms that explain onset of civil war and conflict. Regan and Norton (2005) present an argument, explaining some of these differences.

for Collier and Hoeffler (2004) an indicator of identity – a mechanism that is not included in their model for civil war onset.

4.2 Through what mechanism does ethnicity work as in indicator?

The theoretical framework presented earlier in the thesis created the foundation for three separate arguments. I argued that ethnicity is an indicator if grievance, as done by Collier and Hoeffler (2004) and Fearon and Laitin (2003). Second it was argued that ethnic fractionalization is an indicator of opportunity, and last, that ethnicity is an indicator in itself, due to the salience of ethnic identity. 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, using an oblique promax rotation.[45] Two separate factor analyses was presented, the difference between the two analyses is the variable measuring ethnic fractionalization. Table 3 includes the variable ETHFRACIPOLATED, and Table 4 includes ETHFRACWAR, other than this the two factor analyses are identical. Since the result of the two factor analyses are very similar they are commented on correspondingly, the few differences between them are commented on consecutively. The two factor analyses identifies four different factors, these are given names based on theoretical belonging of the variables that loads on the different factors.[46] After identifying the factor as either a grievance or opportunity factor, relevant theory and belonging hypothesis are discussed, before moving on to the next factor. The result demonstrates that ethnic fractionalization is not an indicator of either grievance or opportunity, but an indicator in itsef.

The variables that load on the first factor in the both Table 3, and Table 4 are variables that theoretically belong within the opportunity explanation for civil war onset. GDP PER CAPITA and AGRICULTURE have the highest loadings on the factor. These variables do theoretically say something about a countries wealth, and the factor is therefore called *Wealth*.[47] Comparing Table 3 and Table 4 EDUCATION SPENDING do no longer load on the *Wealth* factor in Table 4. The result of this has little to say for the theoretical identification of the

[45] Factor scores under 0.35 have been omitted from the tables; these scores are so low that they are not relevant for the understanding of the hypothetical variables made out of the loading of the different variables on the factor.
[46] All together the four factors explain 0.733 of the variance to the variables in Table 3, and 0.709 in Table 4.
[47] BUREAUCRATIQ QUALITY, is by Jakobsen, De Soysa, and Jakobsen (2013) found to be an indicator for *Capacity*. This factor is not found in this thesis.

factor, because it loads weakly compared to other clear opportunity variables. The same goes for POPULATION GROWTH that loads on the fourth factor in Table 3, but on the *Wealth* factor in Table 4.

Table 3. Eigenvalue, rotated,variable loadings, unique variance, ETHFRACIPOLATED, 1985– 2010

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

Arguing that ethnic fractionalization is an indicator of opportunity for civil conflict onset, ethnic groups are seen extreme forms of interest groups that have a common preference in politics (Fearon 2006). Ethnic groups must be seen as rational actors that calculate and maximize the utility before taking part in or starting a civil conflict (Toft 2002). Following a rational actor perspective, ethnic collective violence is based on a belief that each individual within the ethnic group have given wants, goals, or tastes, and individuals will choose the most efficient way to reach their goals (Hechter, Friedman, and Appelbaum 1982). Further, ethnicity is an indicator of opportunity since the group must be of a certain size in order to rebel. This argument was based on findings presented by Montalvo and Reynal-Querol (2005), and Ellingsen (2000). The regression analysis in Table 2 supports their findings due to

the significant effect of the curvilinear relationship between ethnic fractionalization and civil conflict onset.

As both Table 3 and Table 4 demonstrate ethnicity does not load in the *Wealth* factor. As a result of this hypothesis H_2: *Ethnicity is a mechanism of opportunity*, must be rejected. ETHFRACWAR, the variable coded based on the curvilinear relationship between ethnic fractionalization and civil conflict onset strengthens the rejection of hypothesis H_2 since some level of polarization is taken into account. Ethnic fractionalization is *not* a mechanism of opportunity and rational actor theory.

Table 4. Eigenvalue, rotated, variable loading, and uniwue variance, ETHFRACWAR, 1985– 2010

Variables	***Wealth***	***Empowerment***	***Grievance***	***Identity***	**Uniqueness**
GDP per capita¤	0.927				0.096
Agriculture	-0.950				0.225
Bureaucratic quality	0.737				0.328
Life expectancy	0.9334				0.203
Rural population	-0.896				0.280
Law and order	0.617				0.401
Educational spending			0.493		0.610
Military		-0.828			0.360
Fuel export		-0.811			0.391
Polity		0.807			0.205
Empowerment rights		0.755			0.207
Population¤			-0.948		0.244
Physical integrity rights			0.663		0.232
Ethfracwar				0.985	0.027
Population growth	-0.517				0.558
Eigenvalue	6.230	2.347	1.343	1.071	**N=1185**

The second and third factors are both grievance factors, named respectively *Empowerment rights*, and *Grievance.* Factor number two is a grievance factor based on the theoretical argument that belonging to the variables MILITARY, POLITY, and EMPOWERMENT RIGHTS. The variable FUEL EXPORT was theoretically argued to be an indicator for opportunity, but

since all the other variables that load on the factor have a strong belonging in the grievance theory the factor is identified as a grievance factor. FUEL EXPORT can also be seen as a grievance indicator based on the theoretical argument. States will be unwilling to give up, or have an interest in stronger control over areas where there are valuable natural resources. Due to the states interest, and their willingness to get what they are interested in, a feeling of grievance may develop among those that are affected by the interests of the states. Finding natural resources such as oil, gas, diamonds, and gold can lead to a rearrangement or compulsory moving for those living in the area (Toft 2002). Based on this the level of fuel export might result in grievances among those that are affected. Further, countries might not have good governance, and the economic advantages might only benefit a share of the population. This in turn might lead to a relative deprivation or inequality (Davies 1962; Gurr 1970, Cederman, Wimmer, and Min 2009, 2010). The third factor *Grievance* has high loadings of variables that clearly are theoretical grievance variables. In Table 3 the factor is made up of POPULATION, and PHYSICAL INTEGRITY RIGHTS, in Table 4 EDUCATION SPENDING also loads on this factor. There are no changes in which factors that have the highest loading, and therefore the factor has the same name in both of the tables.

Explaining civil war as a result of grievance has its roots in the relative deprivation theory. The deprived actor school assumes that anger is a result of deprivation, and that it has a direct effect on the likelihood of a person's participation in violent action (Davies 1962; Gurr 1970). Frustration explanations for civil war onset is linked to ethnicity by arguing that grievance is based in ethnic identity, and that a common identity in itself is not sufficient in order to mobilize for civil war onset, a common grievance must also be present (Toft 2002; Murshed and Gates 2005). In light of the relative deprivation school the gap between what an individual expects, and what one gets results in grievances. If this is experienced along ethnic lines an ethnic group can gather, and fight collectively for what each of the individuals feel they are entitled to. Inequality between ethnic groups is also seen as something that can result in grievance within an ethnic group. Ethnic groups that are discriminated have a higher likelihood of turning violent than groups that are not politically discriminated (Cederman, Wimman and Min 2009, 2010). According to Sen (1973) there is a close link between inequality and rebellion. In their model for civil war Collier and Hoeffler (2004) used ethnic fractionalization as a mechanism of grievance, when they included it in their grievance model for civil war onset. The results from the factor analysis show that Collier and Hoeffler (2004) where wrong when assuming that ethnicity is an indicator for grievance. Ethnicity does not

load on either of the grievance factors, *Empowerment* or *Grievance.* As a result of this it follows that hypothesis H_1: *Ethnicity is a mechanism of grievance*, is rejected. Linking together ethnicity and grievance is not the right way to go about when connecting ethnic fractionalization as and civil war onset.

It is the fourth factor that gives an answer to what ethnic fractionalization is an indicator of. In Table 3 ETHFRACIPOLATED loads on factor four together with the variables EDUACATION SPENDING, and POPULATION GROWTH. It is ETHFRACIPOLATED that has the highest loading on the factor, and based on the theoretical explanation of ethnicity as an indicator of identity this factor is named *Identity.* In Table 4 it is only one variable that load on factor four, this is ETHFRACWAR which emphasize that the factor should be named *Identity*.

Ethnic fractionalization is argued to be an indicator in itself for civil conflict onset based on the importance of ethnic identity. According to Wolff (2006) ethnic identity has an enormous power when it becomes the predominant identity in a person's personality. As mentioned earlier, it is the fundamental factor that ethnicity rest upon that makes it so important for a person, these are factors such as language, religion, heritage and territory (Ellingsen 2000; Toft 2002). Based on Toft (2002), ethnic identity is an indicator in itself, this is because much of ethnic group identity rests upon the territory where they live. Territory is important for the survival of the group. If an ethnic group feel that their territory is threatened the salience of their shared identity becomes apparent, and the group must mobilize in order to protect their territory and secure their survival and identity (Toft 2002). According to primordialists, ethnic identity is enough to mobilize for civil conflict onset in itself, based on the fundamental characteristics of ethnicity, "ancient hatred" and differences between groups. They argue that there are no other mechanisms needed for civil war onset than common ethnic identity. Only identity can explain civil war onset, not for instance political and economic differences (Blimes 2006; Østby 2011). The result from the factor analysis in both Table 3, and Table 4 give support to hypothesis H_3: *Ethnicity is an indicator for civil conflict in itself.*

By confirming hypothesis H_3, that ethnicity is an indicator in itself the research question has been answered, ethnicity work as an indicator through the mechanism of identity in civil conflict onset. The results from the factor analysis show that Collier and Hoeffler (2004), and Fearon and Laitin (2003) are incorrect when they operationalize ethnic fractionalization as an indicator for grievance, nor is it an indicator of opportunity as indicated by Ellingsen (2011).

That ethnicity is not an indicator of grievance or opportunity has consequences for the conceptual understanding of civil conflict onset as a result of grievance or opportunity. Collier and Hoeffler (2004), and to some extent Fearon and Laitin (2003) conceptualize civil war as a result of grievance or opportunity, it follows based the results from the factor analysis that identity also must be built-in when modeling civil conflict onset.[48] The result of this is that a new model for civil war onset must be put forward, civil conflict as a result of grievance, opportunity or identity, and not only grievance or opportunity as done by Collier and Hoeffler (2004), and Fearon and Laitin (2003).

But is it as simple as to explain civil war as a result of grievance, opportunity, or identity? According to primordialists it is, but arguing that ethnic identity is a mechanism for civil war onset in itself has been critiqued by those who do not believe in the primordial explanation of the link between ethnicity and civil conflict onset. Primordialists fail to account for the differences in the level of conflict over time, and also why there are multiethnic societies that live in peace (Østby 2011). Following this criticism of primordialism, identity as a mechanism for civil war onset is not enough in itself to start a conflict.

Despite arguing strongly for the link between of territory and ethnic identity, Toft (2002) also built upon grievance and opportunity. She argued that ethnic groups are rational actors that do a cost-benefit analysis before entering a civil war. This links together identity and opportunity as mechanisms for civil war onset. Further, if an ethnic groups territory is threatened, so are the chances for survival of an ethnic group, this leads to frustration and grievance. According to Murshed and Gates (2005: 122) it is not enough to have a common identity, there must also be a shared feeling of grievance. If a person identifies with a group that is victim of discrimination, the more likely it is that the individual will participate in violent actions in order to defend their ethnic identity. In this case a shared ethnic identity and grievance function as a mobilization factor (Harff and Gurr 1993). Based on this argument there are no clear distinction between whether it is identity or grievance or opportunity that causes civil war. The likelihood of civil conflict onset increases if there are a combination of mechanisms involved. In other words, grievance and opportunity can still be mechanisms for civil war

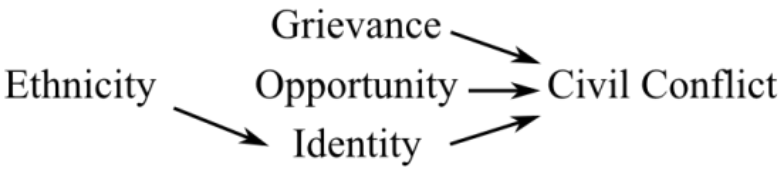
[48] As demonstrated in the regression analysis it is a significant relationship between ETHFRACIPOLATED and civil conflict onset. With civil war onset as the dependent variable no significant relationship to ETHFRACIPOLATED was found. As mentioned earlier this is in accordance to the findings done by Collier and Hoeffler (2004), and Fearon and Laitin (2003). But reviewing the literature on civil war onset other scholars have found a significant relationship between ethnic fractionalization and civil war (Sambanis 2001; Regan and Norton 2005), based on this the following argument can be applicable to civil war onset, as well as civil conflict.

onset, but ethnicity is not an indicator for either of them and must be conceptualized as a mechanism in itself.

If civil war is operationalized in a perspective of either grievance or opportunity an important mechanism for civil war onset is left out, the mechanism of identity. This is shown in two factor analyses where ethnic fractionalization is identified to be an indicator in itself, due to the salience off identity for civil conflict onset. But identity is not a mechanism that in itself can explain civil war onset. The arguments above show that it is problematic to frame civil conflict onset in "either-or" terms. This is supported by several scholars (Ellingsen 2000, 2011; Gurr 2000; Berdal 2005; Østby 2008).

In other words, recognizing the importance of identity when explaining civil war onset, cannot be the same as ruling out the importance of grievance and opportunity, these are still important mechanisms in order to clarify onset of civil conflict. Neither is the conceptual distinction for understanding civil conflict as a mechanism of grievance or opportunity helpful. Civil conflict cannot be understood through either opportunity or grievance as argued, and operationalized by Collier and Hoeffler (2004), and in part by Fearon and Laitin (2003). In order to explain civil war the mechanism of identity must also be incorporated into the model, this is demonstrated through the results of the factor analyses where ethnicity is not an indicator of grievance as it operationalized by Collier and Hoeffler (2004) and Fearon and Laitin (2003), and neither an indicator of opportunity as indicated by Elingsen (2011). Ethnicity is an indicator in itself due to the salience of identity. The result are presented in Figure 3 where it is shown that ethnicity is an indicator of identity, and the arrows pointing from the different mechanisms demonstrate that civil war is a result of all three mechanisms, grievance, opportunity, and identity.

Figure 4. Overview results and argument



Widening the conceptual understanding of civil war is confirmatory with Gurr (2000), and Ellingsen (2000) argument. Both argue that civil war onset cannot be seen as a result of identity, grievance or opportunity, rather it should be seen as a result of identity, grievance, and opportunity. According to Gurr (2000: 66) "It is simplistic to argue that one kind of motivation is primary and the others subsidiary".

All in all, based on the results from the factor analysis, it is concluded that ethnic identity is an indicator for civil conflict onset in itself due to the salience of identity. But legitimate critique of the primordial perspective on ethnicity and civil conflict, alongside an argument that civil war is a result of several mechanisms show that identity is not a sufficient mechanism in itself to explain civil war, grievance and opportunity must also be included in an explanation. With ethnicity as an indicator for civil war onset in itself, further research cannot be conceptualized as a result of grievance or opportunity, as done by Collier and Hoeffler (2004). Identity must be included in a causal model for civil conflict onset, and civil conflict must be seen as a result of grievance, opportunity *and* identity, not grievance, opportunity *or* identity (Ellingsen 2000, 2011; Gurr 2000).

5 Conclusion

In this thesis I have answered the research question, *through what mechanism does ethnic fractionalization work as an indicator for civil conflict onset?* In earlier literature there has been a tendency to operationalize ethnic fractionalization as a measure for grievance when explaining civil conflict onset, often in a perspective of conceptualizing civil conflict as a consequence of either grievance or opportunity (Fearon and Laitin 2003; Collier and Hoeffler 2004; Ellingsen 2011). The main contribution of this thesis to the research field of civil war onset is to answer whether ethnicity can be operationalized as an indicator of grievance, or if it is an indicator of the mechanism opportunity, or a mechanism in itself due to the salience of ethnic identity.

Based on the results from the analysis I have demonstrated that ethnic fractionalization is an indicator of civil conflict onset in itself, due to the salience of ethnic identity. This means that Collier and Hoeffler (2004) and Fearon and Laitin (2003) are incorrect when using ethnic fractionalization as an indicator of grievance. As a consequence of the results it follows that further research must include identity in a causal model when conceptualizing civil conflict onset. Civil conflict cannot be conceptualized as a result of either grievance or opportunity as done by Collier and Hoeffler (2004). Following this, I argue that ethnic identity as a mechanism for civil war in itself due to the salience of identity, but that it is not adequate to start a civil war. This is mainly due to the primordialists' failure to explain why not all ethnic groups fight each other, if the only mechanism needed for civil conflict inset is a shared ethnic identity. By this it follows that the mechanisms for civil conflict onset cannot be seen in an "either-or" perspective. Civil conflict must be seen as a result of grievance, opportunity and identity.

Answering the research question is one of two contributions this thesis provides to the research of civil war onset. The second contribution is an updated version of the ELF-index. The ELF-index gives the level of ethnic fractionalization within a country by calculating the likelihood of two randomly drawn individuals in a country coming from different ethnic groups. By interpolating the old and new value in the ELF-index, the changing composition of ethnic fractionalization is reflected, which gave the variable ETHFRACIPOLATED. The second variable was also coded based on the relationship between ETHFRACIPOLATED, and civil conflict onset. Based on the curvilinear relationship between the two, the variable

ETHFRACWAR was coded so that it was pinpointed towards conflict. In order to test the new ELF-index and the empirical relationship between ethnic fractionalization and civil conflict, a logistic regression analysis was performed, and presented in the chapter of empirical analyses. Using time-series cross-sectional data for the period 1985–2010, results showed a significant curvilinear relationship between ethnic fractionalization and civil conflict onset, but not for civil war as dependent variable.

Following this the hypotheses were tested. In order to do this, two separate exploratory factor analyses were performed, the difference between them being the coding of the variable measuring ethnic fractionalization. One included ETHFRACIPOLATED, the other one ETHFRACWAR, both gave the same results. In order to answer the research question I theorized three possible links between ethnicity and civil conflict onset earlier in the thesis. This resulted in three mutually exclusive hypotheses. The first hypothesis H_1: *Ethnicity is an indicator of grievance,* was theorized based on the relationship between ethnicity, relative deprivation and inequality. The second hypothesis H_2: *Ethnicity is an indicator of opportunity,* was mainly based on the connection between ethnicity, and the rational actor theory. Last, the third hypothesis H_3: *Ethnicity is an indicator for civil conflict in itself,* here I argued based on primordialists view, and the important link between territory and ethnic survival and that identity in itself is enough for civil conflict onset.

Based on the results from the factor analyses, hypothesis H_3 was confirmed. Ethnicity is an indicator in itself, due to the salience of identity. This means that ethnic fractionalization cannot be used as an indicator for grievance as done by Collier and Hoeffler (2004), nor is at an opportunity indicator as indicated by Ellingsen (2011). By concluding that ethnic fractionalization is an indicator for civil conflict onset in itself, it follows that future research must take into account the aspect of ethnicity when conceptualizing civil conflict onset. This results in another implication for future research. Since identity in itself cannot explain civil conflict onset, due to the shortcomings of the primordialists' connection between ethnicity and civil conflict, civil conflict cannot be seen as a result of either grievance, frustration or identity, it must be seen as a result of all three. Civil conflict must be regarded as a result of grievance, opportunity and identity, in other words "It is simplistic to argue that one kind of motivation is primary and the others subsidiary" Gurr (2000: 66).

6 References

Abouharb, M. Rodwan and Anessa L. Kimball (2007). "A New Dataset on Infant Mortality Rates, 1816–2002". *Journal of Peace Research*, 44(6): 743–754.

Alanso, Ana Maria (1994). "The Politics of Space, Time, and Substance: State Formation, Nationalism, and Ethnicity". *Annu. Rev. Anthropol*, 23: 379–405.

Alesina, Alberto, Arnaud Devleeschauwer, Willian Easterly, Sergio Kurlat, Romain Wacziarg (2003). "Fractionalization". *Journal of Economic Growth*, 8(2): 155–194.

Allen, Karen (2012, November). *Samburu Country Killings: Kenya's 'valley of death'*. Downloaded 25.May 2013 from http://www.bbc.co.uk/news/world-africa-20392510.

Atlas Narodov Mira (1964). Moscow: Glavnoe upravlenie geodezii i kartografii.

Beck, Nathaniel, Johnathan N. Katz and Richard Tucker (1998). "Taking Time Seriously: Time-Series Cross-Section Analysis with a Binary Dependent Variable". *American Journal of Political Science*, 42(4): 1260–1288.

Beck, Nathaniel (2001). "Time-Series Cross-Section Data". *Statistica Neerlandica*, 55(2): 111–133.

Berdal, Mats (2005). "Beyond greed and grievance – and not to soon …". *Review of International Studies*, (4): 687–698.

Blimes, Randall L. (2006). "The Indirect Effect of Ethnic Heterogenity on the Likelihood of Civil War Onset". *Journal of Conflict Resolution*, 50(4): 536–547.

Buhaug, Halvard (2006). "Relative Capacity and Rebel Objective in Civil War". *Journal of Peace Research*, 43(6): 691–708.

Cederman, Lars-Erik and Luc Girardin (2007). "Beyond Fractionalization: Mapping Ethnicity onto Nationalist Insurgencies". *American Political Science Review*, 101(1): 173–185.

Cederman, Lars-Erik, Anderas Wimmer and Brian Min (2009). "Ethnic politics and armed conflict. Aconfigurational analysis of a new global dataset". *American Sociological Review*, 72(2): 316–337.

Cederman, Lars-Erik, Anderas Wimmer and Brian Min (2010). "Why do Ethnic Groups Rebel? New Data and Anlysis". *World Politics*, 62(1): 87– 119.

Chesterman, Simon (2001). "Introduction: Global Norms, Local Contex" in Simon Chesterman (ed.), *Civilians in War*. Boulder CO: Lynne Rienner.

Chirozza, Giacomo (2002). "Is there a Clash of Civilizations? Evidence from Patterns of International Conflict Involvement, 1946–97". *Journal of Peace Reseach*, 39(6): 711–734.

CIA World Factbook (2012). "Country specific data" Last consulted January 2013 https://www.cia.gov/library/publications/the-world-factbook/.

Cingranelli, David L. and David L. Richards (2008). *The Cingranelli-Richards Human Rights Dataset Version 2008.03.12*. Downloaded 28. January 2013 from http://www.humanrightsdata.org.

Collier, Paul (1999). "On Economic Consequences of Civil War". *Oxford Economic Papers*, 51(1): 168–193.

Collier, Paul (2001). "Implications of Ethnic Diversity". *Economic Policy*, 16(32): 127– 166.

Collier, Paul, Lani Elliot, Håvard Hegre, Anke Hoeffler, Marta Reynal-Querol, and Nicolas Sambanis (2003). *Breasking the Conflict Trap: Civil War and Development Policy.* Washington, DC: World Bank and Oxford University Press.

Collier, Paul and Anke Hoeffler (2004). "Greed and Grievance in Civil War". *Oxford Economic Papers*, 56(4): 563– 595.

Collier, Paul, Anke Hoeffler and Nicholas Sambanis (2005). "The Collier-Hoeffler Model of Civil War Onset and the Case Study Project Research Design" in Paul Collier, Nicholas Sambanis (ed.), *Understanding Civil War: Evidence and Analysis – Europe, Central Asia, and Other Regions.* Washington, D.C.: World Bank.

Collier, Paul (2007). "Economic Causes of Civil Conflict and Their Implications for Policy" in Chester A. Crooker, Fen Osler Hampson and Pamela All (ed.), *Leasing the Dogs of War. Conflict Management in a Divided World*. Washington, DC: United States Institute of Peace Press.

Collier, Paul, Anke Hoeffler and Dominic Rohner (2009). "Beyond Greed and Grievance: Feasibility and Civil War". *Oxford Economic Papers*, 61(1): 1–27.

Conteh-Morgan, E. (2004). *Collective political violence: an introduction to the theories and cases of violent conflicts*. London: Routledge.

Davies, James C. (1962). "Towards a Theory of Revolution". *American Sociological Review*, 27(1): 5–19.

Conway, James M. and Allen I. Huffcutt (2003). "A Review and Evaluation of Exploratory Factor Analysis Practices in Organizational Reseach". *Organizational Research Methods*, 6(2): 147–168.

Correaltes of War 2 Project (2013). Last consulted April 2013 http://www.correlatesofwar.org/.

DeRouen, Karl R. and David Sobek (2004). "The Dynamics of Civil War Duration and Outcome". *Journal of Peace Research*, 41(3): 303–320.

Dixon, Jeffey (2009). "What Causes Civil Wars? Integrating Quantitative Research Findings". *International Studies Review*, 11(4): 707–735.

Ellingsen, Tanja (2000). "Colorful Community or Ethnic Witches' Brew? Multiethnicity and Domestic Conflict during and after the Cold War". *Journal of Conflict Resolution*, 44(2): 228–249.

Ellingsen, Tanja (2011). "Ethnicity Matters, but What Kind? A Review of the Quantitative Literature of Ethnicity and Conflict" in Tor Georg Jakobsen (ed.), *War: An Introduction to Theories and Research on Collective Violence.* New York: Nova Science Publisher.

Esteban, Joan-Maria and Debraj Ray (1994). "On the measurement of polarization". *Econometrica*, 62(4): 819– 851.

Fabrigar, Leandre R., Duane T. Wegener, Robert C. MacCallum, Erin J. Strahan (1999). "Evaluating the Use of Exploratory Factor Analysis in Psychological Research". *Psychological Methods*, 4(3): 272– 299.

Fearon James, D., (2003). "Ethnic and Cultural Diversity by Country ". *Journal of Economic Growth*, 8 (2): 195–222.

Fearon, James D. and David D. Laitin (2003). "Ethnicity, Insurgency, and Civil War". *American Political Science Review*, 97(1): 75–90.

Fearon, James D. (2006). "Ethnic Mobilization and Ethnic Violence" in Barry R. Weingast and Donald A. Wittman (ed.), *The Oxford Handbook of Political Economy*. New York: Oxford University Press.

Fearon, James D., Kimuli Kasara and David D. Laitin (2007). "Ethnic Minority Rule and Civil War Onset". *American Political Science Review*, 101(1): 187–193.

Fenton, Steve (2010). *Ethnicity* (2nd ed.). Cambridge: Polity.

Filzmoser, Peter, Karel Hron, Clemens Reimann and Robert Garret (2009). "Robust factor analysis for compositional data". *Computers and Geosciences*, 35: 1854–1861.

Floyd, Frank J. and Keith F Widaman (1995). "Factor Analysis in the Development and Refinement of Clinical Assessment Instrument". *Psychological Assessment*, 7(3): 286–299.

Ford, J. Kevin, Robert C. MacCallum, and Marianne Tait (1986). "The Application of Exploratory Factor Analysis in Applied Psychology: A Critical Review and Analysis". *Personnel Psychology*, 39(2): 291–314.

Geertz, Clifford (1973). *The Interpretation of Cultures*. New York: Basic Books.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson,Margareta Sollenberg, and Håvard Strand (2002). "Armed Conflict 1946–2001: A New Dataset". *Journal of Peace Research*, 39(5).

Gurr, Ted Robert (1970). *Why Men Rebel*. Princeton, N.J.: Princeton University Press.

Gurr, Ted Robert (1983). "The Political Origins of State Violence and Terror: A Theoretical Analysis" in Michael Stohl and George A. Lopez (ed.), *Government Violence and Repression: An Agenda for Research*. New York: Greenwood Press.

Gurr, Ted R. (2000). *People versus State: Minorities at Risk in the New Century*. Washington, D. C: United States Institute of Peace Press.

Hamilton, Lawrence C. (1992). *Regression with Graphics. A Second Course in Applied Statisitcs*'. California: Duxbury Press.

Hamilton, Lawrence C. (2013). *Statisitcs with STATA. Version 12*. USA: Books/Cole Cengage Learning

Harbom, Lotta and Peter Wallensteen (2012). "Armed Conflict, 1989–2011". *Journal of Peace Research*, 49(4).

Harff, Barbara and Ted Robert Gurr (1993). *Ethnic Conflict in World Politics*. New York: Westview Press.

Hay, Colin (2002). *Political Analysis: A Critical Introduction*. Basingstoke: Palgrave.

Hechter, Michael, Debra Friedman and Malka Appelbaum (1982). "A Theory of Ethnic Collective Action". *International Migration Review*, 16(2): 412– 434.

Hegre, Håvard and Nicholas Sambanis (2006). "Sensitivity Analysis of the Empirical Literature on Civil War Onset". *Journal of Conflict Resolution*, 50(4): 508– 535.

Hegre, Håvard, Tanja Ellingsen, Scott Gates, and Nils Petter Gleditsch (2001). "Toward a democratic civil peace? Democracy, political change, and civil war, 1816–1992". *American Political Science Review*, 95(1): 33–48.

Horowitz, Donald L. (2000). *Ethnic Groups in Conflict*. Berkeley: University of California Press.

Hox, Joop, J. (2010). *Multilevel Analysis: Techniques and Applications* (2nd ed.). New York: Routledge.

ICRG (2011). *International Country Risk Guide*, http://www.prsgroup.com.

Jacoby, Tim (2008). *Understanding Conflict and Violence. Theoretical and interdisciplinary approaches.* London: Routledge.

Jakobsen, Tor Georg, Indra De Soysa and Jo Jakobsen (2013). "Why do Poor Countries Suffer Costly Conflict? Unpacking Per Capita Income and the Onset of Civil War". *Conflict Management and Peace Science*, 30(2): 140–160.

Joshua Project (2012). "Country specific data". Last consulted January 2013 http://joshuaproject.net/.

Kanbur, Ravi, Prem K. Rajaram, and Ashutosh Varshney (2011). "Ethnic Diversity and Ethnic Strife. An Interdisciplinary Perspective". *World Development*, 39(2): 147– 158.

Kalyvas, Stathis N. (2008). "Ethnic Defection in Civil War". *Comparative Political Studies*, 41(8): 1043–68.

Kelly, Willian A. (1981). "A Generalized Interpretation of the Herfindahl Index". *Southern Economic Journal*, 48(1): 50–57.

Kim, Jae- On and Charles W. Mueller (1978a). *Introduction to Factor Analysis. How It Is and How To Do It*. London: Sage Publications.

Kim, Jae- On and Charles W. Mueller (1978b). *Factor Analysis. Statistical Methos and Practical Issues.* London: Sage Publications.

Kladermans, Bert (1984). "Mobilization and Participation; Social-psychological expansions of resource mobilization theory". *American Sociological Review*, 49(5): 583–600.

Laitin, David and Daniel Posner (2001). "The Implications of Constructivism for Constructing Ethnic Fractionalization Indices". *APSA-CP*, 12(1): 13–17.

Lake, David A. and Donald Rothchild (1998). "Spreading fear: The genesis of transnational ethnic conflict" in David A. Lake and Donald Rothchild (ed.), *The international spread of ethnic conflict: Fear diffusion and escalation*. Princeton, NJ: Princeton University Press.

Levinson, David (1998). *Ethnic Groups Worldwide: A Eeasy Reference Handbook*. Phoenix: Oryx Press.

Library of Congress (2012). *Country Studies*. Last Consulted August 2012 http://lcweb2.loc.gov/frd/cs/.

Midtbø, Tor (2012). *STATA. En entusiasitsk innføring.* Oslo: Universitetsforlaget.

Montalvo, Jose G., and Marta Reynal-Querol (2005). "Ethnic Polarization, Potential Conflict, and Civil Wars". *American Economic Review*, 95(3): 796–816.

Moses, Jonathon M. and Torbjørn L. Knutsen (2007). *Ways of Knowing*. London: Palgrave.

Murshed, S. Mansoob (2002). "Conflict, Civil War and Underdevelopment: An Introduction". *Journal of Peace Research*, 39(4): 387– 393.

Murshed, S. Mansoob and Scott Gates (2005). "Spatial-Horizontal Inequality and the Maoist Insurgency in Nepal". *Review of Development Economics*, 9(1): 121– 134.

Murshed, S. Mansoob and Tadjoeddin, Mohammad Zulfan (2009). "Revisiting the Greed and Grievance Explanations for Violent Internal Conflict". *Journal of International Development*, 21(1): 87– 111.

Nordås, Ragnhild (2012). "The Devil in the Demography? Religion, Identity, and War in Côte d'Ivorie" in Jack A. Goldstone, Erik P. Kaufmann, and Monica Duffy Toft (ed.),

Political Demography. How Population Changes Are Reshaping International Security and National Politics. Boulder, Colorado: Paradigm Publishers.

Olson, Mancur (1965). *The Logic of Collective Action. Public Goods and the Theory of Groups.* Cambridge, Mass: Harvard University Press.

Pevhouse, Jon C. and Jason D. Brozek (2008). "Time- Series Analysis" in Janet M. Box-Steffenmeier, Henry E. Brady, and David Collier (ed.). *The Oxford Handbook of Political Methodology*. Oxford: Oxford University Press.

Polity IV Project (2011). "Polity IV Project: Political Regime Characteristics and Transitions, 1800–2010", downloaded January 2013 from http://www.systemicpeace.org/polity/polity4.htm .

Regan, Patrick M., and Daniel Norton (2005). "Greed, Grievance, and Mobilization in Civil Wars". *Journal of Conflict Resolution,* 49(3): 319– 336.

Ross, Michael. (2004). "What do we know about natural resources and civil war?". *Journal of Peace Resource*, 41(3): 337–56.

Ross, Michael (2006). "A Closer Look at Oil, Diamonds, and Civil War". *Annual Review of Political Science,* 9: 265–300.

Sambanis, Nicholas (2001). "Do Ethnic and Nonethnic Civil Wars Have the Same Causes? A Theoretical and Empirical Inquiry (Part 1)". *Journal of Conflict Resolution*, 45(2): 259– 282.

Sambanis N (2004). "Poverty and the organization of political violence: A review and some conjectures". Working paper. New Haven, CT: Yale University.

Sen, Amartya Kuman (1973). *On Economic Inequality*. New York: Oxford University Press.

Smith, Amy Erica (2009). "Legitimate Grievances: Preferences for Democracy, System Support, and Political Participation in Bolivia". *Latin American Research Review*¸44(2): 102–126.

Soysa, Indra De, and Nils Petter Gleditsch (1999). *To Cultivate Peace: Agricultire in a World of Conflict,* (PRIO report). Oslo: International Peace Research Institute.

Soysa, Indra De (2002). "Paradise is a Bazaar? Greed, Creed, and Governance in Civil War, 1989–99". *Journal of Peace Research,* 39(4): 395–416.

Stouffer, Samuel A., Edward A. Suchman, Leland De Vinney, Shirley A. Starr, and Robin M. Williams, Jr. (1949). *The American Soldier: Adjustment to army life (Vol. 1).* Princeton: Princeton University Press.

Taylor, Charles L. and Michael C. Hudson (1972). *World Handbook of Political and Social Indicators*. Ann Arbor, MI: ICPSR.

Theisen, Ole Magnus (2012). *Renewable Resources Scarcity, Natural Disasters, and the Possibility of Collective Violence* (Doctoral theses, NTNU). Trondheim: NTNU.

Tilly, Charles and Sidney Tarrow (2007). *Contentious Politics*. London: Paradigm Publisher.

Themnén, Lotta and Petter Wallensteen (2012). "Armed Conflicts, 1946–2011". *Journal of Peace Research*, 19(4): 565– 575.

Toft, Monica Duffy (2002). "Invisible territory, geographic concentration, and ethnic war". *Security Studies*, 12(2): 82–119.

Ulleberg, Pål and Hilmar Nordvik (2001). *Faktoranalyse*. Trondheim: Tapir.

Vanhanen, Tatu (1999). "Domestic Ethnic Conflict and Ethnic Nepotism: A Comparative Analysis". *Journal of Peace Research*, 36(1):55–73.

Walter, Barbara F. (2004). “Does Conflict Beget Conflict? Explaining Recurring Civil War”. *Journal of Peace Research*, 41(3): 371–388.

Weinstein, Jeremy M. (2005). “Resources and the Information Problem in Rebel Recruitment”. *Journal of Conflict Resolution,* 29(4): 598–624.

Williams, Robert (1975). "Relative Deprivation", in Lewis A. Coser (ed.), *The Idea of Social Structure: Paper in Honor of Robert L. Merton*. New York: Harcourt Brace Jovanovich.

Wolff, Stefan (2006). *Ethnic Conflict. A Global Perspective*. New York: Oxford University Press.

World Bank (2013a). *Population*. Downloaded February 2013, from http://data.worldbank.org/indicator/SP.POP.TOTL.

World Bank (2013b). *Population Growth*. Downloaded February 2013, from http://data.worldbank.org/indicator/SP.POP.GROW.

World Bank (2013c). *Military Expenditure*. Downloaded February 2013, from http://data.worldbank.org/indicator/MS.MIL.XPND.GD.ZS.

World Bank (2013d). *GDP per Capita*. Downloaded February 2013, from http://data.worldbank.org/indicator/NY.GDP.PCAP.CD.

World Bank (2013e). *Fuel Exports*. Downloaded February 2013, from http://data.worldbank.org/indicator/TX.VAL.FUEL.ZS.UN.

World Bank (2013f). *Agriculture*. Downloaded February 2013, from http://data.worldbank.org/indicator/NV.AGR.TOTL.ZS.

World Bank (2013g). *Life Expectancy at birth*. Downloaded February 2013, from http://data.worldbank.org/indicator/SP.DYN.LE00.IN.

World Bank (2013h). *Rural Population*. Downloaded February 2013, from http://data.worldbank.org/indicator/SP.RUR.TOTL.ZS.

World Bank (2013i). *Public spending on education*. Downloaded February 2013, from http://data.worldbank.org/indicator/SE.XPD.TOTL.GD.ZS.

Østby, Gudrun (2008). "Polarization, Horizontal Inequalities and Violent Civil Conflict". *Journal of Peace Reserach*, 45(2): 143–62.

Østby, Gudrun (2011). *Horizontal Inequalities and Political Violence* (Doctoral theses, Universitetet i Oslo). Oslo: Universitetet i Oslo.

Appendix A

Tables excluded from the analysis section

Table A1. Binary regression, coefficient (st.error), civil conflict onset

Variables	
EthfracIpolated	4.695***
	(1.623)
EthfracIpolated^2	-3.003*
	(1.592)
Lagged Incidence	0.879***
	(0.313)
Constant	-4.985***
	(0.378)
N=4263	

Note: ***significant at 1%, **significant at 5%, and *significant at 10%. The units are clustered by country using Huber-White robust standard errors (White, 1980).

Table A2. Logitsic Regression, coefficient (st.error), civil conflict/war, lag incidence

Variables	Model 1 Conflict	Model 2 Conflict	Model 4 War	Model 5 War
Population❖¤	0.315***	0.369***	0.105	0.130
	(0.114)	(0.108)	(0.139)	(0.155)
GDP per capita❖¤	-0.370***	-0.462***	-0.681***	-0.731***
	(0.143)	(0.148)	(0.216)	(0.206)
Fuel exporter❖	0.857**	0.876***	0.816	0.828
	(0.343)	(0.335)	(0.617)	(0.615)
Mountain	0.005	0.002	0.017*	0.015
	(0.007)	(0.006)	(0.009)	(0.010)
Authocracy❖[a]	-0.037	0.312	-0.613	-0.605
	(0.324)	(0.318)	(0.493)	(0.499)
Democracy❖[a]	0.609	0.551	-1.127*	-1.440*
	(0.408)	(0.373)	(0.839)	(0.832)
Empowerment rights❖	-0.060	-0.059	-0.082	-0.085
	(0.037)	(0.037)	(0.072)	(0.070)
EthfracIpolated❖	1.396***	8.372***	-0.180	2.322
	(0.497)	(2.402)	(0.879)	(4.590)
EthfracIpolated^2 ❖		-7.277***		-2.756
		(2.349)		(5.144)
Temporal dependence				
Lag Incidence	0.382	0.248	-1.218	-1.260
	(0.322)	(0.321)	(1.155)	(1.143)
Constant	-6.923***	-8.315***	-0.951	-1.340
	(2.161)	(2.137)	(3.386)	(3.651)
Pesudo LL	-331.504	-326.648	-140.699	-140.407
Observations	2734	2734	2754	2754
Years	1985-2010	1985-2010	1985-2010	1985-2010

Note:❖Lagged one year, ¤Log transformed, [a] Anocracy as reference category,
[b] Probability of F-value, *significant at 10%, **significant at 5%, and ***significant at 1%. The units are clustered by country using Huber-White robust standard errors (White, 1980).

Appendix B

Descriptive statistics

Table B1. Frequences, civil conflict onset

Onset	Frequency	Percent
0	4,141	96.62
1	145	3.38
Total	4,286	100

Table B2. Frequences civil war onset

Onset	Frequency	Percent
0	4,2417	98.81
1	51	1.19
Total	4,298	100

Table B3. Frequences dumy variable based on POLITY2

Categories	Frequency	Percent
Autocracy	1,097	27.96
Anocracy (ref)	1,192	30.38
Democracy	1,634	41.65
Total	3,923	100

Table B4. Frequences dummy variable fuel export

Categories	Frequency	Percent
0 (<1/3)*	3,086	83.02
1 (>1/3)*	631	16.98
Total	3,717	100

*of total merchandise export

Table B5. Descriptive statistics

Variables	*N*	Min	Max	Mean	Std. Dev.
GDP per capita¤	4122	4.164	11.627	7.677	1.619
Agriculture	3767	0.035	96.577	17.557	15.529
Bureaucratic quality	3314	0	4	2.148	1.183
Life expectancy	4271	26.819	82.933	65.535	10.816
Rural population	4276	0	94.943	47.568	23.596
Law and order	3314	0	6	3.682	1.492
Educational spending	2052	0	16.059	4.519	1.836
Military	3112	0.046	117.388	2.693	3.504
Polity	3923	-10	10	2.210	7.052
Empowerment rights	3905	0	14	8.238	4.248
Population¤	4276	12.021	21.014	15.871	1.679
Pgysical integrity rights	3915	0	8	4.800	2.297
Ethfracipolated	4292	0.001	0.960	0.431	0.261
Ethfracwar	4266	0.14	89	0.732	2.385
Population growth	4274	-7.533	18.588	1.670	1.538
Moutn	3682	0	94.3	17.485	21.837

Note: ¤ Log transformed

Appendix C

Graphs, figures, and explanations

Figure C1. Screeplot, Table 3

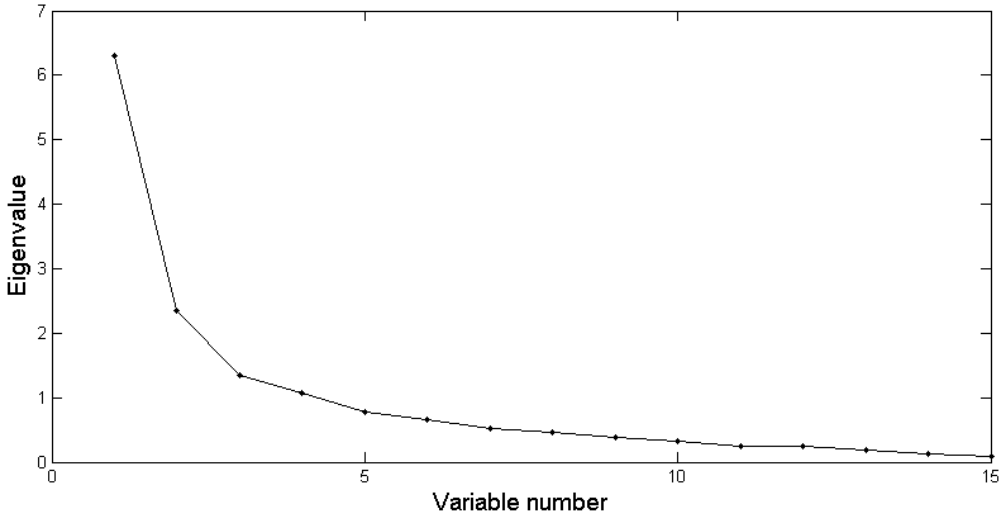
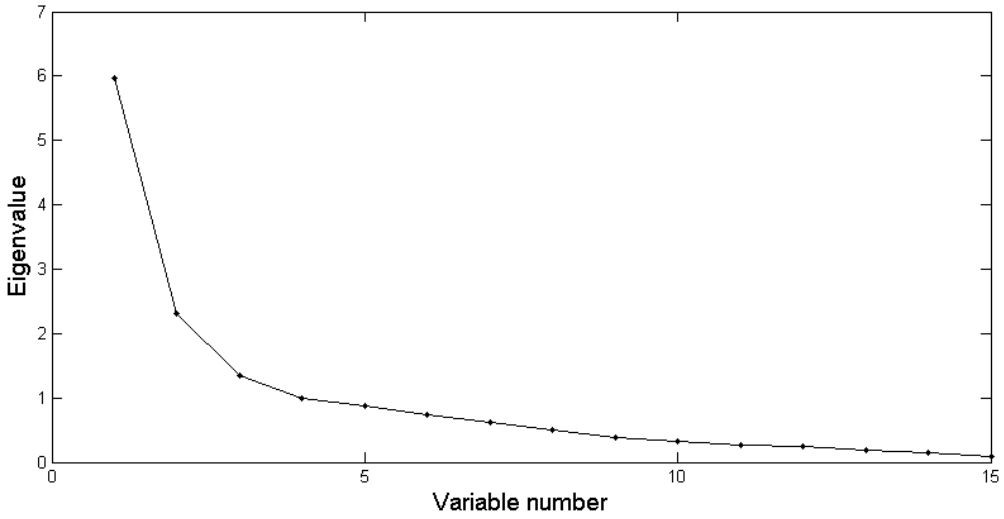


Figure C2. Screeplot, Table 4



C3. Natural Cubic Splines

The reason for using for using a natural cubic spline with three knots to control for temporal dependence is that there is no theoretical reason for assuming that there is a linear impact of time on the probability of conflict. The same assumption is done for the PEACE YEARS

variable included in the analysis, and why I chose to a natural cubic spline with three knots in the thesis, only including the regression analysis with lagged incidence variable in the appendix as a robust check. Natural cubic spines fit cubic polynomials to a predetermined number of subintervals of variable. These polynomials are joined as "knots," and the number and placement of the knots specified by the analyst. Forcing the splines, and their first and second derivatives, to agree at each of the knots, enforces smoothness. Each knot only uses up one degree of freedom, so that we can flexibly fit a cubic spline using up only a very few degrees of freedom. The estimated spline coefficients can then be applied to trace out the path of duration dependence (Beck et al. 1998: 1270).